Rezension zu:

Stockmann, Reinhard/Meyer, Wolfgang (Hg.): The Future of Evaluation. Global Trends, New Challenges, Shared Perspectives. Hampshire: Palgrave Macmillan, 2016. 393 Seiten, 75,00 £, ISBN 978-1-137-37636-7

Hellmut Wollmann[1]

Der von Reinhard Stockmann und Wolfgang Meyer herausgegebene Band geht auf eine am Centrum für Evaluation an der Universität des Saarlandes am 14. und 15. Juni 2012 anlässlich dessen zehnjährigen Bestehens veranstaltete Tagung zurück. Die Neuerscheinung gibt einen überaus informativen, anregenden und lesenswerten Ein-, Über- und Ausblick auf die ‚globale' Entwicklung der Evaluationsforschung.

Zunächst sei ein sozusagen ‚wissenschaftsaußenpolitisches' Verdienst der neuen Publikation hervorgehoben. Dadurch, dass die Aufsätze des Sammelbandes in englischer Sprache geschrieben sind und dieser bei einem der international führenden angelsächsischen Fachverlage erscheint, leistet er, von zwei deutschen Wissenschaftlern initiiert und ediert, einen bemerkenswerten, sozusagen ‚deutschen' Beitrag zum internationalen Fachdiskurs, in dem ansonsten Publikationen angelsächsischer, insbesondere US-amerikanischer Provenienz bestimmend vorherrschen.

Auf rund 350 Seiten versammelt der Band 25 Aufsätze, die von 31 Autor(inn)en aus 18 Ländern aller Kontinente verfasst sind. In der thematischen und personellen Komposition des Bandes konnten sich die Herausgeber offenkundig auf die umfangreichen eigenen Forschungsarbeiten und -kontakte stützen, über die sie und das von ihnen geleitete Saarbrücker Centrum im Felde der nationalen und internationalen Evaluationsforschung – mit einem Schwerpunkt auf Entwicklungspolitik und deren Evaluierung – verfügen.

Ihre Vertrautheit mit Entwicklungsländern (im einschlägigen Jargon gesprochen: mit Ländern des ‚Global South') schlägt darin zu Buche, dass die Leserin/der Leser mit Entstehung, Stand und Perspektiven der Evaluationsforschung in Ländern vertraut gemacht wird (so China, Ägypten, Uganda, Costa Rica), die – jedenfalls im ‚Mainstream' der Evaluationsliteratur – kaum ins Blickfeld gerückt werden. Hierbei sei insbesondere der von Susan Tamondong geschriebene Übersichtsaufsatz zu den Ländern des ‚Global South' hervorgehoben (S. 51ff.). Allerdings hat dieser ‚Bias towards Development Cooperation', auf den die Herausgeber eingangs selber hinweisen (S. 18), den Preis, dass (zumal in einem Band, dessen Seitenzahl um seiner Lesbarkeit willen begrenzt ist) die Länder des ‚Global North', sieht man von dem Aufsatz von Donna Mertens (S. 214ff.) ab, weniger eindringlich vertreten sind.

In der ‚Rekrutierung' und Zusammensetzung der Autor(inn)en des Bandes ist den Herausgebern eine fruchtbare interdisziplinäre und interinstitutionelle Mischung von Fachleuten aus universitären und außeruniversitären Forschungseinrichtungen ebenso wie aus öffentlichen Verwaltungen und internationalen Förderinstitutionen gelungen.

In ihrem einleitenden Aufsatz formulieren die Herausgeber vorab als Leitthese des Bandes: „During the last decades evaluation has become an important instrument for rational governance… In general the history of evaluation is a remarkable and globally effective success story" (S. 9).

Zur Strukturierung des Bandes und der in ihm versammelten Beiträge geben sie vor allem fünf Suchfragen und ‚Indikatoren' vor (S. 12f.), nämlich a) die „Nachfrage" nach Evaluierung (etwa Anforderungen aus der Staats- und Verwaltungsmodernisierung, b) das „Angebot" von Evaluation (etwa Entstehung und Verbreitung

1 Humboldt-Universität zu Berlin

© Waxmann Verlag GmbH. Nur für den privaten Gebrauch.

einschlägiger universitärer und außeruniversitärer Forschungseinrichtungen), c) die „Institutionalisierung“ von Evaluation (etwa die Entstehung von Evaluationsgesellschaften), d) die „Professionalisierung“ von Evaluation (etwa Einrichtung einschlägiger universitärer Studien- und Ausbildungsgänge) und e) die Verbreitung und Nutzung von Evaluationsergebnissen.

