

## Portfolioevaluierungen: Herausforderungen und aktuelle Erfahrungen

Frühjahrstreffen des Arbeitskreises Forschungs-, Technologie-  
und Innovationspolitik der Gesellschaft für Evaluation e.V.  
am 11. Juni 2013, Wien

*Iris Fischl,<sup>1</sup> Marianne Kulicke,<sup>2</sup> Jan Wessels<sup>3</sup>*

Evaluationen in der Vergangenheit waren – mit wenigen Ausnahmen – zumeist auf einzelne Programme fokussiert. Zunehmend jedoch gibt es bei Auftraggebern von Evaluationen ein Interesse ‚über den Tellerrand‘ einer einzelnen Maßnahme hinauszublicken bzw. die Wirkungen eines Bündels an Maßnahmen gesamthaft zu bewerten. Zwar zählt zu Evaluationen von Einzelprogrammen oft auch die Bewertung des Programmumfelds (i.S. von ähnlichen Programmen, Komplementarität von Bundes- und Landesprogrammen), jedoch steht sie meist nicht im Zentrum, sondern ist ein Zusatzmodul.

In Österreich beziehen sich in jüngster Vergangenheit immer mehr Evaluationen auf ein definiertes Portfolio an Maßnahmen bzw. auf definierte Akteure und deren Förderungen. Eine solche Entwicklung lässt sich für Deutschland nicht feststellen. Dies dürfte – legt man die Größe der Länder zugrunde – aus einer höheren Anzahl und Bandbreite an Fördermaßnahmen in Österreich resultieren, während in Deutschland zumindest im Bereich der Innovationsförderung zuletzt das Ziel einer Reduktion der Programmvierfalt durch Zusammenlegen komplementärer Programme zu einem Programm (Zentrales Innovationsprogramm Mittelstand, ZIM) oder Wegfall kleinerer Maßnahmen verfolgt wurde.

Ausgangspunkt der Definition eines Portfolios aus einem Bündel mehrerer Maßnahmen können dabei Themen oder Branchen sein, aber auch Förderungsinstrumente oder -agenturen oder – wie in Deutschland – einzelne Regionen. Maßnahmenportfolios sind zumeist historisch gewachsene Strukturen, wodurch sich auch ein zunehmender Bedarf an der Evaluation ganzer Maßnahmenbündel bzw. Maßnahmenportfolios ergibt.

Vor diesem Hintergrund führte der Arbeitskreis Forschungs-, Technologie- und Innovationspolitik der DeGEval am 11. Juni 2013 in Wien sein Frühjahrstreffen zum Thema „Portfolioevaluierungen: Herausforderungen und aktuelle Erfahrungen“ durch. Über 30 Evaluierende, Auftraggeber, Programmmanager/-innen bzw.

---

1 KMU FORSCHUNG AUSTRIA, Wien

2 Fraunhofer Institut für System- und Innovationsforschung (ISI), Karlsruhe

3 Institut für Innovation und Technik (iit) in der VDI/VDE Innovation + Technik GmbH, Berlin

-träger/-innen diskutierten Erfahrungen aus aktuellen Portfolioevaluierungen und erhielten einen Überblick über Herausforderungen, methodische Möglichkeiten, Nutzen und Lerneffekte von solchen Analysen. Das Treffen wurde von der KMU Forschung Austria (KMFA) organisiert und fand bei der Wirtschaftskammer Österreich (WKO) statt.

Nach Einführung durch Iris Fischl (KMFA) berichteten zunächst Evaluierende aus Österreich und Deutschland in vier Vorträgen über aktuelle Erfahrungen in diesem Feld.

Der erste Beitrag von Iris Fischl (KMFA) bezog sich auf eine Portfolioevaluation auf regionaler Ebene (Monetäre Unternehmensförderung der Stadt Wien). Im Mittelpunkt stand der Umgang mit einem Maßnahmenportfolio als Evaluationsgegenstand, das sich aus sehr unterschiedlichen monetären Unternehmensförderungen – verankert in vier Förderagenturen – zusammensetzte. Thema war der Umgang mit der Diskrepanz zwischen Einzelprogramm- und Portfolioperspektive, handelte es sich doch um eine Portfolioevaluation, die auch Analysen auf Einzelprogrammebene anstellte. Iris Fischl stellte zwei von der KMFA entwickelte „Positionierungstools“ vor, die das Evaluierungsteam maßgeblich bei der schrittweisen Überprüfung der Konsistenz der Programmziele, der Identifizierung von Komplementaritäten und Redundanzen im gesamten Maßnahmenportfolio sowie der Diskussion mit dem Auftraggeber und der Steuerungsgruppe der Evaluation unterstützten.

Der Vortrag von Helmut Gassler, Joanneum Research, bezog sich auf eine Studie zur Bewertung des Gesamtportfolios der aws (austria wirtschaftsservice), Technologieprogramme mit Prüfung der Synergien untereinander sowie Bewertung der Funktion und Positionierung im gesamten österreichischen Fördersystem. Neben der Portfoliobetrachtung sollte auch eine Bewertung der 14 Einzelmaßnahmen (in vier Programmfamilien) erfolgen. Auch er betonte die Notwendigkeit einer Strukturierung der zu evaluierenden Maßnahmen als Ausgang einer Portfoliobetrachtung. In der Studie erfolgte eine Analyse der Zielsysteme pro Programm und Familie, um deren Kernziele zu identifizieren und vergleichbar zu machen. Nach Beseitigung von Redundanzen in Zielformulierungen und Zielklärung mit der aws bei unklaren Formulierungen konnten Komplementaritäten oder Überschneidungen in und zwischen Programmfamilien und Einzelmaßnahmen identifiziert werden. Im nächsten Schritt wurden Komplementaritäten in der Inanspruchnahme der aws-Programme aus Nutzersicht herausgearbeitet (Basis: Sonderauswertung der Unternehmensbefragung aus der Systemevaluierung). Dies erfolgte mittels einer Netzwerkanalyse (i.S. einer Vernetzung der Programme untereinander), um häufige aws-interne Kombinationen zu identifizieren. Diese Methodik kam auch für die Bewertung der Positionierung der aws-Technologieprogramme im Innovationssystem zum Einsatz, um aus Nutzerperspektive häufige Kombinationen von aws- und anderen bundesweiten Förderangeboten sowie Synergien und Überschneidungen aufzuzeigen.

Herausforderung der Studie war der fehlende Zugang zu Daten der aws über die Fördernehmer, ein nicht nur für Portfolioevaluierungen typisches Problem. In der Diskussion wurde das Bild des „Fördersupermarkts“ als Synonym für die Breite und leichte Zugänglichkeit zu Förderangeboten in Österreich geprägt. Die aws-Technologieprogramme sind dagegen sehr kleinteilig angelegt („Fachhandel“ für

spezifische Zielgruppen), die die finanzielle Komponente mit begleitenden Betreuungsangeboten kombiniert.

