

Rezension zu:

Balzer, Lars/Beywl, Wolfgang: *evaluiert. Planungsbuch für Evaluationen im Bildungsbereich*. Bern: hep verlag, 2015 (1. Auflage). 200 Seiten, 27,00 EUR, ISBN 978-3-0355-0339-5

*Maria Gutknecht-Gmeiner*¹

1. Autoren, Buch und Zielgruppe

Wolfgang Beywl und Lars Balzer sind beide bekannt als Evaluationsforscher. Ihr neues gemeinsames Buch „evaluiert. Planungsbuch für Evaluationen im Bildungsbereich“ basiert daher auf jahre- bzw. jahrzehntelanger Beschäftigung mit dem Thema Evaluation sowie auf einschlägiger Lehrtätigkeit. Es ist aus Veranstaltungsskripten und Studienbriefen hervorgegangen, hat mehrere Vorgängerversionen und wurde immer wieder in der Praxis erprobt und weiterentwickelt. Auch Evaluationsforscher wie Michael Patton und Bob Stake haben, wie die Autoren selbst sagen, „Spuren“ hinterlassen – dies wird im Buch auch an dem stark nutzenorientierten Zugang zu Evaluation, der Einbeziehung von Stakeholdern sowie der gleichwertigen Behandlung qualitativer Methoden erkennbar.

Das Buch versteht sich nicht nur als Einführung in die Evaluation sondern „ist ein praxisnaher Leitfaden für die Planung und Durchführung einer Bildungsevaluation und bietet Ihnen eine systematische Anleitung zum Erwerb bzw. Ausbau der genannten [Evaluations] Kompetenzen“ (S. 7). Die Autoren haben sich also einen expliziten Bildungsauftrag gegeben, sie wollen einschlägiges Evaluationswissen und -können vermitteln, wobei es um die fachlichen Grundlagen geht (Selbst- und Sozialkompetenzen z.B. werden nicht angesprochen). Das gesamte Buch ist didaktisch aufgebaut, weshalb auch empfohlen wird, die Kapitel in der angebotenen Reihenfolge zu lesen oder besser gesagt: durchzuarbeiten. Das Buch richtet sich v.a. an Studierende sowohl im Erststudium als auch in Weiterbildungsstudiengängen, in denen es einen Bezug zur Evaluation gibt, aber auch an

Praktiker(innen) in Bildungsorganisationen, die ihr Wissen zu Evaluation erweitern wollen.

2. Inhalte des Buchs/Überblick

Das Buch ist in 13 Kapitel gegliedert, wobei die ersten beiden Kapitel eine Sonderstellung einnehmen: als Auftakt wird in Kapitel 1 mit einer aus der Praxis von Bildungseinrichtungen gegriffenen *Fallbeschreibung*² das Feld aufgemacht. Es geht dabei um die Evaluation eines Kurses, der Basiswissen zur Erstellung von Internetseiten vermittelt. Dabei wird rasch ersichtlich, wie komplex Evaluation sein kann und welche Überlegungen bereits im Vorfeld angestellt werden müssen, um zweckmäßige und brauchbare Evaluationen zu gestalten. Im Dialog zwischen den für den Kurs Verantwortlichen werden die wichtigsten Fragestellungen (in den ersten Schritten angelehnt an die 4-Ebenen-Systematik von Kirkpatrick (Kirkpatrick/Kirkpatrick 2006) angesprochen, denen sich die Evaluation von Programmen im Bildungsbereich stellen muss. Auf die Fallbeschreibung wird in weiterer Folge immer wieder Bezug genommen.