Aus der imposanten Fülle, Vielfalt und Dichte des Bandes können und sollen in dieser kurzen Besprechung zwei Debattenstränge ausgewählt und angesprochen werden, die sich als ‚rote Fäden‘ durch mehrere Aufsätze ziehen.

Zum einen sei die Diskussion zu Entwicklung, Stand und künftigen Perspektiven der Evaluationsforschung aufgegriffen. Mehrere Beiträge vermitteln einleuchtend das Bild, dass die Evaluationsforschung in den USA und in Europa, also im ‚Global North‘, nach dem ‚Aufbruch‘ in den 1960er Jahren – in ihrer weiteren Expansion maßgeblich durch die vom New Public Management inspirierten Politik- und Verwaltungsmodernisierung und ihren Evaluierungsmaximen gewissermaßen ‚endogen‘ vorangetrieben worden ist, während sie in den Entwicklungsländern (des ‚Global South‘) im Wesentlichen durch internationale Förderinstitutionen und Organisationen (wie World Bank und OECD) und deren Nachfrage nach Evaluierung ihrer Fördermaßnahmen, also ‚exogen‘, bestimmt worden ist (S. 351). Hinsichtlich der künftigen Entwicklung der Evaluationsforschung werden für die Länder des ‚Global North‘ mehrere und unterschiedliche Einflussfaktoren diskutiert. Zum einen macht Frans L. Leeuw (S. 254ff.) auf neue Politik- und Evaluierungsfelder aufmerksam, die sich insbesondere aus der Digitalisierung der nationalen und internationalen Kommunikation und deren Regulierung durch nationale Regierungen und supranationale Regime (‚Digital Policies‘) ergeben. Zum anderen sieht Jan-Eric Furubo (S. 228ff.) voraus, dass insbesondere die europäischen Länder vor dramatischen Politikbrüchen stehen werden („turbulent times“, er deutet künftig noch massivere Flüchtlingsströme aus „several countries south of the Mediterranean“ an) und dass dadurch auch die (bislang eher auf ‚inkrementale‘ Politikveränderungen geeichte) Evaluationsforschung vor neue inhaltliche und methodische Herausforderungen gestellt werden wird. Des Weiteren folgert Donna Mertens (S. 214ff.) eine Neufokussierung der Evaluationsforschung daraus, dass sich diese (normativ) verstärkt den Fragen der „social justice und human rights“ zuwende und sich einem dieser Problemstellung methodisch und analytisch angemessenen „Process and Participatory“-Ansatz („involving communiy members in the evaluative process“) verschreiben solle.

Als weitere sich durch mehrere Beiträge ziehende Diskussionslinie sei die Verbreitung und Nutzung der Evaluationsergebnisse, also das ‚Lernen‘ in Politik und Verwaltung, aufgegriffen.

Übereinstimmend (und einleuchtend) wird davon ausgegangen, dass, wie Carol Weiss, die Ahnherrin der ‚Knowledge Utilization Research‘ schon in den 1980er Jahren beobachtete, der Fundus an verfügbarem Evaluierungswissen rasant anwächst und zuletzt vermöge des Zugriffs auf die elektronischen Medien (Internet, Social Media) regelrecht explodiert, wodurch sich die Frage nach ihrer Nutzung nur noch dringlicher und drängender stellt (Stockmann/Meyer, S. 16). In der Frage, wie die Chancen der Nutzung dieses enormen politischen, administrativen und gesellschaftlichen Lernpotenzials zu sichern und zu steigern seien, werden im Wesentlichen zwei Positionen erörtert.

Einerseits wird darauf abgehoben, dass die Nutzungschance durch die ‚Akzeptanz‘ der Ergebnisse durch von der Evaluation ‚Betroffene‘, und diese wiederum durch deren frühzeitige Beteiligung und Einbindung in den Evaluationsprozess, gesteigert werden können. Dies ist der Kern von Michael Pattons vielzitiertem Konzept einer ‚User Focused Evaluation‘. An dieses knüpft Peter Dahler-Larsen (S. 33ff.) und besonders nachdrücklich Donna Mertens (eine Doyenne der US-amerikanischen Evaluationsforschung) (S. 214ff.) mit ihrem (bereits erwähnten) Plädoyer für einen ‚transformativen‘ Ansatz, der „social justice and human rights“ und die Betroffenenbeteiligung als normative Richtschnur für die Zielsetzung und methodische Anleitung für Evaluation an.

