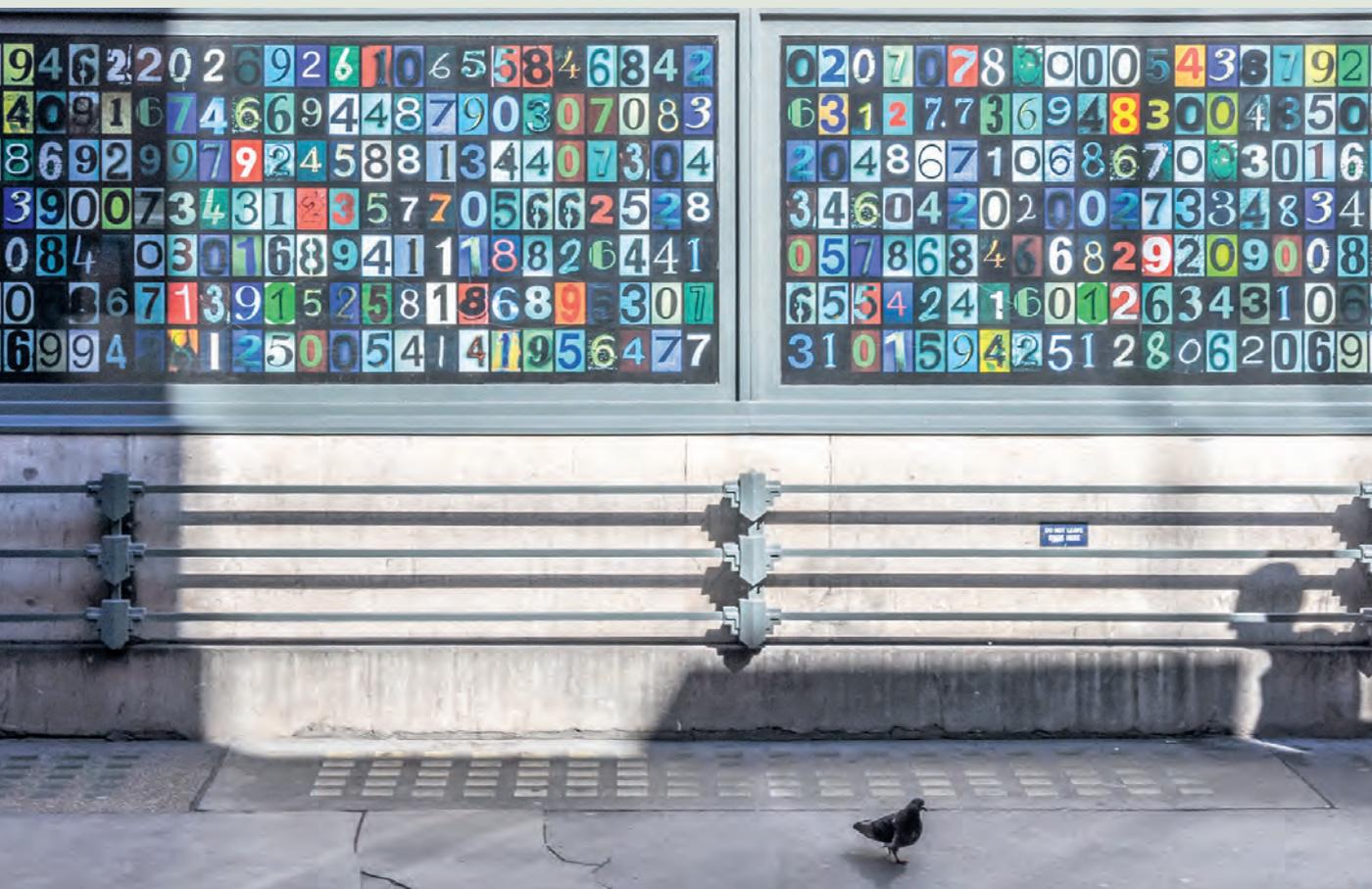


evaluiert

Planungsbuch für Evaluationen im Bildungsbereich



Lars Balzer
Wolfgang Beywl



Inhaltsverzeichnis

Vorwort.	7
1 Einstieg ins Thema	11
2 Was ist Evaluation?	14
2.1 Wissenschaftliche Evaluation statt Alltagsbewertung	15
2.2 Evaluation als wissenschaftliche Dienstleistung statt (Grundlagen-)Forschung	18
2.3 Evaluation als Qualitätsmanagement bzw. Controlling?	25
2.4 Evaluation als Beitrag zur gesellschaftlichen Entwicklung	26
3 Evaluationsprozess und Auftrag	28
4 Bestimmung des Evaluationsgegenstandes	34
5 Bestimmung der interessierten Akteure und die Rolle der Evaluierenden	45
5.1 Geld- und Auftraggebende der Evaluation	47
5.2 Finanzierende und Träger des Bildungsangebots	48
5.3 Programmleitende und -mitarbeitende	49
5.4 Teilnehmende	49
5.5 Vorgesetzte, Arbeitgebende, Arbeitskollegen	50
5.6 Sonstige Akteure	50
5.7 Evaluierende	51
5.8 Akteure im Zusammenspiel	54
6 Bestimmung der Evaluationszwecke und -fragestellungen	57
6.1 Evaluationszwecke	57
6.1.1 Funktionale Unterscheidung von Evaluationszwecken	57
6.1.2 Evaluationszwecke nach Verwendungsabsichten	60
6.1.3 Zusammenspiel von Funktion, Zweck und Rolle der Evaluation	65
6.1.4 Verdeckte Zwecksetzungen	67
6.2 Evaluationsfragestellungen	69
6.2.1 Festlegen der Fragestellungen	69
6.2.2 Formulierung von Fragestellungen	72
6.2.3 Arten von Fragestellungen	74
7 Festlegung der Bewertungskriterien	76
8 Auswahl von Erhebungsdesign und -methoden	84
8.1 Erhebungsdesigns	85
8.1.1 Ein-Gruppen-Designs	85
8.1.2 Mehr-Gruppen-Designs	92
8.2 Datenerhebungsmethoden	97
8.3 Erhebungsdesign und Datenerhebungsmethoden im Zusammenspiel	102

9	Durchführung der Erhebungen	105
10	Datenauswertung, Interpretation und Bewertungssynthese	111
11	Berichterstattung	117
11.1	Adressierte	117
11.2	Zeitpunkte	119
11.3	Verantwortung und Zuständigkeiten	120
11.4	Auswahl geeigneter Berichterstattungsformate	122
11.4.1	<i>Schriftliche Berichte</i>	123
11.4.2	<i>Mündliche Berichterstattungsformate</i>	126
12	Nutzung der Ergebnisse	129
13	Evaluation der Evaluation	137
	Literaturverzeichnis	143
	Verzeichnis der Abbildungen	157
	Verzeichnis der Tabellen	157
	Verzeichnis der Übungsaufgaben	158
	Onlineressourcen für die Evaluation	159
	Musterlösungen zu den Übungsaufgaben	161
	Glossar	184
	Stichwortverzeichnis	193

Vorwort

Evaluation hat in den letzten 20 Jahren auch im deutschsprachigen Raum eine steile Karriere gemacht. Ihre Wurzeln und großen Potenziale liegen insbesondere auch im Bildungsbereich. Evaluation kann helfen, Angebote vom Eltern-Kind-Turnen über den Computerkurs bis hin zur Facharzt-Weiterbildung und viele Hunderte mehr systematisch zu beschreiben und zu bewerten, meist mit dem Zweck, diese zu verbessern.