Dr. Heike Belitz, Deutsches Institut für Wirtschaftsforschung (DIW), berichtete über Vorgehensweise und Ergebnisse einer Studie zur gesamtwirtschaftlichen Bewertung der Innovationsförderung des Bundes in Deutschland für den Mittelstand. Kern war die Gegenüberstellung technologieoffener und technologiespezifischer Förderansätze in ihrer Inanspruchnahme und Wirkung. Betrachtet wurde ein nach Instrumenten und Adressaten heterogenes Portfolio: Programme für FuE-Projekte, Forschungsinfrastrukturen, Beratungen und Dienstleistungen sowie Management-Netzwerke. Neben sekundärstatistischen Daten bildeten eine umfangreiche Primärdatenerhebung bei Fördernehmern und Expertengespräche die Informationsbasis. Eine Erkenntnis für Portfolioevaluierungen aus dieser Studie ist die Notwendigkeit der Gruppenbildung von Förderadressaten mit einer Typisierung nach Programmnutzung, Innovationsaktivitäten/-strategien und Größe, um Aussagen zu Wirkungen einzelner Maßnahmen treffen zu können. Nach Heike Belitz ergibt sich für solche Evaluierungen die Notwendigkeit, vorhandene Einzelprogrammevaluationen in den Förderkontext einzubetten, die Ziele der Einzelprogramme für eine Messbarkeit zu konkretisieren und die Wirkungen weiterer Rahmenbedingungen neben der Förderung auf die Zielerreichung der Förderung zu berücksichtigen. Sie betonte, dass Portfolioevaluierungen gerade für Fördergeber neue Perspektiven eröffnen (Betrachtung geförderter KMU statt einzelner Projekte) und den Einsatz neuer methodischer Ansätze ermöglichen.

Leonhard Jörg skizzierte in seinem Beitrag aus Sicht einer Förderagentur (Österreichische Forschungsförderungsgesellschaft, FFG) den Prozess der Strukturierung einer breiten Palette an Förderangeboten als Portfolio. Die Grundvoraussetzung für eine Portfolioevaluation erscheint zunächst trivial: Man benötigt innerhalb einer Agentur die Strukturierung einer Vielzahl von Maßnahmen als Portfolio, um diese in ihrem Zusammenspiel und ihrer Wirksamkeit zu bewerten. Diesen Prozess der Definition eines Portfolios hat die FFG beschritten, wobei vier Problembereiche eine gesamthafte Bewertung ihres Förderangebots erforderten:

- Steuerung (Grenzen der Steuerungshebel auf Einzelprogrammebene durch Vielfalt),
- Legitimation (Auseinanderfallen von Wirkungsanspruch und zuordenbarer Wirkung),
- Wettbewerb (Beeinträchtigung der Wirksamkeit durch Wettbewerb zwischen Programmen),
- Allokation (Nebeneinander von Programmen mit großen Schnittmengen nach Zielen und Adressaten).

Die FFG (einziger bundesweiter Projektträger in Österreich) bietet in 27 Förderlinien 59 meist themenspezifische Programme an, die aber thematisch keineswegs überschneidungsfrei sind. Ein Strukturierungsschritt war die Analyse der Themenadressierung nach Bereichen und ihres Förderumfangs (z.B. wie viele Mittel flossen in den Bereich Produktion aus einzelnen FFG-Programmen? Welches Gewicht haben produktionsspezifische gegenüber themenoffenen Programmen?). Der Prozess der Strukturierung umfasste auch die Definition organisatorischer Regelungen

und Maßnahmen für ein in sich stimmiges Portfolio. Ein Baustein bildet dabei die Standardisierung mit vier Instrumenten und deren Ausrichtung auf den definierten Zielkorridor: Themenoffene Projektförderung, Förderung von Strukturen und Netzwerken, von Humanressourcen sowie von Themenschwerpunkten. Ferner zählt dazu die Definition des operativen Portfoliomanagements, wofür Thementeam für die wichtigsten sechs Themen gebildet wurden.

Eine Evaluation des so als Portfolio strukturierten Förderangebots der FFG sollte folgende Aspekte bewerten: Wahl der eingesetzten Instrumente, Funktionieren der instrumentenübergreifenden Steuerung und des Zusammenspiels zwischen Instrumenten in der Abwicklung, Umsetzbarkeit einer Portfolio- gegenüber der Einzelprogrammperspektive, Wirksamkeit des implementierten Portfoliomanagements in der Praxis sowie Mehrwert des Portfolios aus Kundenperspektive.

Die Strukturierung als Portfolio ist noch in der Umsetzung, ohne dass in der Förderpraxis bereits Änderungen oder organisatorische Modifikationen in größerem Umfang erfolgten.

In der abschließenden Plenumsdiskussion führten unter der Moderation von Thomas Jud die Diskutanten Dr. Johannes Gadner, Dr. Michael Wiesmüller, Leonhard Jörg und Helmut Gassler zentrale Erfahrungen zusammen. Als wichtige Motivation zur Portfoliobildung wurde von Seiten der Fördergeber vor allem der vermutete Effizienzgewinn für die Technologie- und Innovationspolitik angeführt. Nicht alle zu evaluierenden Portfolios sind aber als solche gezielt aufgebaut, nicht selten werden historisch gewachsene De-facto-Portfolios evaluiert. Als Herausforderungen nannten die Evaluierenden primär Schwierigkeiten bei der Datenverfügbarkeit, außerdem zu knapp bemessene Projektlaufzeiten für hoch komplexe Maßnahmen- und Wirkungsgefüge. In diesem Kontext wurde auch auf die Vorteile begleitender Evaluationen (längere Beschäftigung und Durchdringung des Untersuchungsgegenstandes) hingewiesen. Als weitere Herausforderung betonten die Diskutanten die Verteilung der Verantwortlichkeiten für unterschiedliche Portfolioelemente auf mehrere Auftraggeber, die jeweils individuell mögliche Empfehlungen der Evaluation umsetzen können. Schließlich wurde ergänzend zum diskutierten Portfoliobegriff auch ein Portfolioverständnis eingebracht, welches – ausgehend insbesondere von europäischen Fördermaßnahmen – stärker auf der Projektebene unterschiedliche Projektschwerpunkte und -ergebnisse als Portfolio eines Programms versteht.

## Komplexität und Evaluation von Programmen der beruflichen Bildung

Frühjahrstagung des Arbeitskreises Berufliche Bildung im April 2013

*Michael Kalman,<sup>1</sup> Jürgen Götzhaber<sup>2</sup>*

Am 19.04.2013 fand die Frühjahrstagung 2013 des Arbeitskreises Berufliche Bildung im Bundesinstitut für Berufsbildung (BIBB), Bonn statt. Rund 20 Expertinnen und Experten tauschten sich zum Thema „Komplexität und Evaluation von Programmen der beruflichen Bildung“ aus.

Die Frühjahrstagung griff dabei ganz bewusst das Motto ‚Komplexität und Evaluation‘ der DeGEval-Jahrestagung 2013 auf. Mit Blick auf Programme der beruflichen Bildung wurde beleuchtet, wie sich Komplexität aus unterschiedlichen Perspektiven darstellt, in welchen Dimensionen und Faktoren sie sich manifestiert und welchen Herausforderungen sich Auftragnehmende und Auftraggebende gegenübersehen. Zunächst beleuchtete *Christoph Acker* das Thema aus der Sicht eines Auftraggebenden einer Evaluation anhand des Programms ‚Jobstarter Connect‘. Herr Acker ist Leiter dieses Programms in der Programmstelle JOBSTARTER beim Bundesinstitut für Berufsbildung (BIBB).

