



In zehn Schritten zu einer guten Evaluation

Bildungsplanung ist ohne systematische Beschreibung und Bewertung von Bildungsmaßnahmen nicht mehr denkbar. Was aber zeichnet eine gute Evaluation aus? Wie häufig werden Ergebnisse geschönt? Und wie gut ist die Evaluationsqualität in der Berufsbildung?

Von Lars Balzer und Wolfgang Beywl. Lars Balzer ist Leiter der Fachstelle Evaluation am EHB. Wolfgang Beywl ist Leiter der Professur für Bildungsmanagement sowie Schul- und Personalentwicklung an der PH FHNW.

Evaluation begegnet einem an vielen Orten. Dabei wird ganz Unterschiedliches darunter verstanden. In nahezu jedem Buch, das sich intensiver mit diesem Thema befasst, findet sich eine andere Definition, und je nach wissenschaftlicher Disziplin wird der Begriff unterschiedlich akzentuiert. Zudem werden heute «allzu leicht [...] auch triviale Formen der Rückmeldung zu Evaluationen» (Böttcher). Das provoziert Kritik wie: «sich epidemisch ausbreitende, neue Krankheit, die Evaluitis» (Frey) oder «Vorsicht vor Evaluationismus!» (Kappler).

In diesem Text – und dem zugrunde liegenden Sachbuch – wird Evaluation als eine wissenschaftliche Dienstleistung verstanden, die für festgelegte Zwecke und

nach begründeten Kriterien insbesondere Programme und Projekte beschreibt und bewertet. Die Bewertung geschieht systematisch, transparent sowie nachvollziehbar und basiert auf Daten und Informationen, die mithilfe sozialwissenschaftlicher Methoden gewonnen werden. Nach unserem Verständnis müssen Personen für diese Tätigkeit besonders qualifiziert sein. Dabei gehen sie in zehn Schritten vor, die kein rigide einzuhaltendes Strickmuster vorgeben, sondern den spezifischen Rahmenbedingungen der konkreten Evaluation anzupassen sind.

Im ersten Schritt ist der Evaluationsgegenstand zu bestimmen. Es wird eine für Aussenstehende nachvollziehbare

Erstbeschreibung erstellt. Relevante Dimensionen sind zum Beispiel Name des Bildungsangebots, Initiatoren, Verantwortliche, Mitarbeitende, Zielgruppen, Inhalte, Durchführungspläne, Materialien und Ressourcen. Evaluationsgegenstände sind in der Regel konkrete Projekte, können aber auch gesamte Organisationen sein. Oft gilt es, die verfolgten Projektziele auszuformulieren. Deren Überprüfung kann Thema der Evaluation werden.

Im zweiten Schritt werden die am Evaluationsgegenstand interessierten Akteure, die sogenannten Stakeholder bzw. Beteiligte und Betroffene, bestimmt. Beteiligte sind Personen, Gruppen oder Organisationen, die in Bezug auf den Evaluations-

gegenstand aktiv oder einflussreich sind, zum Beispiel Finanzierende, Verantwortliche oder Mitarbeitende. Betroffene sind insbesondere Personen mit wenig Einfluss, beispielsweise die Zielgruppen. Akteure, die Informationsinteressen und gegebenenfalls Erwartungen in Bezug auf die Evaluation hegen, sind potenziell Nutzende. Auch die Evaluierenden gehören zu den Akteuren: Interne Evaluation wird von Personen der den Evaluationsgegenstand tragenden Organisation gesteuert. Externe Evaluationen führen Evaluierende durch, die von ausserhalb kommen.

Im dritten Schritt geht es um die Festlegung der Evaluationszwecke sowie die Formulierung spezifizierter, klar orientierender und priorisierter Evaluationsfragestellungen. Es werden entscheidungsorientierte, rechenschaftslegungsorientierte, verbesserungsorientierte und wissensgenerierende Evaluationen unterschieden. Zweck einer entscheidungsorientierten Evaluation ist es, Richtungsentscheidungen wie Ausweitung und Reduzierung eines Angebotes zu begründen. Bei der Rechenschaftslegung geht es darum, den

Keinesfalls dürfen Bewertungskriterien ad hoc, rein intuitiv, methodisch unkontrolliert oder nicht nachvollziehbar festgelegt werden.

Wert eines Evaluationsgegenstandes zu bestimmen. Beides geschieht in der Regel durch Bilanzierung im Rahmen einer sogenannten summativen Evaluation. Die formative Evaluation hingegen beschafft Informationen für Verbesserungen des Evaluationsgegenstandes. Die auf den Zwecken aufbauenden Fragestellungen der Evaluation sollen in Worte fassen, welche Informationen die vorgesehenen Nutzen benötigen. Es kann um Themen gehen wie Bedarfsgerechtigkeit von Angeboten, Kundenzufriedenheit, Lernerfolg, Transfernutzen oder Kosten-Nutzen-Angemessenheit.