Diejenigen, die Bildungsprogramme finanzieren oder verantwortlich leiten, geben eine Evaluation in Auftrag, sei es an interne Mitarbeitende oder an Externe. Diese Evaluatorinnen und Evaluatoren müssen über einschlägiges Evaluationswissen und -können verfügen: wie sie zu relevanten und beantwortbaren Evaluationsfragestellungen kommen; wie sie Evaluationspläne an diese Fragestellungen anpassen; wie sie Evaluationskriterien festlegen; welche Rahmenbedingungen sie bei der Realisierung einer Evaluation zu beachten haben; welche methodischen Klippen sie umschiffen müssen; wie sie gewonnene Daten verarbeiten; wie sie die Nutzung der erzielten Ergebnisse fördern können und vieles Weitere mehr.

Dieses Buch ist ein praxisnaher Leitfaden für die Planung und Durchführung einer Bildungsevaluation und bietet Ihnen eine systematische Anleitung zum Erwerb bzw. Ausbau der genannten Kompetenzen. Wenn Sie sich mit Unterstützung dieses Buches erarbeiten, wie Sie nützliche und glaubwürdige Evaluationen planen und durchführen, dann bauen Sie auf Hunderte vielfach erfolgreicher, oft auch sehr herausfordernder Lerngeschichten auf. Personen, die mit verschiedenen Vorgängerversionen dieses Buches gearbeitet haben, hatten es mit unterschiedlichen Problemstellungen und Hindernissen der (Bildungs-)Evaluation zu tun, und konnten vielfach entsprechende Lösungen für diese finden. Die vorliegenden Texte und Aufgaben haben ihre Wurzeln in Veranstaltungsskripten und Studienbriefen (z. B. Balzer & Beywl, 2014; Beywl & Balzer, 2009). Diese sind in zahlreichen Aus- und Weiterbildungsveranstaltungen erprobt und laufend verbessert worden.

So manche Darstellung und Erläuterung ist erst auf Drängen von engagierten Teilnehmenden entstanden, die uns überzeugten, dass wir unsere theoretischen Vorstellungen zu und praktischen Erfahrungen mit Evaluation genauer ausführen müssen. Andere Passagen – insbesondere die Übungsaufgaben – sind nach kritischen Rückmeldungen der Teilnehmenden gründlich überarbeitet worden. Vielfach haben wir Verbindungen zwischen den Kapiteln stärker betont, um dem Anspruch, Ihnen Evaluation als eine systematische wissenschaftliche Vorgehensweise zu vermitteln, gerecht zu werden. Wir haben z. B. die Lösungen zu den Aufgaben inhaltsanalytisch ausgewertet und dabei beurteilt, in welchem Maße sie zur Lernzielerreichung beitragen und ggf. Verbesserungen vorgenommen.