Dem steht die Vorstellung gegenüber, dass eine Verbesserung der Nutzungschance von Evaluierungsergebnissen in erster Linie durch die Sicherung ihrer methodisch-wissenschaftlichen Validität und durch ihre in dieser Validität gründenden Glaubwürdigkeit („credibility“) erzielt werden könne. Um diese zu erreichen und zu erhalten, sei die institutionelle und kognitive Unabhängigkeit der Evaluationsforschung und Evaluationsforscher(innen) unerlässlich, die am ehesten durch deren akademisch-universitäre Einbindung bewirkt werden könne, und umgekehrt durch ihre Nähe oder gar Einbindung in die Binnenstruktur der staatlichen usw. Auftraggeber von Evaluation strukturell infrage gestellt sei.

© Waxmann Verlag GmbH. Nur für den privaten Gebrauch.

Unter dem Stichwort eines „gap between science and utility“ wird dieses Problem und Dilemma von den Herausgebern ausführlich und eindringlich diskutiert (S. 238ff., S. 345). Sie neigen eher der zweiten Position zu, indem sie vor dem Risiko einer ‚Entwissenschaftlichung‘ („de-scientification“, S. 356) der Evaluationsforschung warnen und sich für ihre sozialwissenschaftliche, nach Möglichkeit universitäre oder hochschulnahe Verankerung verwenden (womit sie vermutlich nicht zuletzt das von ihnen gegründete und geleitete Saarbrücker Centrum modellhaft vor Augen haben). Demgegenüber sei auch insoweit noch einmal auf Donna Mertens verwiesen, die ihr Plädoyer für eine die Beteiligten und gesellschaftliche Praxis einbeziehende Evaluationsstrategie mit der Vorstellung und Forderung verbindet, dass durch einen der jeweiligen Ausgangssituation und Konstellation der Evaluierung jeweils angemessenen und gebotenen ‚Methodenmix‘ („use of mixed methods“, S. 218) nicht nur die der Akzeptanz förderliche Praxis- und Betroffenennähe, sondern zugleich auch die wissenschaftliche Validität eingelöst und damit das Spannungsverhältnis und ‚Dilemma‘ zwischen Wissenschaftlichkeit und praktischem Nutzen, zwischen ‚Science and Utility‘, überbrückt und aufgehoben werden können.

In einem ausgezeichneten (der/dem zeitknappen und raschen Leser(in) als Eingangs- und Vorablektüre empfohlenen) Schlusskapitel fassen die Herausgeber die Ergebnisse der 25 Kapitel ihres Bandes entlang der von ihnen vorgeschlagenen fünf Leitfragen zusammen, indem sie die Entwicklung der Evaluationsforschung in den USA und den europäischen Ländern (des ‚Global North‘) und in den Entwicklungsländern (‚Global South‘) – in ihren Unterschieden und Gemeinsamkeiten – gegenüberstellen und ihre differenzierten Aussagen in einer übersichtlichen und informativen Tabelle (S. 351 ff.) konzis kondensieren. Alles in allem sehen sie ihre eingangs formulierte These („a lot of evidence can be delivered for calling evaluation an ongoing global success story“ (S. 10)) durch die Beiträge des Bandes bestätigt.

Der inhaltsreiche und anregende Band verdient Beachtung und Verbreitung nicht nur in der deutschen, sondern darüber hinaus auch gerade in der angelsächsischen und anglophonen ‚Evaluation Community‘.

UNSERE BUCHEMPFEHLUNG

Bernd Groot-Wilken, Kevin Isaac, Jörg-Peter Schräpler (Hrsg.)

Sozialindices für Schulen

Hintergründe, Methoden und Anwendung

Der Band gibt einen umfassenden Überblick der Diskussion über Schulsozialindices. Die Autoren beschreiben verschiedene Möglichkeiten der Berechnung und Anwendung und geben einen Einblick in den aktuellen Stand der Dinge im deutschsprachigen Raum. Die Herausgeber möchten damit die Diskussion um faire(re) Vergleiche sowohl auf der methodischen als auch administrativen Umsetzungsebene weiterführen und Impulse für Weiterentwicklungen geben.

Beiträge zur Schulentwicklung, 2016, 216 Seiten, br., 29,90 €, ISBN 978-3-8309-3486-8

E-Book: 26,99 €, ISBN 978-3-8309-8486-3

www.waxmann.com

© Waxmann Verlag GmbH. Nur für den privaten Gebrauch.