Bereits das Programm ‚Jobstarter Connect‘ selbst, der Evaluationsgegenstand, ist komplex – so der Tenor seiner Präsentation mit dem Titel „Programme, Projekte und Politik: Komplexe Anforderungen an die Evaluation“. Die Grundidee ist, dass mit bundeseinheitlichen Ausbildungsbausteinen Jugendlichen und jungen Erwachsenen der Übergang in die duale Berufsausbildung erleichtert werden soll. Wie Acker ausführte, sind diese Ausbildungsbausteine in ihrer Lernergebnisorientierung und Kompetenzbasierung an der Ganzheitlichkeit des Berufsbildes und am Konzept der beruflichen Handlungsfähigkeit orientiert und beruhen auf den geltenden Ausbildungsrahmenplänen und Rahmenlehrplänen. Sie werden eingesetzt in bestehenden Maßnahmen der Berufsvorbereitung, der geförderten außerbetrieblichen Ausbildung und der Nachqualifizierung. Damit ist eine enorme Vielfalt angesprochen, die Acker in einer anschaulichen Grafik seiner Präsentation durch 16 unterschiedliche Maßnahmentypen visualisierte. Zugleich unterscheiden sich die regionalen Kontexte, in denen die Ausbildungsbausteine umgesetzt werden. Acker nannte als

---

1 Mitglied Sprecherteam des AK Berufliche Bildung, Kalman Consult, Berlin

2 Mitglied Sprecherteam des AK Berufliche Bildung, Evaluelab GmbH, Hamburg

weitere Komplexitätsmerkmale die unterschiedlichen Durchführungsstellen der Projekte und divergierende regionale sowie bundesweite Akzeptanzen und Rechtsauffassungen zu den Ausbildungsbausteinen. Schließlich sind Parallelen und Überschneidungen zu weiteren Programmen wie die modulare Nachqualifizierung der ‚Perspektive Berufsabschluss‘ oder ‚DECVET – Entwicklung eines Leistungspunktesystems in der beruflichen Bildung‘ unverkennbar. Zur Komplexität tragen auch die zahlreichen Zielsetzungen von ‚Jobstarter Connect‘ bei. Hierzu gehören z.B. so unterschiedliche Ziele wie „Die Verbesserung des Übergangs von jugendlichen Altbewerberinnen und Altbewerbern in eine duale Ausbildung“ oder „Eine bessere Verzahnung bestehender Teilbereiche des Berufsbildungssystems“ (siehe 2. Förderrichtlinien zur Durchführung des Programms JOBSTARTER – CONNECT vom 25. Juni 2009).

So sieht Acker mit Blick auf die Evaluation mehrere „Herausforderungen der Wirkungsanalyse“ – in komplexen Wirkkonstellationen ist es schwierig, die Ursachen von Wirkungen klar zuzuschreiben:

- Sind es die Ausbildungsbausteine selbst?
- Die sie rahmenden Maßnahmen?
- Die günstige oder ungünstige regionale Marktsituation?
- Die Professionalität des Projektpersonals? etc.

Der beauftragten Evaluation kommt jedenfalls im ‚lernenden System‘ des Erprobungsprogramms ‚Jobstarter Connect‘ eine wichtige Funktion zu. So sollen Effekte der Einführung des Instruments Ausbildungsbausteine auf das Übergangssystem hinsichtlich unterschiedlicher Dimensionen ermittelt und Empfehlungen zur Programmumsetzung gegeben werden. Eine „systematische Rückkopplung von Evaluationsergebnissen in den Erprobungsprozess“ wird erwartet.

Perspektivenwechsel: Die entsprechende Auftragnehmendensicht wurde in der Folge von *Stefan Ekert*, Geschäftsführer der InterVal GmbH, Berlin, am Beispiel der Evaluation des Programms ‚Jobstarter Connect‘ präsentiert. Ekert machte in seiner Präsentation „Externe Evaluation von ‚JOBSTARTER CONNECT‘ – eine komplexe Mehrebenen-Evaluation“ deutlich, dass komplexe Anforderungen durch ein komplexes Evaluationsdesign zu adressieren sind, welches zugleich die Vielschichtigkeit des Evaluationsgegenstandes zu berücksichtigen hat.

Ekert machte die Komplexität des Evaluationsgegenstandes an drei Punkten fest:

- Heterogenität der Projekte,
- Programmebene mit Programmphasen und -umfeld,
- Herausforderungen der Implementierung von zertifizierten Ausbildungsbausteinen in das bestehende Berufsbildungssystem.

Die Ausbildungsbausteine werden in unterschiedlichsten Projekten, die sich hinsichtlich ihres regionalen Umfelds, der Lernorte, der beteiligten Akteure, der Zielgruppen usw. unterscheiden, umgesetzt. Da das Programm einen Erprobungscharakter hat, ist diese Heterogenität und in der Folge diese Komplexität erwünscht, weil sich die Programmstelle daraus entsprechend vielfältige Erkenntnisse verspricht.

Die Programmarchitektur selbst fügt weitere Komplexitätsmerkmale hinzu. So wirken zahlreiche Programmakteure mit, angefangen beim Fördermittelgeber Bundesministerium für Bildung und Forschung (BMBF), der Programmstelle beim BIBB über den Begleitausschuss, weitere Expertinnen und Experten bis hin zu den zahlreichen Projektakteuren. Die unterschiedlichsten Funktionen sind zu koordinieren, wie Vergabe, Kontrolle, fachliche Begleitung, Erfahrungsaustausch, Monitoring, Diskurs – und Evaluation. Das auftragnehmende Evaluationsteam soll nicht nur nach dem Stand der Technik evaluieren, sondern auch Kommunikationsprozesse analysieren und fördern, beraten und den Input von Ergebnissen in Reflexionsprozesse gewährleisten. Daneben ist das Programm in unterschiedliche Phasen eingeteilt – dies eine weitere Dimension von Komplexität, nämlich Vorerfahrungen, Programmstart, Erprobung, Reflexion und Optimierung, die von vielfältigen fachpolitischen Rahmenbedingungen und fachwissenschaftlichen Diskursen gerahmt werden. Schließlich sollen die erprobten Ausbildungsbausteine in das bestehende Berufsbildungssystem implementiert werden und den Rang anrechenbarer, prüf- und zertifizierungsfähiger Module erhalten – was durchaus kontrovers diskutiert wird.

Wie gehen die Evaluierenden mit diesen Komplexitäts-Herausforderungen um? Die Antwort von Ekert lässt sich zu drei Aspekten verdichten, nämlich Kommunikation nach innen und außen, Klassifizierung in einer Mehrebenenbetrachtung und Anwendung eines Methodenmix. Die Ergebnisse der Untersuchung werden dann synoptisch den verschiedenen Ebenen zugeordnet.

Der Kommunikation kommt in Evaluationsvorhaben naturgemäß eine bedeutende Rolle zu. Dabei geht es auch um die Verständigung darüber, was eine Evaluation kann und was nicht. Angesichts der Komplexität der Evaluationsgegenstände sind schnelle und einfache Antworten nicht zu erwarten. Die Bildung unterschiedlicher Ebenen entspricht der Ordnungsfunktion von Evaluationsdesigns. So identifiziert Ekert vier Ebenen: das ‚reine‘ Konzept, die konkret vorliegenden Ausbildungsbausteine, das Programm und die Projekte. Schließlich verspricht sich Ekert durch die „Nutzung vielfältiger empirischer Zugänge zur Erfassung dessen, was vor Ort und im Umfeld passiert“ im Rahmen der Anwendung eines Methoden-Mix valide Ergebnisse, die über eine fachliche Bewertung den verschiedenen Ebenen mittels Synopse zugeordnet werden.