Der Titel dieses Buches ist ein Vorgriff auf das, was Sie nach einer systematisch durchgeführten Evaluation erreicht haben, und markiert den Abschluss eines Dreischritts:

- **evaluiert?** – ein legitimer, vielfach notwendiger Checkpoint, der in der Lebensphase eines Bildungsprogramms anzuwenden ist
- **evaluiert!** – ein Planungsentschluss, der gefasst wird, wenn die Evaluation einer Bildungsmaßnahme als möglich und nutzenversprechend eingeschätzt wird
- **evaluiert.** – der Schlusspunkt hinter der Durchführung einer Evaluation, an den sich gelegentlich ein neuer Dreischritt anschließt.

Mit dem im Untertitel genannten «Bildungsbereich» sind zahlreiche auch benachbarte Felder angesprochen. Zum einen sind dies alle Bildungsstufen von der frühen Bildung über die allgemeinbildenden Schulstufen und die grundständige Berufsbildung bis hin zur hochschulischen Bildung sowie der allgemeinen und beruflichen Weiterbildung. Zum anderen sind dies Handlungsfelder der sozialen Arbeit, des Gesundheitsbereichs, der Verbraucherarbeit, der Beratung, des Coachings und viele andere mehr. Wichtig ist, dass es zumindest im weitesten Sinne um intentionales, also zielgerichtetes Lernen geht, das im Rahmen von Programmen stattfindet. Was hier unter einem Programm verstanden wird, schlagen Sie am besten gleich im Glossar am Ende des Buches nach – das Wort bzw. der Wortbestandteil wird über 400-mal in diesem Buch genutzt.

Adressaten dieses Buches sind alle, die sich die fachlichen Grundlagen der Evaluation erarbeiten wollen, um Evaluationen im Bildungsbereich Schritt-für-Schritt planen und schließlich erfolgreich umsetzen zu können. Angesprochen sind zum einen Studierende in den grundständigen Bachelor-, besonders den Masterstudiengängen im Bereich der Bildungswissenschaften und auch der sozialen Arbeit sowie ihrer Bezugsdisziplinen, u. a. Psychologie, Soziologie und Wirtschaftswissenschaft. Zum anderen richtet sich das Buch an Weiterbildungsstudierende in Programmen, die ein substanzielles Element zur Evaluation enthalten, z. B. im Bereich des Projekt-, Change- oder Qualitätsmanagements. Schließlich gehören auch Fachpersonen in Bildungsorganisationen, die ihr Wissen und Können im Prozess der Arbeit auf- und ausbauen wollen, zu den Hauptadressaten. Erfahrene Evaluierende können dieses Buch im Rahmen eines kombinierten Schulungs- und Beratungsangebotes einsetzen. Präsenzzeiten der Teilnehmenden können – wie von uns vielfach umgesetzt – durch die vorbereitende Lektüre und evtl. in Kombination mit der Bearbeitung der Übungsaufgaben reduziert werden. Durch ein solches Konzept eines «inverted classroom» kann diese Präsenzzeit effektiver für kollegiale Beratung und Coaching durch Evaluationsexpertinnen und -experten genutzt werden. Dabei kann z. B. eine reale Evaluation über zehn – jeweils durch ein Buchkapitel erläuterte – Schritte geplant und ggfs. danach tatsächlich umgesetzt werden.

Detaillierte Lernziele dieses Buches finden sich je zu Beginn der → [Kapitel 2–13](#). Durch ein intensives Studium dieses Buches können Sie u. a. ...

- wissenschaftliche Evaluation definieren,
- eine systematische Schrittfolge für eine für den Auftrag maßgeschneiderte Evaluation planen,
- den jeweiligen Evaluationsgegenstand differenziert beschreiben,
- die im Rahmen einer Evaluation zu berücksichtigenden Akteure identifizieren,
- Evaluationszwecke und -fragestellungen präzise formulieren,
- Kriterien für die systematische Bewertung klären und anwenden,
- geeignete Erhebungsdesigns auswählen und Erhebungsmethoden bestimmen,
- die Erhebungsarbeiten effizient und flüssig in den Bildungsalltag einbetten,
- Auswertungs- und Bewertungsarbeiten konzipieren und abstimmen,
- Evaluationsberichte auf die Adressierten zeitlich und inhaltlich abstimmen,
- die effektive Nutzung der Evaluation und ihrer Ergebnisse vorbereiten
- sowie Voraussetzungen für eine Evaluation der Evaluation – die sogenannte Meta-Evaluation – schaffen.

Die grundlegenden Kompetenzen der Datenerhebung – etwa in Bezug auf die drei Hauptgruppen Inhaltsanalyse, Beobachtung und Befragung – müssen Sie entweder bereits aus anderen Studien mitbringen oder durch Bearbeitung von

einschlägigen Lehrbüchern zu empirischen Methoden evtl. in Kombination mit entsprechenden Studienangeboten erwerben. Gleiches gilt für die qualitativen und quantitativen Auswertungsverfahren. In den entsprechenden Kapiteln dieses Buches machen wir Literaturempfehlungen, welche durch Internetquellen ergänzt werden.