Kapitel 2 bietet eine recht ausführliche *Einführung*, in der Evaluation definiert und das (Be-)Werten als das besondere Charakteristikum von Evaluation herausgestrichen wird. Es wird wissenschaftliche Evaluation einerseits von Alltagsbewertungen, andererseits von (Grundlagen-)Forschung unterschieden. Dabei werden auch die Begrifflichkeiten geklärt, wobei die Autoren den Begriff ‚Evaluationsforschung‘, der lange als Synonym für Evaluation galt, nun als Forschung über/zu Evaluation definieren. Der Dienstleistungscharakter, der Nützlichkeitsprimat sowie die gesellschaftliche

1 Impulse – Evaluation und Organisationsberatung, Wien

2 Das Fallbeispiel in Kapitel 1 findet sich ähnlich in Balzer 2005: 1ff.

Relevanz von Evaluation werden hervorgehoben. Von Qualitätsmanagement und Controlling wird Evaluation als eigener Ansatz abgegrenzt.

In den nächsten 11 Kapiteln wird dann der gesamte Evaluationsprozess nachgezeichnet. Dies beginnt in Kapitel 3 mit der *Auftrags- und Planungsphase*. Die Autoren argumentieren für „maßgeschneiderte“ Vorgehensweisen anstatt standardisierter Evaluationen „von der Stange“, um dem jeweiligen Evaluationsgegenstand und den Informationsbedürfnissen der Betroffenen und Beteiligten gerecht werden zu können. Sie empfehlen eine gründliche Klärung des Evaluationsauftrags sowie eine systematische Abfolge von zehn Schritten für die Planung und Umsetzung einer Evaluation, die in den darauffolgenden Kapiteln näher erläutert werden.

Kapitel 4 widmet sich der *Bestimmung des Evaluationsgegenstandes*, d.h. seiner Eingrenzung (nicht zuletzt in Abhängigkeit von den Fragestellungen), der Klärung der Ziele bzw. des Zielsystems des Evaluationsgegenstands sowie der Erarbeitung und Nutzung eines logischen Modells. Hier wird auch der Programmbaum als anschauliches und einfach einzusetzendes Modell dargestellt.

Um die *Bestimmung der interessierten Akteure und die Rolle der Evaluierenden* geht es in Kapitel 5, ein wichtiger Schritt, um der Einbettung von Evaluation in einen sozialen Zusammenhang gerecht zu werden. Das Erkennen der verschiedenen Betroffenen und Beteiligten, die Interessens- und Rollenklärung zwischen den verschiedenen Stakeholdern sowie die Notwendigkeit der (möglichst frühen und aktiven) Einbindung des Evaluationsumfeldes sind daher wichtige Themen des Kapitels. Ausführlich wird auf die verschiedenen Aufgaben und Rollen der Evaluierenden eingegangen sowie auf die sich daraus ergebenden (Kompetenz-)Anforderungen. Dabei werden die verschiedenen Evaluationsarten und ihr Einfluss auf die Position und den Auftrag der Evaluierenden vorgestellt.

Als nächster Schritt kommt es in Kapitel 6 zur *Bestimmung der Evaluationszwecke und -fragestellungen*. Es wird ausführlich auf Evaluationszwecke sowohl in zeitlich-funktionaler Hinsicht als auch in Bezug auf den Verwendungszusammenhang – hier wird die Trias Verbesserung, Entscheidung, Wissensgenerierung erläutert – eingegangen und zum besseren Verständnis das Zusammenspiel dieser verschiedenen Funktionen, Zwecke und ‚Rollen‘ (vgl. Definition im „Glossar der Evaluation“, verfügbar unter: <http://eval-wiki.org/glossar/Evaluationsrolle>, letzter Zugriff: 01.07.2016) dargestellt, wobei ‚formativ‘ und ‚summativ‘ als Schlüssel-

begriffe nur en passant eingeführt werden. Dann wird das Finden und Formulieren von Fragestellungen als zentrales Element einer Evaluation thematisiert. Als Anregung und erste Orientierung gibt es eine Übersichtstabelle mit ‚generischen‘ Fragestellungen für Bildungsprogramme. Sehr positiv ist der Hinweis auf die verschiedenen Arten von Fragestellungen – von rein deskriptiven, über evaluative zu attributiven. Diese Unterscheidung fehlte bislang in deutschsprachigen Einführungswerken zu Evaluation, trägt jedoch zur Klarheit im Evaluationsauftrag und -prozess entscheidend bei.