In einem weiteren Referat wurde das Tagungsthema durch *Peter Jablonka*, Geschäftsführer der Forschungsgruppe SALSS (Berlin/Bonn), im Zusammenhang mit der wissenschaftlichen Begleitung des Modellversuchsprogramms ‚Neue Wege in die duale Ausbildung – Heterogenität als Chance für die Fachkräftesicherung‘ behandelt. Jablonka machte deutlich, dass sich bei der wissenschaftlichen Begleitung Komplexitätsgesichtspunkte noch einmal anders darstellen als bei dezidierten Evaluationen, weil häufig die Evaluation eingebettet ist in einen vielfältigen Aufgabenkranz, der auch Unterstützungsleistungen für laufende Prozesse der Programm- bzw. Projektdurchführung beinhaltet.

Das Modellversuchsprogramm hat sich zum Ziel gesetzt, „innovative Wege in die Ausbildung und eine spätere Berufstätigkeit aufzuzeigen“ und dabei „die zunehmende Heterogenität der Jugendlichen als Herausforderung und Chance“ zu begreifen (siehe die entsprechende Informationsbroschüre des BIBB, S. 3). Auch hier

zeigt sich ein weiteres Mal, dass der Evaluationsgegenstand – das Modellversuchsprogramm mit seinen diversen Projekten – komplex ist. Jablonka macht dies an den zahlreichen Stakeholdern und Betroffenen fest. Auf der Programmebene agieren die Programmsteuerung und Fachwissenschaftliche Betreuung des BIBB sowie die wissenschaftliche Begleitung bestehend aus Forschungsgruppe SALSS und den Kooperationspartnern Universität Magdeburg und dem Institut für sozialwissenschaftliche Beratung (isob). Die Aufgaben der wissenschaftlichen Begleitung sind vielfältig und reichen von der prozessbegleitenden Unterstützung der Projektakteure bei Selbstevaluation, Monitoring und einheitlichen Berichtsformaten bis hin zu Erkenntnis- und Ergebnistransfer in Politik, Berufsbildungspraxis und Berufsbildungsforschung. Schließlich die Projektebene: Hier wurden 17 verschiedene Vorhaben ausgewählt, die von unterschiedlichsten Akteuren in einer breiten regionalen Streuung gestaltet werden. Die Themen dieser Projekte könnten unterschiedlicher nicht sein – angefangen bei „Chance Pflegeberuf“ – Unterstützung und gezielte Vorbereitung von noch nicht ausbildungsreifen Jugendlichen für die Altenpflegeausbildung“ (Träger: AWO Landesverband Saarland e.V.) bis „S’CoolWiki: Förderung der Fach-, Medien- und Sozialkompetenz über Web 2.0 im Elektro- und IT-Handwerk“ (Träger: Elektro Technologie Zentrum Stuttgart (etz). Die wissenschaftliche Begleitung bedient sich in ihrer Arbeit sowohl formativer als auch summativer Evaluationsverfahren – auch dies verdichtet die Komplexität weiter.

Jablonka arbeitete in seiner Präsentation drei weitere Aspekte der Komplexität heraus:

- „Heterogenität“ als Gegenstand des Programms (und der wissenschaftlichen Begleitung/Evaluation): Heterogene (Lern-)Gruppen weisen einen Doppelcharakter zwischen Herausforderung und Chance auf. Nicht zuletzt für die Lehr- und Fachkräfte kann es schwierig sein, ihr didaktisches Konzept in uneinheitlichen Klassen zu realisieren, da die Unterschiede zwischen den Teilnehmenden weit größer sind als in homogenen Lerngruppen. Zugleich wird Heterogenität jedoch, wenn sie erfolgreich gemanagt werden kann, als Chance begriffen.
- Die Intervention: Vielfalt der Aktivitäten der Modellversuche. Die Vielfalt von Instrumenten und Handlungsansätzen in den verschiedenen Vorhaben des Programms erscheint als „ungeordnete Komplexität“, die durch Generalisierung zu acht, später zu vier Handlungsfeldern gebannt wurde. Dabei wurde das Instrument Synopse gewählt, um die verschiedenen Projekte den Handlungsfeldern gegenüberzustellen. Am Beispiel des Handlungsfeldes „Ausbildungsmanagement“ demonstrierte Jablonka, wie die Systematisierung als Ausgangsbasis für die nächsten Schritte genutzt wurde. Das Handlungsfeld wurde mit Unterpunkten konkretisiert, denen wiederum konkrete Handlungsschritte zugeordnet wurden. Damit erhielten die verschiedenen Projektakteure konkrete Handlungsempfehlungen, die nunmehr in einer „geordneten Komplexität“ erscheinen. Zur weiteren Unterstützung der Projekte wurde ein Zielsystem entwickelt, welches den Besonderheiten der regionalen Vorhaben gerecht wird, zugleich aber auch einen Gesamtüberblick über die jeweiligen Zielerreichungsgrade ermöglicht.
- Verstetigung und Transfer: Hier hat die wissenschaftliche Begleitung eine ganze Reihe von Produkten zu erstellen wie Handreichungen für die Praxis (KMU,

Bildungsdienstleister und Schulen), Empfehlungen an die Politik (optimierte Nutzung von vorhandenen Förderinstrumenten, Optimierung von Förderinstrumenten) und Beiträge zum wissenschaftlichen Diskurs (Heterogenität in den Bildungswissenschaften, Wirksamkeit von Interventionen).

In den intensiven Diskussionen ging es u.a. um Herangehensweisen und Strategien, wie mit Komplexität umzugehen ist. Die Gedankenfigur „Reduktion von Komplexität“ spielte hierbei immer wieder eine Rolle. Es wurde kritisch gefragt, wie eine adäquate Reduktion von Komplexität, wie z.B. durch Einbeziehung von Leitdiffenzen nach Luhmann, aussehen kann. Demgegenüber wurde jedoch auch der Standpunkt vertreten, dass Komplexität transparent gemacht, nicht jedoch reduziert werden sollte. In diesem Zusammenhang wurde auch vor einer „Überreduzierung von Komplexität“ gewarnt. Ein weiterer Aspekt ist der Konstruktionscharakter von Komplexität: Programme, Projekte und deren komplexe Struktur werden entwickelt, konstruiert, die ‚Baupläne‘ sind rezipierbare Konzepte; sie können dann dekonstruiert und transparent gemacht werden, ggf. unter Einsatz programmtheoretischer Verfahren. In diesem Zusammenhang wurde auch die „Evaluationstauglichkeit“ von Programmdesigns thematisiert. Einigkeit bestand darin, dass Komplexität und Evaluation von Programmen der beruflichen Bildung ein hochrelevantes Thema ist, welches die Professionalität von Evaluierenden herausfordert.