Es empfiehlt sich, das Buch von vorne nach hinten durchzuarbeiten. Das → **Kapitel 1** enthält eine kleine Fallbeschreibung dazu, welche Herausforderungen sich schon ganz zu Beginn der Evaluation einer Bildungsmaßnahme stellen. Das Beispiel ist bewusst einfach gewählt. Dennoch entfaltet sich in seinen fiktiven Dialogen bereits die Komplexität, deren Bewältigung auch für erfahrene Evaluierende immer wieder eine große Herausforderung darstellt. In vielen der über 40 Beispielkästen der nachfolgenden Kapitel wird auf dieses Eingangsszenario zurückgegriffen.

Die folgenden → **Kapitel 2-13** führen in die Methodologie und Fachsprache der Evaluation ein. Vielfach wird in späteren Kapiteln das Wissen der vorangehenden vorausgesetzt. Eine Besonderheit besteht darin, dass alle Fachbegriffe nach bestimmten Regeln definiert sind und sie bei jeder Verwendung im Buch stets dieselbe Bedeutung haben. Sie finden die Begriffserklärungen im angehängten Glossar, das – wie einige andere im Literaturverzeichnis aufgeführte aktuelle Publikationen – auf dem noch umfangreicheren Online-Glossar des Evaluationsinstituts Univation basiert (vgl. Schobert, Beywl & Niestroj, 2015, im Erscheinen), welches unter http://eval-wiki.org/glossar/Eval-Wiki:_Glossar_der_Evaluation verfügbar ist.

Die meisten Kapitel enthalten Übungsaufgaben. Diese ermöglichen es, Ziele unterschiedlicher Anspruchsniveaus zu verfolgen, von der Erarbeitung des Begriffsverständnisses über die Anwendung von Wissen auf vorgegebene Fallbeispiele bis hin zum Transfer auf reale Praxisbeispiele aus den Bildungsfeldern, in denen Sie tätig sind. Im Anhang des Buches finden Sie passende Musterlösungen. Für Autodidakten empfiehlt es sich, alle Übungsaufgaben durchzuarbeiten. Wenn Sie in Lerngruppen arbeiten, versuchen Sie dies gemeinsam auch bei schwierigen Übungsaufgaben z.B. des Kapitels 7. Wenn Sie an einer Präsenzveranstaltung teilnehmen, in der Sie von Evaluationsexpertinnen und -experten begleitet werden, ist auf die Bearbeitung der Übungsaufgaben eventuell zu verzichten, da sie dann bei der Anwendung des Wissens und dem Transfer ins Praxisfeld direkt unterstützt werden können.

Zum Beleg von Originalaussagen werden in diesem Buch an verschiedenen Stellen fremdsprachige Textstellen zitiert. Zum leichteren Verständnis haben wir diese sinngemäß ins Deutsche übersetzt und die übersetzten Textteile in den Text integriert.

Sollten Sie bei der Lektüre des Buches auf Fehler oder unklare Passagen stoßen, schreiben Sie uns bitte eine E-Mail an: evaluiert@lars-balzer.info

Die Entstehung der Texte in diesem Buch geht teils zurück bis in die Mitte der 1990er-Jahre. Im Modellprogramm «Selbstevaluation in der Kinder- und Jugendhilfe» im Rahmen des Bundesjugendplans (D) fanden erste mehrwöchige Kurse statt. Michael Patton und Bob Stake, die nachfolgend mehrfach zitiert werden, führten damals einige Kurse durch, von deren Aufbau und Materialien sich Spuren auch noch in diesem Buch finden.

An der langen Entstehungsgeschichte des vorliegenden Buches haben viele Personen mitgewirkt. Seine ersten Weiterbildungsveranstaltungen zu Evaluation