Der *Festlegung von Bewertungskriterien* ist ein eigenes Kapitel (7) gewidmet. Dies verdeutlicht, wie wichtig den Autoren die Bewertungsfunktion von Evaluation ist: Sie sehen Bewertung als den „professionelle[n] Kern der Evaluation und auch ihr Alleinstellungsmerkmal gegenüber anderen Ansätzen wie dem Controlling oder dem Qualitätsmanagement“ (S. 82). Kriterien sowie Kriterienpunkte (Schwellenwerte) sind daher aus Sicht der Autoren „unverzichtbar“ (S. 76) und sollten im Vorfeld festgelegt werden. Die Autoren nehmen keine normative Setzung vor, wer im Evaluationsprozess Kriterien bestimmen und/oder die Bewertung vornehmen darf/soll, sondern zeigen die verschiedenen Möglichkeiten auf – von der alleinigen Verantwortlichkeit der Evaluierenden über den Einbezug von Betroffenen und Beteiligten des evaluierten Programms bis zur Ausklammerung der Bewertung aus dem eigentlichen Evaluationsprozess. Es wird mit dem nicht einlösbaren Anspruch der „absoluten Objektivität“ aufgeräumt und die Notwendigkeit eines planvollen, systematischen, methodisch kontrollierten und nachvollziehbaren Vorgehens bei der Bewertung eingefordert.

Das Buch kommt ohne detaillierte Beschreibungen von Datenerhebungsmethoden und -instrumenten aus (siehe unten). Sehr wohl aber wird in Kapitel 8 die *Auswahl von Erhebungsdesigns und -methoden*, die jeder Evaluationsplanung zugrunde liegt, ausführlich erläutert. So werden die verschiedenen quantitativen Erhebungsdesigns – mit Ein-Gruppen- und Mehr-Gruppen-Designs und verschiedenen Messzeitpunkten – beschrieben und ihre Aussagekraft und Einsatzmöglichkeiten diskutiert. Dann gehen die Autoren auf qualitative und quantitative Erhebungsmethoden ein. Sie legen jeweilige Stärken und Schwächen dar und – daraus folgend – ihre Eignung für verschiedene Evaluationszwecke und Fragestellungen. Entscheidend ist es, aus Sicht der Autoren, eine sorgsame, für

den jeweiligen Kontext passende Design- und Methodenwahl zu treffen.

Der *Durchführung der Erhebungen* ist Kapitel 9 gewidmet, dabei wird vor allem auf eine gute Einbindung in den Alltag des Bildungsprogramms/der Bildungseinrichtung hingewiesen sowie auf den Umgang mit Störungen und Konflikten. In Kapitel 10 geht es um die *Datenauswertung, Interpretation und Bewertungssynthese*. In Bezug auf quantitative und qualitative Auswertungsmethoden verweisen die Autoren wieder auf vorhandene Literatur. Sie widmen sich zuerst den Grundlagen jeder Analyse: der Auswertungsplanung sowie der Beschreibung von Datenquellen, Stichprobe(n) und Erhebungsdurchführung. Den folgenden Prozess der eigentlichen Auswertung und Interpretation gliedern sie in elf an Carol Weiss (1998) angelehnte Analyseschritte, die für qualitative und quantitative Verfahren gelten, verschiedene Analysegesichtspunkte beschreiben und in denen das Beschreiben und ‚Zählen‘ nur den Anfang darstellen. Das Kapitel endet mit einer kurzen Behandlung der Bewertungssynthese, die es ermöglicht, aus den Einzelbewertungen von Kriterien eine Gesamtbewertung eines Evaluationsgegenstands zu ziehen.