Bewertungskriterien sind Voraussetzung für jede systematische und faire Be-

wertung. Sie werden im vierten Schritt festgelegt. Sie bieten eine Bezugsbasis, auf die sich Urteile über die Güte oder Tauglichkeit von Evaluationsgegenständen stützen. Operationalisierte Kriterien sollen vor Beginn der Datenerhebung, allerspätestens aber vor Beginn der Datenauswertung festgelegt werden. Vielfach wird es notwendig sein, Stakeholder nach ihren Bewertungsmaßstäben zu fragen und die anzulegenden Kriterien auszuhandeln. Keinesfalls dürfen Bewertungskriterien ad hoc, rein intuitiv, methodisch unkontrolliert oder nicht nachvollziehbar festgelegt werden.

Im fünften Schritt planen die Evaluierenden das konkrete methodische Vorgehen und setzen es um. Ein zum Umfeld des Evaluationsgegenstandes passendes Erhebungsdesign sowie angemessene Erhebungsmethoden werden ausgewählt und ein Datenerhebungsplan wird entworfen. Er legt fest, welche Daten, auf welche Art und Weise, bei wem und wann erhoben werden sollen. Das Erhebungsdesign klärt die logistische und zeitliche Strukturierung des Evaluationsprojektes sowie den methodischen und technischen Vorgang des Datenbeschaffens. Eine wichtige Designentscheidung ist, zu wie vielen Zeitpunkten und bei wie vielen Gruppen Daten erhoben werden. Es werden daraufhin Erhebungsmethoden eingesetzt, um die Daten zur Beantwortung von Fragestellungen zu gewinnen. Für die Wahl der Methoden gibt es kein Patentrezept; man muss sich nach den gegebenen Fragestellungen und Rahmenbedingungen richten.

Sorgfältig geplant beginnen im sechsten Schritt die Erhebungen. Dabei ist die logistische und technische Seite ebenso zu beachten wie der Umgang mit Störungen. Evaluation wird nicht unter Laborbedingungen durchgeführt, sondern ist angewandte Sozialwissenschaft im dynamischen Lern- und Arbeitsfeld, das durch die Evaluation auch beeinflusst wird. Die Herausforderung besteht darin, gleichzeitig flexibel auf die Impulse des Feldes zu reagieren und wissenschaftlich systematisch zu verfahren. Bei der Durchführung der Datenerhebungen konkurrieren Ansprüche auf Gül-

Jeder zweite Forscher fühlt sich vom Auftraggeber unter Druck gesetzt

Zwischen Auftraggebenden und Evaluierenden können diverse Abhängigkeiten bestehen. Sei es, dass interne Evaluationen von weisungsbefugten übergeordneten Hierarchiestufen in Auftrag gegeben werden oder dass extern Evaluierende wirtschaftlich von weiteren Aufträgen abhängig sind. Dies birgt die Gefahr, dass interessen-geleitet inhaltliche Erwartungen an die Evaluationsergebnisse und -berichte geknüpft werden. Diese Gefahr besteht nicht nur in der Theorie. In einer 2014 durchgeführten anonymen Befragung von Mitgliedern der Schweizerischen Evaluationsgesellschaft gab rund die Hälfte der 123 Antwortenden an, dass sie sich während ihrer Evaluationstätigkeit bereits beeinflusst oder unter Druck gesetzt gefühlt hat. Für 90 Prozent von diesen war das kein Einzelfall. Bei 60 Prozent der Befragten verlangten Auftraggebende eine inhaltlich verzerrende Darstellung von Ergebnissen. Mit diesem Dilemma umzugehen, ist nicht leicht. 12 Prozent der Befragten meinten, dass sie beim jüngsten Beeinflussungsvorfall keine Änderungen am Bericht vorgenommen hätten. 72 Prozent reagierten mit Änderungen ohne falsche oder ungenaue Darstellung, immerhin 9 Prozent mit Änderungen, die zu falscher oder ungenauer Darstellung führten. Bei 7 Prozent geschah etwas anderes infolge der Beeinflussung. Mehr als die Hälfte der Befragten sieht Chancen zur Prävention darin, schon zu Beginn der Evaluation mögliche negative Resultate zu diskutieren und ein grösseres Verständnis der Zweckbestimmung, Funktion und Gütemassstäbe der Evaluation zu schaffen.

Pleger, L., Sager, F. (2016, im Druck): Don't tell me cause it hurts – Beeinflussung von Evaluierenden in der Schweiz. In: Zeitschrift für Evaluation (Nr. 1).

tigkeit, Zuverlässigkeit und Objektivität mit ökonomischen Überlegungen, Termintreue oder der Auslösung von Beteiligungsbereitschaft bei den Stakeholdern.