konnte Wolfgang Beywl Ende der 1990er-Jahre in der «Werkstatt für Demokratie und Öffentlichkeit» in Bonn (D) sowie der Jugendakademie Vlotho (D) durchführen; es folgte eine Serie am Bundesinstitut für Erwachsenenbildung in Strobl (A), bevor 2002 die Kurse im Rahmen des heutigen Master-Programms «Evaluation» am Zentrum für universitäre Weiterbildung der Universität Bern (CH) begannen. Er lehrt seit 2011 im Master Educational Sciences (Universität Basel und Pädagogische Hochschule FHNW) und startete 2012 den Zertifikatslehrgang «Bildungsevaluation» an der PH FHNW am Campus Brugg-Windisch (CH). Lars Balzer begann 2001 mit Evaluationskursen an der Universität Koblenz-Landau (D). Weitere an der Pädagogischen Hochschule Heidelberg (D), am Eidgenössischen Hochschulinstitut für Berufsbildung (CH), an der Universität Bern (CH), an der Fachhochschule Nordwestschweiz (CH), an der Berner Fachhochschule (CH), am Institut de formation continue (LUX) sowie an der Universität Johannesburg (SA) folgten. Gemeinsam haben beide Autoren im Rahmen zweier Studiengangkohorten des «Masters in Management und Coaching im Bildungs- und Sozialwesen» der Universität Luxemburg in die Evaluation des Bildungs- und Sozialbereichs eingeführt, und dies – wie schon in einigen Fortbildungen zuvor – konsequent mithilfe der Vorgängerversionen des Buches, das Sie nun in der Hand halten. Von den Hunderten Teilnehmenden dieser Evaluationsweiterbildungen haben uns zahlreiche Personen wertvolle Rückmeldungen gegeben, uns auf Lücken und Schwächen der Erklärungen und der Übungsaufgaben hingewiesen. Es war ein Privileg, mit diesen zusammenzuarbeiten. Viele unserer Kolleginnen und Kollegen aus dem Eidgenössischen Hochschulinstitut für Berufsbildung EHB IFFP IUFP, dem Institut Weiterbildung und Beratung der Pädagogischen Hochschule FHNW sowie von Univation in Köln haben uns ebenfalls inspiriert, direkte Rückmeldungen oder auch nur ganz nebenbei wichtige Hinweise gegeben, von denen hoffentlich auch Sie als Leserinnen und Leser für Ihre Evaluationspraxis profitieren können.

Wolfgang Beywl und Lars Balzer, Bern und Zollikofen, im Mai 2015

Einstieg ins Thema

KAPITEL 1

Was ist eigentlich eine Evaluation?

Stellen Sie sich einen Weiterbildungskurs zum Thema «Wie erstelle ich meine erste Homepage?» vor, in dem an zwanzig Abenden Basiswissen über die Erstellung von Internetseiten vermittelt wird.

Eine kleine Fallgeschichte

Nachdem dieser Kurs als Pilotprojekt gestartet ist, fragen sich die Weiterbildungsverantwortlichen in einem Teamgespräch, wie gut ihre Arbeit eigentlich ist. Nach Kursende gehen die Teilnehmenden ihrer Wege, man trifft sie nicht mehr, ob mehr Wissen vorhanden ist als zuvor und ob tatsächlich eigene Internetseiten erstellt worden sind, ist den Weiterbildungsverantwortlichen normalerweise nicht bekannt. «Wir brauchen eine Evaluation, um zu überprüfen, ob wir gute Arbeit leisten», schlägt einer der Verantwortlichen vor. Bevor sie einen Evaluationsauftrag an eine interne oder eine externe Evaluationsfachperson formulieren, diskutieren sie die Möglichkeiten.

Herr Schmidt ist von dieser Idee sogleich begeistert: «Lasst uns doch am Ende des Kurses unsere Teilnehmenden fragen, ob sie mit uns und unserer Arbeit zufrieden sind. Wenn z. B. 85 Prozent derselben mit unseren Veranstaltungen zufrieden sind, gerne zu weiteren Veranstaltungen kommen würden und uns weiterempfehlen, können wir doch mit Fug und Recht sagen, dass wir Gutes geleistet haben.» – «Das ist nicht schlecht», ist man sich schnell einig, «doch fast hilfreicher und interessanter wäre zu wissen, was den restlichen 15 Prozent oder auch den Zufriedenen nicht gefallen hat, damit wir uns weiter verbessern können!»

Erfolg = Zufriedenheit?

«Darüber hinaus», stellt Frau Zbinden fest, «sollten wir mehr wissen als nur, wie zufrieden unsere Teilnehmenden sind. Ich glaube, bei uns geht es auch mal ganz lustig zu und her, und das gefällt allen. Zufriedenheit ist wichtig, denn das erhöht die Chance, dass unsere Kunden wiederkommen. Aber hauptsächlich sollen sie doch etwas lernen.» Kopfnicken macht die Runde. «Wir könnten unsere Teilnehmenden also ergänzend befragen, ob sie etwas gelernt haben.» – «Wobei: Können das unsere Teilnehmenden wirklich selbst beurteilen? Wie wäre es anstelle dessen mit einem Wissenstest am Ende der letzten Stunde? So könnten sie sich einerseits selbst überprüfen, und wir könnten andererseits erkennen, was von den Inhalten tatsächlich behalten worden ist!» – «Ja, das würde uns sicher helfen», ist die einhellige Meinung.

Erfolg = kurzfristiges Wissen?

«Aber was heißt eigentlich ›behalten‹?», wirft Frau Lavric ein. «Ein Wissenstest im Anschluss an eine intensive Fortbildung überprüft vielleicht das Kurzzeitgedächtnis oder den Fleiß, also wie intensiv sich jemand auf die Abschlussprüfung vorbereitet hat. Ist aber nicht vielmehr interessant, was längerfristig noch gewusst wird?» Ein Hauch von Verzweiflung macht sich breit. «Sollen wir unsere Teilnehmenden vielleicht zwei Monate später noch einmal testen?», fragt Herr Schmidt ungläubig, gar nicht mehr so begeistert wie noch zu Beginn der Diskussion. «Und wie sollen wir sie dazu motivieren? Der ganze Aufwand!» Ein leises Stöhnen in der Runde ist nicht zu überhören.