Es folgt in Kapitel 11 die *Berichterstattung*, die auf die verschiedenen Personen(gruppen), an die sich die Berichterstattung wenden kann („Adressierte“, S. 117), zugeschnitten werden muss und rechtzeitig zu erfolgen hat. Verantwortung und Zuständigkeiten werden noch einmal geklärt, insbesondere was das Schlussfolgern und Bewerten angeht sowie das Formulieren (möglicher) Empfehlungen und die daraus folgenden Entscheidungen über die weitere Programmumsetzung/-entwicklung (Letztere werden eher nicht bei den Evaluierenden verortet.) Schließlich werden (und auch das wurde in der bisherigen deutschsprachigen Literatur kaum behandelt) die verschiedenen Berichtsformate, unter denen schriftliche Berichte (für die Qualitätsanforderungen definiert werden) nur eine Möglichkeit darstellen, ausführlich beschrieben.

Der *Nutzung der Ergebnisse* ist Kapitel 12 gewidmet. Die Autoren gehen kurz auf den Prozessnutzen von Evaluation ein, um sich dann v.a. der instrumentellen Nutzung von Evaluationen zuzuwenden. Eine erfolgreiche Nutzung der Ergebnisse wird im Sinne der Nutzungsorientierung als Teil des Evaluationsprozesses gesehen, auf den auch die Evaluierenden einen Einfluss haben – und nicht als etwas, was rein nachgelagert erfolgt. Sie erfordert eine vorausschauende Planung, zentraler Faktor ist der Einbezug der Betroffenen und Beteiligten. Verbreitung und Rezeption werden als notwendige, aber nicht

hinreichende Bedingungen für die Nutzung gesehen. Auch auf die Gefahren der Nichtnutzung bzw. des Missbrauchs weisen die Autoren hin.

Mit der geforderten Dokumentation der Nutzung leiten die Autoren über zu Kapitel 13, das sich der *Evaluation der Evaluation* annimmt. Nach einer Unterscheidung zwischen Evaluations-synthese, Meta-Analyse und Meta-Evaluation werden die „Standards für Evaluation“ als Kriterien für die Meta-Evaluation präsentiert. Die Meta-Evaluation als Schlussphase des Evaluationszyklus wird v.a. in Hinblick auf die Qualitätssicherung argumentiert, hier hätten die Autoren noch stärker auf die interne oder externe Weiterentwicklung des Vorgehens bei der Evaluation verweisen können.

Im Anhang des Buches finden sich ein ausführliches Literaturverzeichnis sowie Hinweise zu Onlineresourcen, Musterlösungen zu den Übungsaufgaben, ein Glossar, das auf dem von Univation erstellten Online-Glossar (vgl. eval-wiki.org/glossar/Eval-Wiki:_Glossar_der_Evaluation) basiert, sowie ein Stichwortverzeichnis.

3. Diskussion

Das Buch besticht durch seine gute theoretische Fundierung, die auch die internationalen Status quo der Evaluationsforschung berücksichtigt, und die praxisorientierte und eingängige Aufbereitung. Es zeichnet sich dadurch aus, dass die Spezifika der Evaluation sehr klar herausgearbeitet und die komplexen Anforderungen an Evaluation und Evaluator(innen) in gut verständlicher Art und Weise präsentiert werden.

Als roter Faden ziehen sich das Bewerten und die Nutzungsorientierung als grundlegende Merkmale von Evaluation sowie die Einbettung von Evaluation in einen sozialen Kontext durch das Buch. Darin hebt es sich positiv ab von anderen, ähnlichen Einführungen in die Evaluation, in denen diese Dimensionen oft wenig oder gar nicht thematisiert werden. Auffällig ist auch, dass es kein Kapitel mit Beschreibungen gängiger Methoden gibt. Die Autoren selbst geben als Begründung an, „dass es sehr viel gute Literatur gibt, die die unterschiedlichen Erhebungsmethoden ausführlich beschreibt. Diese Publikationen können für den Evaluationsprozess verwendet werden“. Stattdessen werden Erhebungsdesigns und Auswahl von Methoden intensiver behandelt, die aus Sicht der Autoren von entscheidender Bedeutung sind, oft aber „zu wenig Beachtung finden“ (S. 84). Dieser Fokus ist eine sehr sinnvolle Entscheidung, vermeidet er doch Redundanzen zu bestehender Methodenliteratur (die zudem ausführlicher sein kann als ein Kapitel in einem Evaluationslehrbuch), während er Lücken

in Bezug auf die grundlegenden Fragen der empirischen Vorgehensweise schließt.