Nach Abschluss der Erhebungen sind im siebten Schritt die Daten systematisch zu erfassen, zu kontrollieren und auszuwerten, indem sie verdichtet, überschaubar dargestellt, verbal zusammengefasst und in Bezug zu den Fragestellungen in deutende, erklärende Sinnzusammenhänge

Es ist breit anerkannt, dass die Evaluationsdurchführung selber bereits vielfältig Nützliches für die Praxis leisten kann.

gestellt werden. Dann sind Schlussfolgerungen zu ziehen und Bewertungen auf Grundlage der Bewertungskriterien vorzunehmen. Es ist wichtig, dass die Bewertungen aus den empirischen Daten unter Rückgriff auf die zuvor festgelegten Kriterien begründet werden.

Damit eine Evaluation Nutzen für die Praxis entfalten kann, muss über ihre Ergebnisse berichtet werden. Dies erfolgt im achten Schritt. Dabei ist zu klären, wem, wie, wann und in welcher Form berichtet werden soll.

Evaluationen lohnen sich nur dann, wenn sie für und durch die Praxis genutzt werden, was im neunten Schritt vorbereitet wird. Unter Nutzung einer Evaluation wurde lange Zeit allein die Verwendung ihrer Ergebnisse verstanden. Inzwischen ist breit anerkannt, dass die Evaluationsdurchführung selber bereits vielfältig Nützliches für die Praxis leisten kann.

Im zehnten Schritt kann die Evaluation selbst Gegenstand einer Evaluation werden. Es sollte eine professionelle Verpflichtung sein, die geleistete Arbeit zu bewerten und Verbesserungsmöglichkeiten für künftige Evaluationen abzuleiten.

Balzer, L., Beywl, W. (2015): *evaluiert – Planungsbuch für Evaluationen im Bildungsbereich*. Bern, hep verlag.

Evaluationen in der Berufsbildung: «Steigende Professionalisierung»

Es wird viel evaluiert in der schweizerischen Berufsbildung. Allerdings ist es schwierig, sich einen Überblick über Intensität und Qualität zu verschaffen. Datenbanken wie Aramis, Daris, forschungportal.ch oder die Projektdatenbank der Schweizerischen Koordinationsstelle für Bildungsforschung ermöglichen nur einen eingeschränkten Einblick. Dennoch lassen sich folgende Arten unterscheiden:

- Periodisch durchgeführte, externe Schulevaluationen zur Überprüfung von Qualitätsstandards sowie zur Identifikation von Stärken und Entwicklungsmöglichkeiten einzelner Berufsfachschulen – oft sehr standardisiert und eher Auditierung als Evaluation.
- Interne Schulevaluationen, die bis hin zur Ebene des Unterrichts ansetzen – eher selten und kaum öffentlich zugänglich.
- Einzelevaluationen, die von öffentlichen oder privaten Institutionen in Auftrag gegeben werden und für konkrete Projekte und Programme interne Informationsbedürfnisse befriedigen.
- Einzelevaluationen, die verpflichtend durchgeführt werden bei Projekten, die finanzielle Unterstützung des Bundes nach Artikel 54/55 BBG erhalten.
- Nationale Evaluationen des SBFJ zur Überprüfung und Begleitung innovativer Elemente des Berufsbildungssystems.

Eine Bestimmung der gelieferten «handwerklichen» Qualität und insbesondere der Nützlichkeit der Evaluationen ist bei fehlendem Gesamtüberblick nicht möglich. Es gibt aber Hinweise auf eine steigende Professionalisierung:

- Die Universität Bern bietet seit 2002 ein Weiterbildungsprogramm zur

Evaluation an mit Einzelkursen und drei Universitätsabschlüssen (CAS, DAS, MAS). Das Programm wurde von mehreren Hundert Personen besucht. Darüber hinaus existieren Weiterbildungsangebote anderer Anbieter.

- Im Register der Schweizerischen Evaluationsgesellschaft (SEVAL), die über 500 Mitglieder zählt, finden sich heute über 250 Einträge. Viele stammen von Mitarbeitenden aus Büros mit einem Kerngeschäft in der Evaluation.
- Die SEVAL erliess 2001 Evaluationsstandards. Sie bilden für viele Professionelle eine Grundlage ihrer Arbeit.
- Waren Ausschreibungen vor zehn Jahren oft wenig orientiert an der Fachsprache und in den Zwecksetzungen unbestimmt, werden heute Wirkmodelle gefordert und zahlreiche exakt formulierte Evaluationsfragen vorgegeben. Manchmal werden sogar Instrumente und Stichprobenumfänge festgelegt, was den Zusatznutzen eines extern anregenden Blicks erschweren kann.
- Um aus den Evaluationen zu lernen, ist es wünschenswert, dass Berichterstattungen öffentlich zugänglich gemacht werden. Dass dies zunehmend passiert, ist ein weiterer Hinweis auf die gestiegene Professionalisierung.

Die schweizerische Berufsbildung verfügt über eine lange und ausgeprägte Praxis an Evaluationsaktivitäten, die den internationalen Vergleich nicht scheuen muss. Es wäre wünschenswert, die Erfahrungen zusammenzuführen und einer kritischen Analyse zu unterziehen mit dem Ziel, in der Zukunft wissenschaftlich noch fundiertere und für die Berufsbildung optimal nützliche Evaluationen zu gewährleisten.