## **„Professionalisierung in und für Evaluationen“**

**17. Jahrestagung der DeGEval – Gesellschaft für Evaluation  
e.V.**

**Gemeinsam mit der Schweizerischen Evaluationsgesellschaft  
SEVAL  
vom 10. bis 12. September 2014 an der PH Zürich**

*Die 17. Jahrestagung der DeGEval – Gesellschaft für Evaluation e.V. zum Thema „Professionalisierung in und für Evaluationen“ wird vom 10. bis 12. September 2014 in Zürich stattfinden. Ausrichter der Tagung ist die PH Zürich. Die Veranstaltung wird gemeinsam mit der SEVAL realisiert.*

Warum Professionalisierung? Beide Gesellschaften sind der Überzeugung, dass es nunmehr angemessen ist, sich kritisch und konstruktiv mit unserem eigenen Anspruch, die Professionalisierung von Evaluation zu fördern, auseinanderzusetzen. Hiermit ist ein breites Spektrum von möglichen Themen und Herausforderungen adressiert. Wir werden die Standards der Evaluation diskutieren, die Aus- und Weiterbildungsqualität betrachten, über das notwendige Wissen und Können von Evaluatorinnen und Evaluatoren reflektieren, die Frage der Systematisierung des in Evaluationen gewonnenen Wissens besprechen und unsere Rolle für die gesellschaftliche Entwicklung analysieren. In international vergleichender Perspektive soll es auch darum gehen, sich über Akkreditierung und Zertifizierung im Feld der Evaluation zu informieren und zu fragen, wie sich unsere eigene Professionalität zu der von Auftraggebenden ins Verhältnis setzt.

Wir gehen davon aus, dass ein professionsbezogener Austausch und die damit verbundene Reflexion der eigenen Position eine wesentliche Voraussetzung des eigenen Lernens ist. Der Austausch mit SEVAL dürfte die Qualität unserer Diskurse bereichern. Seit Jahren besteht eine – personelle und organisatorische – Zusammenarbeit mit der SEVAL, die in dieser gemeinsamen Veranstaltung ihren besonderen Ausdruck findet.

Die DeGEval bietet mit ihren Jahrestagungen ein Forum zur Präsentation und Diskussion von neuen wissenschaftlichen Erkenntnissen und praktischen Erfahrungen. Auf ihrer 17. Jahrestagung begrüßt die DeGEval speziell Beiträge, die in grundsätzlicher oder feldspezifischer Weise das Tagungsthema „Professionalisierung“ entfalten. Wir laden Sie herzlich ein, an der Diskussion in Zürich teilzunehmen.

Alle weiteren Informationen zur Vorbereitung der Tagung finden Sie auf unserer Homepage unter [www.degeval.de](http://www.degeval.de).

Sofern Sie einen aktiven Beitrag leisten wollen, bitten wir Sie, ein Abstract für einen Vortrag oder ein Poster (maximal 250 Wörter, Nennung der Autor(inn)en wie folgt: Titel, Vorname, Nachname, Institution, Kontaktdaten, E-Mail-Adresse) **bis zum 31. März 2014** über unsere Konferenzhomepage (<https://www.conftool.com/degeval2014/>) zu übermitteln.

Die Abstracts können zu jedem thematischen Feld der Arbeitskreise eingereicht werden. Sie werden an die Sprecherinnen und Sprecher weitergeleitet, die für die inhaltliche Gestaltung der Workshops verantwortlich sind. Die Sessions werden von der DeGEval und der SEVAL gemeinsam organisiert. Bitte ordnen Sie Ihre Skizze bereits selbst mindestens einem der insgesamt 16 Arbeitskreise zu.

Eine Übersicht über bestehende Arbeitskreise finden Sie unter <http://www.degeval.de/arbeitskreise>.

Bei der Einreichung für einen Vortrag muss ein eindeutiger Bezug zum Tagungsthema „Professionalisierung“ gegeben sein. Sofern Beiträge die Evaluation einer Maßnahme zum Fokus haben, soll die Evaluation und nicht die Maßnahme im Vordergrund des Vortrags stehen.

Vom Tagungsthema abweichende Beiträge zum Thema Evaluation können als Poster präsentiert werden.

Am Mittwoch, 10. September 2014, finden in der Tradition der SEVAL-Jahrestagungen Methodenateliers statt. Zu diesen methodischen Workshops können ebenfalls Abstracts eingereicht werden. Das Methodenatelier soll Evaluatorinnen und Evaluatoren einen Austausch und eine – auch konstruktiv-kritische – Diskussion ihrer Methoden und Forschungstechniken erlauben. Eine Reihe von Evaluationen, die innovative oder beachtenswerte methodische Aspekte enthalten, wird im Rahmen von mehreren parallelen Workshops vorgestellt und erörtert. Auftraggebende können sich für ihre nächsten Ausschreibungen über methodische Ansätze inspirieren und auf den neuesten Stand bringen lassen sowie mögliche Auftragnehmer kennenlernen. Im Vordergrund stehen Methoden und nicht die Inhalte oder Ergebnisse der Evaluationen. Das Methodenatelier findet in Form eines lockeren Informationsaustausches statt. Die Referentinnen und Referenten stellen ihre Ansätze in 20 Minuten mit Folien vor, gefolgt von 25 Minuten für die Diskussion. Jede Präsentation wird wenigstens einmal wiederholt. Die Zuhörenden können frei auch während der Präsentation und Diskussion zwischen den einzelnen Ateliers wechseln.

#### **Einsendeschluss ist der 31. März 2014.**

Sie werden bis zum 31. Mai 2014 über die Beitragsannahme sowie über die Form der Präsentation (Vortrag in einer Session oder Poster) informiert.

Weitere Informationen finden Sie unter: <http://www.degeval.de/>

#### Kontakt:

DeGEval – Gesellschaft für Evaluation e.V.  
Johannes Gutenberg-Universität Mainz  
c/o Zentrum für Qualitätssicherung  
und -entwicklung (ZQ)  
Forum universitatis 4  
D-55099 Mainz

Natalie Salf

Tel.: +49 (0) 6131 / 39-26869  
Fax: +49 (0) 6131 / 39-26868

E-Mail: [info@degeval.de](mailto:info@degeval.de)  
<http://www.degeval.de>

## Considering Gender. Qualitätsvolle Evaluierung durch Einbeziehung der Geschlechterperspektive in Evaluationen

*Maria Gutknecht-Gmeiner,<sup>1</sup> Brigitte Wotha,<sup>2</sup> Angela Wroblewski<sup>3</sup>*

### *Hintergrund und Zielsetzung*

Aus Anlass einer grundlegenden Überarbeitung der Standards für Evaluation in der DeGEval (DeGEval-Standards) hat sich der AK Gender Mainstreaming mit der Relevanz der Geschlechterperspektive für die DeGEval-Standards auseinandergesetzt, um einen Input für die Überarbeitung der Standards zu liefern.<sup>4</sup> Das Ziel der Diskussion im AK Gender Mainstreaming war auf Basis einer systematischen Diskussion der DeGEval-Standards erstens die Notwendigkeit für eine umfassende Betrachtung der Geschlechterperspektive in der Evaluation aufzuzeigen und zweitens konkrete Ansatzpunkte für deren durchgängige Berücksichtigung in der Formulierung der DeGEval-Standards und der begleitenden Materialien (Erläuterungen) zu formulieren.