Erfolg = nachhaltiges Wissen?

«Es wird noch viel schwieriger, bei genauer Überlegung.» Frau Lavric spricht sich in Fahrt: «Selbst wenn nach zwei Monaten noch alles richtig gewusst wird, ist das zwar schön, aber unser eigentliches Ziel ist doch, dass die Teilnehmenden nicht nur bei einem Wissenstest gut abschneiden, sondern dieses Wissen auch tatsächlich anwenden. Denn selbst wenn der Wissenstest zu einem späteren Zeitpunkt erfolgreich ablaufen sollte: Wer sagt uns, dass jemand auch vernünftige In-

Erfolg = Anwendung?

1

ternetseiten erstellen kann? Und vielleicht auch tatsächlich solche erstellt hat? Wissensvermittlung ist nur ein Zwischenziel; schöne, funktionelle und suchmaschinenoptimierte Internetseiten sind hingegen das, was wir eigentlich erreichen wollen.» – «Stimmt, also lasst uns doch die Teilnehmenden zwei Monate nach Abschluss des Kurses anrufen und fragen, ob sie Internetseiten erstellt haben.» Damit ist Herr Malte überhaupt nicht einverstanden: «Das ist zwar eine gute Idee, aber zwei Monate sind eine zu kurze Zeitspanne, und einfach ein paar Internetseiten zu erstellen, ist keine große Herausforderung. Wir müssen die Internetseiten anschauen und selbst nach gewissen Usability-Kriterien überprüfen. Also müssen wir nach den Internetadressen fragen.» – «Das stimmt. Und vielleicht können wir uns schon ein wenig auf die Schultern klopfen, wenn die einfachsten Regeln der Internetseitenprogrammierung eingehalten worden sind. Und wer weiß, vielleicht haben wir nach einer gewissen Zeit ja auch eine Suchmaschinen-Top-Ten dabei?!» – «Das Problem ist nur: Wie bekommen wir das heraus?» – «Was wir also brauchen, ist eine Datenerhebung längere Zeit nach Ende unserer Veranstaltung.» Erneutes Kopfnicken. «Hiermit würden wir weiterkommen!»

Was genau löst den Erfolg aus?

«Aber was wüssten wir dann eigentlich genau?», meldet sich Frau Lavric wieder zu Wort: «Selbst wenn wir das alles in Erfahrung bringen können und feststellen sollten, dass 46 Prozent unserer Teilnehmenden gute Internetseiten erstellt haben: Was bedeutet ein solcher Prozentsatz? Sind das viele, sind das wenige? Wie viele Personen hätten auch ohne Weiterbildung Internetseiten erstellt? Ihr wisst, das Thema ist aktuell, und überall gibt es Do-it-yourself-Bausätze. Könnten wir also mit 46 Prozent zufrieden sein? Und hat unsere Weiterbildung eigentlich einen Anteil am möglichen Erfolg?» – «Wir brauchen einen Vergleichsmaßstab», ist die einstimmige Erkenntnis. Doch das ist leichter gesagt als getan. «Wie wäre es, wenn wir uns mit unseren Kolleginnen und Kollegen in Düsseldorf vergleichen würden? Die haben ein sehr ähnliches Konzept und eine ähnliche Klientel.» – «Das führt uns zwar weiter, aber ob wir absolut betrachtet Erfolg haben, wissen wir nicht. Vielleicht sind wir besser als das Team in Düsseldorf, aber was nutzt das, wenn wir alle eigentlich recht schlecht sind? Umgekehrt könnte ich gut damit leben, im Vergleich etwas schlechter zu sein, aber absolut betrachtet ein gutes Weiterbildungsprogramm anzubieten.»

Wie kann man Wirksamkeit messen?

Jemand hat eine Idee: «Wir brauchen eine andere Personengruppe, die unsere Weiterbildung nicht besucht hat. Dann vergleichen wir die Internetseiten unserer Teilnehmenden mit denjenigen besagter Gruppe. Fällt der Vergleich positiv aus, können wir doch sagen, dass unsere Weiterbildung Erfolg hatte, oder?» Zufriedenheit ist auf den Gesichtern abzulesen. «Und wen nehmen wir als Vergleichsgruppe?» – «Wie wäre es mit den Teilnehmenden des Englischkurses? Die erfahren dort sicher nichts über Internetseitenerstellung.» – «Das stimmt zwar», wirft einer ein, «aber ist das wirklich ein fairer Vergleich? Unsere Teilnehmenden sind motiviert, Internetseiten zu erstellen. Die Teilnehmenden des Englischkurses teilen dieses Interesse nicht. Also könnte es auch sein, dass ein möglicher Erfolg auf diese Motivation zurückzuführen ist und sich dieser auch ohne unseren Kurs eingestellt hätte.»