Die Entscheidung, die Methoden auszuklammern, hat aber aus Sicht der Rezensentin zusätzlich Signalwirkung: Sie macht deutlich, dass Evaluation mehr ist als sozialwissenschaftliche Methoden anzuwenden.³ Mit dem Planungsbuch von Balzer und Bewyl gibt nun ein Werk, in dem es hauptsächlich darum geht, darzulegen, was Evaluation – über die Anwendung sozialwissenschaftlicher Methoden hinaus – ausmacht. Und das ist doch einiges: immerhin 142 Seiten haben die beiden Autoren dazu gefüllt. Gleichzeitig wird damit, so das Buch als Evaluationsleitfaden genutzt wird, der gerade bei Evaluationsnoviz(inn)en beobachtbaren Fixierung auf Erhebungsinstrumente und (unbedachte) Aktion vorgebeugt, zugunsten eines planvollen, analytischen und systematischen Vorgehens.

Was das Buch noch bereichern könnte, wäre eine kurze Abhandlung zu qualitativen Designs. Auch die Verwendung eines Mixed-Methods-Ansatzes ist aus Sicht der Rezensentin eine Designfrage (und nicht nur eine Methodenfrage) und sollte dort diskutiert werden.

Die didaktische Aufbereitung des Buches ist mit Lernzielen zu Beginn jedes Kapitels, vielen Beispielen (insgesamt 42) und Übungsaufgaben (mindestens eine pro Kapitel) sowie Angaben zu vertiefender Literatur sehr gut gelungen. Es ist ein Lehrbuch, wie man es sonst eher aus dem angloamerikanischen Raum kennt: gut verständlich und leicht lesbar geschrieben, mit anschaulichen Fallbeispielen aus der Praxis, Zusammenfassungen und Aufbereitung von zentralen Informationen in abgesetzten Kästen. Die Übungsbeispiele regen zur Reflexion und zur Diskussion an, sie fördern und fordern die gedankliche Durchdringung des ‚Stoffes‘. Das Layout könnte noch verbessert werden. So ist die helltürkise Schrift, die für die Beispielsbeschreibungen verwendet wird, aus Sicht der Rezensentin nicht sehr gut lesbar; Schlüsselaussagen – wie z.B. die Definition von Evaluation (S. 24) sind im Schriftbild weniger stark hervorgehoben als z.B. die Übungsaufgaben, wodurch sie weniger gut sichtbar werden.

Das Buch eignet sich zum Selbststudium oder als Basisliteratur für Evaluationsseminare. Es kann auch Fachpersonal aus dem Bildungsbereich, die Evaluationen durchführen (müssen) als Begleitbuch dienen. Allerdings geht es inhalt-

lich und auch vom Umfang her über einen reinen Praxisleitfaden (wie von Wolfgang Beywl oder seinem Umfeld bereits mehrfach publiziert: vgl. z.B. Beywl/Schepp-Winter 2000; Keller-Ebert/Kißler/Schobert 2005; Beywl/Kehr/Mäder/Niestroj 2008) deutlich hinaus. Das Buch ist nicht nur ausführlicher, sondern auch anspruchsvoller, es geht auf dahinter liegende theoretische Konzepte ein und verbindet Theorie mit Praxis. Es empfiehlt sich also für Personen, die durchaus tiefer in die Materie eindringen wollen und nicht allein eine praxisorientierte Anleitung suchen (auch wenn das Buch als solche genutzt werden kann). Lehrende der Evaluation finden einen Fundus an gut aufbereiteten Schlüsselinformationen, Beispielen und Übungsaufgaben, die entweder direkt übernommen oder dem jeweiligen Kontext angepasst werden können.⁴