Im Vertrag von Lissabon ist Gender Mainstreaming als Verpflichtung für alle Mitgliedstaaten der EU festgeschrieben. Ziel von Gender Mainstreaming ist die Beseitigung geschlechtsspezifischer Disparitäten und Benachteiligungen durch die durchgängige Berücksichtigung der Geschlechterperspektive in allen Politik- und Handlungsfeldern mit dem Ziel der Gleichstellung der Geschlechter.<sup>5</sup> Die Umsetzung dieses Ziels bleibt nicht auf die Durchführung spezifischer Maßnahmen für Frauen beschränkt, sondern erfordert, dass politische Konzepte und Maßnahmen hinsichtlich ihrer Auswirkungen auf die Situation von Frauen und Männern bereits

---

1 IMPULSE – Evaluation und Organisationsberatung, Wien

2 Büro für Stadt- und Regionalentwicklung, Strande

3 Institut für höhere Studien, Wien

4 Dieser Input wurde im März 2013 an den Vorstand der DeGEval und die für die Überarbeitung der Standards zuständige Arbeitsgruppe weitergeleitet. Die Langfassung des Inputs des AK Gender Mainstreaming kann über die Geschäftsstelle der DeGEval (info@degeval.de) oder die Arbeitskreissprecherin (wroblewski@ihs.ac.at) bezogen werden. Wir bedanken uns bei Regina Frey, Anne Rösger, Arn Sauer, Helene Schiffbänker, Martina Schu und Heide Trommer für ihre Rückmeldungen und Anmerkungen zu einer ersten Fassung des Inputs.

5 Europäische Kommission (1996): Einbindung der Chancengleichheit in sämtliche politische Konzepte und Maßnahmen der Gemeinschaft [Komm(1996)67endg.].

in der Planungsphase systematisch überprüft und die möglichen Auswirkungen bei der Umsetzung berücksichtigt werden.

Gender Mainstreaming ist daher auch in der Evaluation zu verankern. Dabei lassen sich zwei Aspekte des Gender-Mainstreaming-Ansatzes unterscheiden. Zum einen meint Gender Mainstreaming in der Evaluation eine durchgängig gleichstellungsorientierte Zugangsweise in allen Evaluierungen. Um dies zu erreichen, sind Daten geschlechtsdifferenziert zu erheben und auszuwerten, Kontext und Ergebnisse systematisch auf mögliche geschlechtsspezifische Unterschiede hin zu untersuchen und die Bewertung der Programmwirkungen in gleichstellungsorientierter Perspektive vorzunehmen. Dies erfordert entweder die Weiterentwicklung bestehender Datengrundlagen und Indikatorensysteme oder die Erhebung spezifischer Primärdaten sowie die spezifische Reflexion der methodischen Vorgangsweise. Zum anderen wird im Rahmen der Evaluation von Gender-Mainstreaming-Implementationsprozessen die Erreichung spezifischer Gleichstellungsziele gemessen. Beide Aspekte erfordern, den Blick auf die Organisationsstruktur und die Handlungsabläufe eines Handlungsfeldes zu legen, wodurch im Allgemeinen die Komplexität des Evaluationsgegenstandes steigt.

### *Begriffsklärungen*

Die geforderte Berücksichtigung der Geschlechterperspektive schließt an ein Verständnis von **Gender** als soziales Konstrukt an. Dabei geht es nicht nur um die Differenzen zwischen den Frauen und den Männern, sondern auch um die Differenzierungen, die gemacht werden, um gesellschaftliche Geschlechtermuster und die damit verbundenen Inklusions- und Exklusionsmechanismen immer wieder neu herzustellen. Die Mitglieder des Arbeitskreises vertreten ein differenziertes Konzept von Gender, das sich nicht auf vermeintlich homogene Genusgruppen bezieht, sondern Gender als immer schon verwoben mit weiteren Strukturkategorien (wie z.B. Alter, sozioökonomischer Status oder ethnischer Hintergrund) denkt. Das Ziel der Strategie Gender Mainstreaming ist die Gleichstellung von Frauen und Männern in ihrer Vielfalt.

Unter **Genderdimension** des Evaluationsgegenstandes werden geschlechtsspezifische Unterschiede in Zugang zu und Teilhabe am Evaluationsgegenstand sowie die geschlechterdifferenzierte Wirkung des Evaluationsgegenstandes verstanden, die durch vorherrschende gesellschaftliche Machtverhältnisse, Strukturen, Werte, Normen und Rollenbilder bedingt sind.

**Genderkompetenz** wird verstanden als die Fähigkeit von Personen, die Genderdimension des Evaluationsgegenstandes zu erkennen und in der Beauftragung bzw. Durchführung der Evaluation zu berücksichtigen. Dies erfordert die Einbeziehung der Methoden und Erkenntnisse der Genderforschung sowie der feministischen Theorie wie auch eine hohe Reflexionsfähigkeit der an der Evaluation Beteiligten.

*Berücksichtigung der Geschlechterperspektive in den DeGEval-Standards*

Die Anmerkungen des AK Gender Mainstreaming orientieren sich an der aktuellen Fassung der Standards sowie deren Erläuterungen und sind jeweils für die Bereiche Nützlichkeit, Durchführbarkeit, Fairness und Genauigkeit formuliert.<sup>6</sup> Dabei wurden die grundsätzliche Ausrichtung oder Formulierung der bestehenden Standards selbst nicht diskutiert.

*Nützlichkeit*

Die Nützlichkeitsstandards sollen die zielgenaue Umsetzung der Evaluation entsprechend der Bedürfnisse der Betroffenen und Beteiligten, insbesondere der vorgesehenen Nutzerinnen und Nutzer, sicherstellen und ihnen Information über alle Aspekte der Evaluation zur Verfügung stellen. Aus Sicht des AK Gender Mainstreaming ist hier die Geschlechterperspektive in die Analyse des Konzepts sowie in die Stakeholderanalyse von Anfang an zu integrieren, um Fehlplanungen der Maßnahme aufzeigen und alle Beteiligten und Betroffenen in die Evaluation einbeziehen zu können.

Nach dem Konzept des Gender Mainstreamings ist im Zusammenhang mit den Nützlichkeitsstandards zentral, die der Evaluation zu Grunde gelegten Werte hinsichtlich ihrer Bedeutung für das Geschlechterverhältnis zu reflektieren. Ein wesentlicher Aspekt dabei ist, den Bewertungsmaßstab für die Genderdimension in der Evaluation darzulegen. Dies erfordert auch aufzuzeigen, mit welchem Bewertungsmaßstab festgestellt wird, dass ein Evaluationsgegenstand bzw. eine Fragestellung gegebenenfalls keine Genderdimension aufweist.

Weiter ist explizit zu prüfen, in welcher Weise Frauen und Männer vom Evaluationsgegenstand und von der Evaluation unterschiedlich betroffen sind bzw. wie Frauen und Männer am Evaluationsgegenstand und an der Evaluation beteiligt sind. Die Geschlechterdimension z.B. in der Wirkungsanalyse erlaubt, geschlechtsspezifisch unterschiedliche Wirkungen festzustellen. Bei der Formulierung von Handlungsempfehlungen sollte das Augenmerk darauf gerichtet werden, inwieweit dadurch festgestellte geschlechtsspezifische Disparitäten abgebaut werden.

Für die Durchführung von Evaluationen ist es aufgrund dieser Anforderungen unabdingbar, dass Evaluatorinnen und Evaluatoren nicht nur methodisch und fachlich kompetent sind, sondern auch Genderkompetenz mitbringen.