Rezension zu:

Großmann, Daniel/Wolbring, Tobias (Hg.): Evaluation von Studium und Lehre. Grundlagen, methodische Herausforderungen und Lösungsansätze. Wiesbaden: Springer VS, 2016. 429 Seiten, 34,99 EUR, ISBN 978-3-658-10885-4

Antonia Baumeister[1]

Der von Daniel Großmann und Tobias Wolbring herausgegebene Sammelband „Evaluation von Studium und Lehre" stellt Grundlagenwissen und neue Forschungsergebnisse zu methodischen und inhaltlichen Fragestellungen der Evaluation an Hochschulen dar. Die 21 Autor(inn)en des Bandes stammen aus verschiedenen Fächern und verdeutlichen bereits die Interdisziplinarität der Thematik. Der Band wird mit einem Geleitwort des Doyens der deutschen Hochschulevaluationsforschung, Hans-Dieter Daniel von der ETH Zürich, eingeleitet. Er skizziert die Entwicklung der Evaluation von Studium und Lehre in Deutschland von der Novelle des Hochschulrahmengesetzes im Jahr 1998 bis heute. Betont wird der hohe wissenschaftliche Anspruch von Evaluationen, der sich auch in den Beiträgen des Bandes widerspiegelt. Der Band gliedert sich in „Grundlagen", „Input", „Prozess" und „Output".

Grundlagen

Begonnen wird mit einem Überblick zu „Stand und Herausforderungen der Evaluation an deutschen Hochschulen" (Daniel Großmann und Tobias Wolbring).

Der zweite Beitrag „Entstehung und Grundlagen der Evaluation" von Reinhard Stockmann äußert interessantes interkulturelles Expertenwissen zur historischen Entwicklung der Evaluation in den USA versus Deutschland. Ausgehend von einem Forschungsmodell der Evaluation lässt Reinhard Stockmann auch kritische Töne an der deutschen Praxis der unabhängigen Evaluation eines Politikfeldes anklingen (vgl. S. 33-34).

Der dritte Beitrag „Evaluation, Kausalität und Validität" von Tobias Wolbring ist methodisch grundlegend und unterstützt das Verständnis über veranschaulichende Beispiele. Manche fachspezifischen methodischen Begriffe, insbesondere zu statistischen Verfahren, z.B. ‚Sensitivitätsanalyse' oder ‚Stratifikation', hätte man noch näher erläutern können. Vielleicht könnte eine zweite Auflage von einem Glossar am Buchende profitieren. Intensiv setzt sich Wolbring mit der Bologna-Reform und ihrer Evaluation auseinander. Kritisch summiert er hier, dass noch keine „belastbare evidenzbasierte Wirkungsevaluation der Bologna-Reform" (S. 83) erfolgt sei. Wenngleich die Forderung nach einer solchen Wirkungsevaluation aus wissenschaftlicher Sicht richtig ist, erscheint es praktisch kaum umsetzbar, internationale Reformen in einer für diesen Rahmen angemessenen Feldstudie zu untersuchen. Andere Bemühungen, Bildungssysteme vergleichend zu evaluieren (z.B. PISA-Studien), kranken an einer Reihe von methodischen Problemen (vgl. Rindermann/Baumeister 2015). Ein Verbund von Hochschulen zu einer Arbeitsgruppe könnte womöglich einen Ex-post-facto-Vergleich von Diplom- und Masterstudiengängen realisieren.

Input

Im ersten Beitrag des Input-Abschnitts „Soziale Determinanten der Studienaufnahme und Fachwahl" von Werner Georg und Tino Bargel werden, basierend auf dem Deutschen Studierendensurvey, die teilweise sehr deutlichen Zusammenhänge zwischen der beruflichen Fachrichtung der Väter und der Fachrichtung der Studierenden aufgezeigt. Aufschlussreich sind auch die Ergebnisse zum Einfluss des Geschlechts auf die Fachwahl sowie die Ergebnisse zu fachkulturellen Distanzen. Folgerichtig spre-

1 Technische Universität Chemnitz

© Waxmann Verlag GmbH. Nur für den privaten Gebrauch.

chen sich die Autoren bei Evaluationen für eine stärkere Berücksichtigung sozialer Merkmale der Studierenden aus.