Systematische Evaluationsplanung ist erforderlich

Ein Kollege mit Wissen über sozialwissenschaftliche Forschung hat dazu einen Einfall: «Wir brauchen eine Vergleichsgruppe, die sich hinsichtlich aller für das Erreichen der Weiterbildungsziele wichtigen Bedingungen wie Motivation oder Vorwissen nicht von den Teilnehmenden unserer Kurse unterscheidet. Eine solche Gruppe können wir ganz leicht auftreiben. Wir haben doch ohnehin mehr Anmeldungen, als wir in einem Monat bedienen können. Wenn wir von den nächsten 100 Anmeldungen 50 zufällig aussuchen, die wir sofort aufnehmen,

und die anderen eben ein wenig später, haben wir die Gruppe bereits gebildet. Nach den Gesetzen der Wahrscheinlichkeit haben wir nämlich eine gute Chance, durch diese zufällige Zuordnung die relevanten Eigenschaften, die wir nicht einmal kennen müssen, gleich auf beide Gruppen zu verteilen. Beide Gruppen unterscheiden sich dann in nichts außer in der Tatsache, dass die eine an der Weiterbildung teilnimmt und die andere eben nicht. Ist unsere Gruppe dann besser, können wir recht sicher sein, dass wir mit unserer Weiterbildung einen Effekt erzielen.»

Doch der nächste Einwand folgt sofort: «Moment! Ich glaube nicht, dass wir lange genug warten können mit unserer Vergleichsgruppe, um längerfristige Resultate überprüfen zu können. Die wollen doch bald loslegen, und wer weiß, was sie in der Wartezeit unternehmen, um schnell ihre gewünschten Internetseiten erstellen zu können.»

«Eigentlich ist mir das zu kompliziert. Die Überprüfung des Erfolges unserer Weiterbildung ist schon wichtig, aber eigentlich hätte ich einfach gerne ein paar Informationen für mich und meine weiteren Planungen, sodass ich von den Teilnehmenden gerne gewusst hätte, was sie für konkrete Verbesserungsvorschläge haben. Es gibt schließlich keine Weiterbildungsmaßnahme, die so gut wäre, dass man sie nicht noch verbessern könnte!»

«Oh Gott, auf was lassen wir uns mit einer solchen Evaluation nur ein? Wie soll eine solche Evaluation irgendwann einmal abgeschlossen werden?»

Es bietet sich an, einen Evaluationsauftrag an hierfür qualifizierte interne Mitarbeitende oder Externe zu vergeben, die über Wissen und Können verfügen, wie sie zu beantwortbaren Evaluationsfragestellungen kommen, wie sie Evaluationspläne an diese Fragestellungen anpassen, welche Rahmenbedingungen sie bei der Realisierung einer Evaluation zu beachten haben, welche methodischen Klippen sie umschiffen müssen, wie sie gewonnene Daten verarbeiten, was mit den erzielten Ergebnissen zu geschehen hat und vieles Weitere mehr.

Dieses Buch bietet Ihnen eine systematische Anleitung zum Erwerb der hierfür erforderlichen Kompetenzen.

KAPITEL 2

Was ist Evaluation?



Lernziele von Kapitel 2:

- Den Begriff der wissenschaftlichen Evaluation definieren können
- Evaluation von Forschung abgrenzen können
- Werte und Bewertung als Besonderheiten von Evaluation an Beispielen erläutern können

Auf dem Weg zur Definition

Auf die Frage, was Evaluation genau ist, gibt es keine einfache Antwort. Nach wie vor dürfte die These von Franklin und Thrasher (1976) gelten, wonach es so viele Evaluationsdefinitionen wie Evaluierende gibt¹: «To say that there are as many definitions as there are evaluators is not too far from accurate» (S.20).

Evaluation wird nach Glass und Ellet (1980, S.211) sehr unterschiedlich definiert und abgegrenzt – und das gilt heute noch ebenso wie damals. Viele theoretische wie auch praktische Aktivitäten werden darunter subsumiert, ohne dass ein allgemeingültiges Evaluationsparadigma vorherrschen würde: «Evaluation is a set of theoretical and practical activities without a widely accepted paradigm.» Und je nach Zuordnung zu einer bestimmten wissenschaftlichen Disziplin wird der Begriff zudem unterschiedlich akzentuiert (Götz, 1998, S.20). Erschwerend kommt hinzu, dass «allzu leicht [...] heute auch triviale Formen der Rückmeldung zu Evaluationen» werden (Böttcher, Kerlen, Maats, Schwab & Sheikh, 2014, S.7).

Diese beinahe beliebige Möglichkeit der Inanspruchnahme macht Evaluation für viele attraktiv, bietet aber auch schier unüberschaubare Ansatzpunkte für Kritik und Polemik. Unterschiedliche Ausdrücke und Aussagen werden in diesem Zusammenhang verwendet und diskutiert, wie «Evaluation als modernes Ritual» (Schwarz, 2006), Evaluation als notwendiges Übel (evaluation as a «necessary evil», Sullivan, 2011), «Evaluations Mania» (Frey, 2007a), Evaluation als sich epidemisch ausbreitende, neue Krankheit, die «Evaluitis» (Frey, 2007b, 2008), und «von der Inquisition zur Evaluation» (Hornbostel, 2008). Es werden sogar Warnungen ausgesprochen: «Vorsicht vor Evaluationismus!» (Kappler, 2010). Oder ist es gar so, dass wir uns zu Tode evaluieren, wie Preußler (2008) fragt – oder dass ein «Evaluationsnotstand» herrscht (Niedermaier, 2012, S.8)?