*Durchführbarkeit*

Kern der Durchführbarkeitsstandards ist die Wahl angemessener Verfahren. Aus einer Geschlechterperspektive ist dabei zum einen die Entwicklung von aussagekräftigen Genderindikatoren zentral. In den letzten Jahren ist insbesondere in der Evaluation von Entwicklungszusammenarbeit sowie Arbeitsmarktpolitik die Anzahl der mit Genderindikatoren durchgeführten Evaluationen stark gestiegen, die als Orientierungsmaßstab für andere Bereiche dienen können. Zum anderen ist aus

---

<sup>6</sup> Die aktuelle Version der Standards und der Begleitmaterialien ist verfügbar unter: <http://www.degeval.de/degeval-standards> [07.08.2013].

einer Geschlechterperspektive essenziell, dass Verfahren eingesetzt werden, durch die die Genderdimensionen in den Fragestellungen adäquat bearbeitet werden können.

### *Fairness*

Den Mitgliedern des AK Gender Mainstreaming erscheint die Annahme, dass Evaluatoredinnen und Evaluatoren völlig unparteiisch sind bzw. ihre Werthaltungen und persönlichen Gefühle vollständig ausklammern können als unrealistisch und wird als problematisch verworfen. Aus einer Geschlechterperspektive erscheint es notwendig, im Rahmen der Evaluation die eigene Rolle als Evaluator/Evaluatoredin laufend zu reflektieren und transparent zu machen. Da viele Wertungen – bewusst oder unbewusst – durch die Sprache erfolgen, gehört es auch zu den grundlegenden Anforderungen an Evaluation, eine geschlechtergerechte Sprache zu verwenden, um unbewusste oder bewusste Diskriminierungen aufgrund von Geschlecht zu vermeiden.

### *Genauigkeit*

Die Genauigkeitsstandards erfordern eine genaue Analyse der geschlechtsspezifischen Ausprägungen der sozialen Realitäten, der ökonomischen und der politischen Rahmenbedingungen, die den Kontext eines Evaluationsgegenstandes betreffen. Alle Informationen sind geschlechtsspezifisch aufzuarbeiten und geschlechterdisaggregiert darzustellen. Nicht geschlechterdisaggregierte gruppenbezogene Daten sind nicht valide.

### *Schlussfolgerungen*

Der AK Gender Mainstreaming versteht die Berücksichtigung einer Geschlechterperspektive in der Evaluation als Qualitätsmerkmal von ‚guter‘ Evaluation. Gute, nützliche, vollständige, faire und genaue Evaluationen können nur mit einer theoretisch fundierten und konsequenten Berücksichtigung der Genderdimension in allen Phasen und Aspekten durch genderkompetente Evaluatoredinnen und Evaluatoren stattfinden.

Der AK Gender Mainstreaming plädiert daher dafür, bei der Überarbeitung der DeGEval-Standards die Geschlechterperspektive durchgehend zu berücksichtigen und diese in den Erläuterungen explizit zu thematisieren. Es geht dabei nicht darum, zusätzlich einen ‚Gender-Standard‘ aufzunehmen, sondern im Sinne des Gender Mainstreamings die Geschlechterperspektive in alle Standards zu integrieren.

## „Schulqualität zwischen Evidenzbasierung und Effekthascherei“

Positionspapier als Ergebnis der Frühjahrstagung 2013 des AK Schulen der DeGEval

*Michael Frais,<sup>1</sup> Angela Stumpp-Marx<sup>2</sup>*

Hinweis: Dieses Positionspapier ist in Abstimmung mit Vertreterinnen der Universität Salzburg, der PH Salzburg sowie der PH Tirol entstanden. Es stellt allgemeine Erkenntnisse aus dem Bereich schulischer Evaluationsverfahren in den Kontext der Diskussionen der diesjährigen Frühjahrstagung des Arbeitskreises Schulen der Gesellschaft für Evaluation (DeGEval) in Salzburg. Da ein Großteil der Teilnehmenden aus Österreich kam, entwickelten sich die diskutierten Inhalte insbesondere vor dem Hintergrund des österreichischen Schulsystems. Die Frühjahrstagung wurde in Kooperation mit der School of Education der Paris-Lodron-Universität Salzburg und der Pädagogischen Hochschule Salzburg am 6./7. Juni 2013 an der Universität Salzburg durchgeführt.

### *Aktivitäten des AK Schulen*

Der Arbeitskreis (AK) Schulen der Gesellschaft für Evaluation führt seit dem Jahr 2010 regelmäßige Frühjahrstagungen zum Thema Schulevaluation durch. Zusätzlich wird diese Thematik auch auf den Jahrestagungen der Gesellschaft für Evaluation im Rahmen von Sessions aus unterschiedlicher inhaltlicher Perspektive diskutiert. Der AK Schulen bezieht bei diesen Veranstaltungen die Sichtweisen von Lehrpersonen, Evaluator(inn)en und Wissenschaftler(inne)n mit ein. Er versteht sich als Forum, welches an der Schnittstelle von Wissenschaft und Praxis grundlegenden Fragestellungen rund um das Thema Evaluation nachgeht.

Ausgehend von der Frage der Wirkungen externer Schulevaluationsverfahren, die im Jahr 2010 aus vorwiegend wissenschaftlicher Perspektive beleuchtet wurde, standen im Jahr 2011 die Themen Qualitätsmanagement und Führung als wichtige beeinflussende Faktoren bei der Durchführung von Evaluationsvorhaben im Fokus. 2012 erörterten wir aus wissenschaftlicher und schulpraktischer Sicht, wie Impulse für Schulentwicklung generiert werden (können).

---

1 Leitung der Evaluationsstelle ESOB, Pädagogische Hochschule Zürich

2 Sächsisches Bildungsinstitut, Radebeul

*Diskussionsstand und Erfahrungswissen im Bereich Schulevaluationen*

Es hat sich in den vergangenen Jahren an vielen Beispielen gezeigt, dass Evaluationen im schulischen Feld zum Teil mit Nutzlosigkeit, Wirkungslosigkeit und Effekthascherei in Verbindung gebracht werden. Insofern schien es dringend geboten, dieser Thematik im Rahmen einer Frühjahrstagung nachzugehen. Dies vor allem auch deshalb, weil weder die beteiligten Institutionen wie Schulaufsicht, Evaluationsfachstellen und Schulen noch die mitwirkenden Evaluator(inn)en ein Interesse an Nutzlosigkeit, Wirkungslosigkeit und Effekthascherei als Ziel oder Ergebnis von Evaluationen im Bildungsbereich haben.

Der weitgehend flächendeckenden Einführung von externen Schulevaluationsverfahren zwischen 2000 und 2010 lag implizit eine mechanistisch orientierte Erwartung an die Wirkungen externer Schulevaluationsverfahren zu Grunde. Externe Evaluation als neues Instrument der Schulaufsicht führe über die Rückmeldung an die Schulen automatisch zu einer Qualitätsverbesserung des schulischen Bildungssystems, so vielerorts die Erwartung. Ständen in Zusammenhang mit der Einführung dieser Verfahren zunächst Akzeptanz- und Implementierungsfragen im Vordergrund, kam mit zunehmender Betriebsdauer die berechnete Frage nach ihren Wirkungen auf. Diese wurde auch vor dem Hintergrund der Kosten, die für externe Schulevaluationsverfahren aufgewendet werden, und im Kontext notwendiger Sparmaßnahmen zur Sanierung der Staats-, Landes- und Kantonshaushalte gestellt.