Das Buch spricht vor allem Personen aus der Weiterbildung/Erwachsenenbildung direkt an: Obwohl Balzer und Beywl sich vorgenommen haben, den gesamten Bildungsbereich abzudecken, ist das Buch doch vor allem auf diesen speziellen Bildungssektor ausgerichtet. Das merkt man deutlich an den Beispielen, aber auch an fehlenden Hinweisen auf die speziellen institutionellen Bedingungen in anderen Sektoren wie Schule oder Hochschule. Das widerspricht aber nicht unbedingt der Eignung des Buches für eine breitere Verwendung – außer man pocht auf eine sektorspezifische Einführung. Die Autoren selbst sehen auch „benachbarte Felder“, wie z.B. soziale Arbeit, Gesundheitsbereich, Konsumentenschutz, Beratung, Coaching, als in ihrem Buch angesprochen, da auch dort „gelernt“ wird. Dem ist zuzustimmen: die inhaltlichen Grundlagen passen prinzipiell auf jede Form von Bildungsprogramm (und mit ein wenig Flexibilität lassen sich auch die Beispiele in andere Bildungskontexte ‚übersetzen‘). Dazu gehören aus Sicht der Rezensentin auch alle Evaluationsgegenstände/-felder, in denen der „Programmarchetyp ‚Beratung, Information und Bildung‘“ (Funnell/Rogers 2011: 352ff.) umgesetzt wird – das Buch kann damit auch über den eigentlichen Bildungssektor hinaus Anwendung finden.

4. Fazit

Das Buch bietet eine theoretisch fundierte, aber gleichzeitig praxisnahe Einführung in die Evaluation für Studierende, ist aber gleichermaßen für

3 Nach wie vor gibt es ja einige ‚Evaluationslehrbücher‘, in denen es (fast) ausschließlich um (vornehmlich quantitative) Erhebungsmethoden geht.

4 Die Rezensentin hat selbst schon an Fallbeispielen und Übungen aus Vorversionen des Buches für Informationsangebote oder Schulungen zu Evaluation Anleihen genommen – mit gutem Erfolg bei den Zielgruppen (siehe z.B. <http://www.sqa.at/mod/page/view.php?id=525>, 01.07.2016).

interessierte Praktiker(innen) sehr gut geeignet und auch für Lehrende als Grundlagentext und/oder Inspiration für die eigene Lehrtätigkeit interessant. Es kann als erste Einführung empfohlen werden – oder auch, da es eigene, neue Akzente setzt, als ergänzender Text zu bereits vorhandenen Einführungswerken.

5. Literatur

- Balzer, Lars (2005): *Wie werden Evaluationsprojekte erfolgreich?* Landau: Verlag Empirische Pädagogik.
- Beywl, Wolfgang/Kehr, Jochen/Mäder, Susanne/Niestroj, Melanie (2008): *Evaluation Schritt für Schritt: Planung von Evaluationen.* Münster: hiba.
- Beywl, Wolfgang/Schepp-Winter, Ellen (2000): *Zielgeführte Evaluation von Programmen: Ein Leitfaden.* Berlin: Bundesministerium für Familie, Senioren, Frauen und Jugend, Referat 501.
- Funnell, Sue C./Rogers, Patricia (2011): *Purposeful Program Theory: Effective Use of Theories of Change and Logic Models.* San Francisco, CA: Jossey-Bass.
- Keller-Ebert, Cornelia/Kißler, Mechthilde/Schober, Berthold (2005): *Evaluation praktisch! Wirkungen überprüfen, Maßnahmen optimieren, Berichtsqualität verbessern.* Darmstadt: hiba.
- Kirkpatrick, Donald/Kirkpatrick, James D. (2006): *Evaluating Training Programs: The Four Levels.* San Francisco, CA: Berrett-Koehler (3. Aufl.).
- Weiss, Carol H. (1998): *Evaluation – Methods for Studying Programs and Policies.* Upper Saddle River, N.J: Prentice Hall (2. Aufl.).
- Glossar der Evaluation. Verfügbar unter: http://eval-wiki.org/glossar/Eval-Wiki:_Glossar_der_Evaluation [01.07.2016].