Für einen differenzierten Umgang mit diesem vermeintlichen Unwort ist es notwendig, einige Definitionen vorzustellen und eine Arbeitsdefinition für dieses Buch zu formulieren.

Wortstamm

Auf der Suche nach einer adäquaten Definition gibt der Wortstamm einen ersten Hinweis. Auch wenn eine unmittelbare Herleitung aus dem Lateinischen nicht angebracht ist (entgegen der weitverbreiteten Meinung existiert das Wort «evaluare» im Lateinischen nicht), legt das Lateinlexikon erste Spuren. Wurzeln finden sich nämlich im lateinischen Wort «valor», das im Deutschen so viel bedeutet wie «bei Kräften sein», «wert sein» oder «gültig sein». Man beachte hierbei die eindeutig positive Konnotation.

¹ Zum Beleg von Originalaussagen werden in diesem Buch an verschiedenen Stellen fremdsprachige Textstellen zitiert. Zum Verständnis ist die Kenntnis der anderen Sprache allerdings nicht erforderlich, da wir alle fremdsprachigen Zitate mindestens sinngemäß ins Deutsche übersetzt und die übersetzten Passagen in den erläuternden Text um das Zitat herum integriert haben.

Den etymologischen Herleitungen des Begriffes von Karbach (1998) folgend, entwickelte sich daraus zunächst das französische «valoir», woraus die Substantivierung «valeur» (im Sinne vom «prix», also auch Wert) entstand. Daraus wurde wiederum das Verb «évaluer» abgeleitet und von diesem das Substantiv «évaluation» («Schätzung», «Ermittlung» oder «Wertbestimmung») gebildet.

Die daraus hervorgegangenen englischen Wörter «evaluate» («bewerten») sowie «evaluation» («Einschätzung», «Auswertung») bilden die Grundlage für die heute im Deutschen gebräuchliche Form des Begriffes.

Eine erste, vom Wortstamm ausgehende Umschreibung von Evaluation lautet also:

«Bestimmung des Wertes einer Sache»
(Bedeutung nach Wortstamm).

Folgt man dieser Bestimmung, so ist Evaluation – zumindest im französischen oder angelsächsischen Sprachraum – eine Bezeichnung für alltägliches menschliches Handeln. Denkend oder sprechend wird auf Basis eines Sinneseindrucks, z. B. des Blicks aus dem Fenster oder der herausgehaltenen Hand, ein Urteil – hier: über das Wetter – abgegeben. Es handelt sich um eine einfache Alltagsbewertung. Obwohl es sich auch in der deutschen Alltagssprache – z. B. in Tageszeitungen, Fernsehinterviews oder Talkrunden – seit einigen Jahren zu etablieren begonnen hat, solche subjektiven Ad-hoc-Bewertungen als «Evaluationen» zu bezeichnen, plädieren wir dafür, «Evaluation» und «evaluieren» für das wissenschaftlich abgestützte, systematische Beschreiben und Bewerten zu reservieren. Was darunter zu verstehen ist, wird nachfolgend präzisiert.

2.1 Wissenschaftliche Evaluation statt Alltagsbewertung

Kromrey (2001) unterscheidet den alltäglichen und den wissenschaftlichen Sprachgebrauch von Evaluation danach, was von wem wie und nach welchen Kriterien bewertet wird. Auf dem Weg zu unserer Definition von Evaluation für dieses Buch gilt es nun, diese und andere relevante Dimensionen zu beschreiben.

Zunächst gilt es zu klären, wer oder was evaluiert werden soll. Dies wird als Evaluationsgegenstand bezeichnet.

Die Menge an potenziellen Evaluationsgegenständen ist beinahe unüberschaubar. Wottawa und Thierau (2003, S. 59) nennen «Personen, Umwelt-/Umfgebungs-faktoren, Produkte, Techniken/Methoden, Zielvorgaben, Programme, Projekte, Systeme/Strukturen, Forschungsergebnisse/Evaluationsstudien» und haben der Evaluation damit ein bereits sehr breites Tätigkeitsfeld eröffnet. Scriven (1981, S. 4) weitet dieses noch aus, indem er jedes Hauptwort eines Wörterbuches zu einem möglichen Evaluationsgegenstand macht: «One can begin at the beginning of a dictionary and go through to the end, and every noun, common or proper, readily calls to mind a context in which evaluation would be appropriate.» Cook und Matt (1990, S. 15) bringen es auf den Punkt: «Alles kann evaluiert werden.»

Entscheidend für die Abgrenzung zur Alltagsbewertung ist, dass Evaluationen zur Bestimmung des Wertes (Güte und Tauglichkeit) eines Evaluationsgegenstandes systematisch, umfassend und objektiv durchgeführt werden sollen, was folgende Definitionen festhalten:

Wer oder was? – Der
Evaluationsgegenstand

Wie? – Die Evaluations-
methoden