Wirkungen externer Evaluationsverfahren auf die Leistungen der Schülerinnen und Schüler sind zurzeit nicht mit zufriedenstellender wissenschaftlicher Genauigkeit nachzuweisen, so lautete eine Erkenntnis aus der Frühjahrstagung 2010 zur Wirkung externer Schulevaluationsverfahren. Allenfalls könnte mit einem sehr hohen finanziellen Aufwand ein gewisser Genauigkeitsgrad erreicht werden, der zur Hypothesenbildung beitragen kann.

Damit einher geht die Erkenntnis, dass externe Evaluationsverfahren häufig nicht zu den gewünschten Impulsen für die schulische Arbeit führen, da die Ergebnisse und Erkenntnisse aus Evaluationen oftmals nicht in ausreichendem Maß für die Verbesserung schulischer Praxis genutzt werden.

*Zukunftsperspektive: Wirkungen und Impulse durch partizipativ gestaltete Evaluationen*

Externe Evaluationsverfahren erfüllen häufig nur das Kriterium der Außenperspektive von Bildungspolitik und Schulaufsicht. Jedoch ist es zwingend notwendig, auch der internen Perspektive jeder einzelnen Schuleinheit angemessenen Raum zu geben. Evaluationen, die aus schulinterner Sicht gestaltet oder mitgestaltet werden, benötigen Raum für eigene Fragestellungen und Perspektiven, die u.a. im Schulprogramm verortet sind. In den vergangenen Jahren hat sich gezeigt, dass die Einbeziehung der Beteiligten und Betroffenen in den Evaluationsprozess zu größerer Akzeptanz und mehr Interesse bei eben diesen Stakeholdern führt und somit eine größere Effektivität in der Umsetzung von Schulentwicklungsaktivitäten auf der Basis empirisch erhobener Daten erreicht werden kann.

Dieser partizipative Ansatz – evaluationstheoretisch auch als interpretativ konstruktivistischer oder als transformativ emanzipatorischer Ansatz zu verorten – erhöht in der Regel die Validität, die Rezeption der Evaluationsergebnisse bzw. des Berichtes durch die Stakeholder sowie Nutzen und Wirksamkeit von Evaluationen. Neben den klassischen sozialwissenschaftlichen Gütekriterien spielen dabei Kommunikation, Intervention, Transparenz und Relevanz eine zunehmend gewichtige Rolle.

Damit Partizipation zu Impulsen führen bzw. sie verstärken kann, benötigen Schulen sowohl im Bereich von Evaluationen als auch im Bereich von Schulentwicklung das entsprechende Know-how, die nötigen Kompetenzen sowie Gestaltungsräume, die innerhalb der durch Bildungspolitik und Schulaufsicht gesetzten Rahmenbedingungen eigenständig genutzt werden können. Gleichzeitig erhöht sich damit für Schulen die Notwendigkeit der Transparenz in Bezug auf ihre Aktivitäten. Zudem muss berücksichtigt werden, dass Gestaltungsräume zu einer Rechenschaftspflicht bis auf Ebene einzelner Schuleinheiten führen.

#### *Schulische Steuerung unter Berücksichtigung partizipativ angelegter Evaluationen*

Wie bereits oben aufgezeigt worden ist, berühren Schulevaluationen, insbesondere solche, die als interne Evaluation oder Selbstevaluation organisiert sind, die Frage schulischer Steuerung. Dabei wird zwischen Top-down- und Bottom-up-Steuerung unterschieden.

Eine Top-down-Steuerung des Schulsystems dient in erster Linie dem Informations- und Steuerungsbedürfnis der Bildungspolitik und Schulaufsicht, aber nur zum Teil den Interessen der einzelnen Schuleinheit. Bei einer ausschließlichen Top-down-Steuerung werden vorhandene Ressourcen auf Schulebene nicht genutzt und somit sinnvolle und effektive Entwicklungsperspektiven behindert oder sogar blockiert. Dies haben auch die Diskussionen auf der diesjährigen Frühjahrstagung in Salzburg deutlich gemacht: Die Selbstwirksamkeit der Schuldirektor(inn)en und Lehrpersonen wurde als ein sehr wichtiger Motivationsfaktor in der schulischen Arbeit bewertet, die sich aber nur dann entfalten kann, wenn die entsprechenden Gestaltungsspielräume vorhanden sind. In der gegenwärtigen Situation schulischer Steuerung in Österreich werden diese Ressourcen anscheinend unzureichend aktiviert, so dass vorhandenes Potenzial zur Generierung eines pädagogischen Mehrwerts nicht oder nur unzureichend genutzt wird. In Deutschland und der Schweiz kann dieses Potenzial zum Teil erheblich besser ausgeschöpft werden.

Wichtig ist in diesem Zusammenhang nochmals hervorzuheben, dass jede Form von Gestaltungsmöglichkeit eine Schule in die Position der Bringschuld transferiert. Konkret bedeutet dies, dass schulische Impulse, die zu Schulentwicklungsmaßnahmen führen, zwingend eine Legitimationspflicht, z.B. durch eine Evaluation, nach sich ziehen und damit wiederum die Schulaufsicht entlasten können. Denkbar wäre beispielsweise, dass sich in Zukunft die Funktionsweise der Schulaufsicht dahingehend ändert, dass Schulen, die ihre Aktivitäten und Entwicklungsprojekte weder transparent machen noch durch Evaluation legitimieren können, eine Art Eingreifmodus der Schulaufsicht auslösen.

*Schulische Steuerung zwischen Innovation und prozessorientierter Vorgehensweise*

Mit einer Top-down-Steuerung werden in der Regel Innovationen oder eine Veränderung der Rahmenbedingungen zu allgemeiner Verbindlichkeit erklärt und müssen von allen Schulen umgesetzt werden. Derartige Steuerungsimpulse setzen Zeichen, lassen sich schnell realisieren und erhalten dadurch ihre Effizienz, dass sie von oben verordnet werden. Selbstverständlich müssen Schulen auf diese bildungspolitischen Innovationsmaßnahmen angemessen reagieren und diese umsetzen.

Gleichzeitig ist eine prozessorientierte Vorgehensweise auf Ebene der einzelnen Schuleinheit notwendig. Nur wenn Zeiträume zur Umsetzung vorhanden sind und ein längerfristiger Zeithorizont eingeplant wird, können Veränderungen aller Art nachhaltig und wirkungsorientiert implementiert werden. Es ist jedoch bekannt, dass solche längerfristigen Prozesse immer wieder durch bildungspolitisch motivierte, kurzfristige und kurzlebige Maßnahmen gestört und häufig nicht mit den notwendigen Ressourcen systematisch zu Ende geführt werden.

Schulen dürfen kein Spielball parteipolitischen Gezänks und von kurzfristigem bildungspolitisch induziertem Aktionismus sein. Folglich muss es darum gehen, eine ausgewogene Balance zwischen bildungspolitisch orientierter Top-down-Steuerung und einer Bottom-up-Steuerung herzustellen, so dass der zum Teil sehr individuelle Kontext jeder einzelnen Schuleinheit Berücksichtigung findet. Dies entspricht der Anforderung an ein Steuerungsmodell, welches eine Top-down-Steuerung im Hinblick auf eine innovativ ausgerichtete Entwicklungsdynamik mit einer prozessorientierten, langsamen, aber dafür auf Wirkung und Nachhaltigkeit ausgelegten Vorgehensweise kombiniert.