Daniel Großmann berichtet in seinem Beitrag („Studienmotivation und Evaluation") von einer Studie zu den Zusammenhängen zwischen Studienmotivation und Evaluation in sozialwissenschaftlichen Masterstudiengängen an der Universität Leipzig. Das Konstrukt der Studienmotivation wird ausführlich definiert. Dabei wird ein Dreikomponentenmodell zur Systematisierung herangezogen. Auf dieser relativ abstrakten Ebene bleibt es schwierig, das Zusammenspiel der drei Komponenten nachzuvollziehen und ihre Bedeutung und Funktion zu durchdringen; eine veranschaulichende Abbildung wäre hilfreich gewesen. Dennoch ist der Beitrag für alle interessant, die sich genauer über das Konstrukt der Studienmotivation, seine Messung sowie seinen Einfluss auf Evaluationen von Studium und Lehre informieren möchten. Der Beitrag ist mit 61 Seiten deutlich länger als die anderen Beiträge des Bandes. Insofern entstand beim Lesen der Eindruck eines ‚Buches im Buch'.

Methodisch ebenfalls exzellent ist die Studie „Messung von studentischem Workload", von der Roger Berger und Bastian Baumeister berichten. In der Studie wurden im Wesentlichen zwei Erfassungsarten sowie zwei längsschnittliche Zeitperioden verglichen: Papiertagebuch versus Smartphone-App-Tagebuch, prospektiv versus retrospektiv, wobei die Retrospektiverhebung zur Erfassung soziodemographischer Variablen um einen Online-Fragebogen ergänzt war. Die Studie ist sehr gut nachvollziehbar und anschaulich beschrieben. Sie liefert aussagekräftige Befunde sowie wertvolle Anregungen zur Messung des studentischen Workloads.

Prozess

Der erste Beitrag des Prozessabschnitts „Lehrveranstaltungsevaluation an Hochschulen" stammt von Heiner Rindermann und widmet sich dem Einfluss von Rahmenbedingungen auf Qualität von Lehre und Ergebnisse von Lehrevaluation. Die zentralen Konzepte der Lehrveranstaltungsevaluation werden erläutert. Studienergebnisse verdeutlichen die Grenzen von Evaluation zur Verbesserung der Lehrqualität. In den Unterkapiteln 4 und 5 stellt Rindermann auf eindrückliche und gleichsam amüsante Art und Weise dar, welche Bedeutung Rahmenbedingungen für die Qualität der Hochschulbildung haben. Ein unbequemer Beitrag, der die Missstände an deutschen Universitäten, auch im internationalen Vergleich, deutlich macht.

Im zweiten Prozessbeitrag „Konstruktvaliditätsprobleme von Lehrevaluationen" zeigen Brigitte Metje und Udo Kelle auf, wie entscheidend *Think-Aloud*- oder *Probing*-Verfahren − als zweite Säule neben quantitativen Methoden der Testkonstruktion − sind, um die Konstruktvalidität von Lehrevaluationsinstrumenten zu prüfen. Auch dieser Beitrag ergänzt die vorherigen um grundlegende Definitionen. Ein empirisches Forschungsprojekt an der Universität Marburg wird vorgestellt, in dessen Rahmen Beantwortungsfehler und Verständnisdivergenzen bei Lehrveranstaltungsevaluationsinstrumenten mithilfe von kognitiven Interviews mit den Beantwortenden aufgedeckt wurden. Die dadurch deutlich eingeschränkte Konstruktvalidität der untersuchten Items wäre mit rein quantitativen Methoden nicht erkannt worden, denn diese Analysen fielen unauffällig aus. Dieser Beitrag ist also insbesondere für die Entwickler(innen) und Anwender(innen) von Lehrevaluationsinstrumenten relevant.

Der dritte Beitrag „Studierendenbefragungen als Panelstudie" stammt von Philipp Pohlenz, Frank Niedermeier, Melinda Erdmann und Juliana Schneider. Die Autor(inn)en erläutern, wie Panelstudien als Längsschnittdaten genutzt werden können, um die Qualitätssicherung an Hochschulen auf eine kausal-empirisch belastbare Basis zu stützen. Studiengänge zielen darauf ab, fachliche und überfachliche Kompetenzen zu vermitteln. Panelstudien tragen u.a. dazu bei, sich an eine kausale Effektivitätsprüfung von Studiengängen anzunähern. Die Potsdamer Befragungen werden jährlich durchgeführt bis zur Exmatrikulation. Danach folgen mehrere Absolventen- und Exmatrikuliertenbefragungen. Die Befragungen liefern umfangreiche Informationen u.a. zu selbsteingeschätzten Kompetenzen, zur Studienmotivation, Studienzufriedenheit, zu wahrgenommenen Schwierigkeiten sowie zu Abbruch- bzw. Wechselgründen. Interessant ist beispielsweise, welche Gemeinsamkeiten und Unterschiede in der Veränderung der Aspekte der Studienmotivation zwischen den untersuchten Studiengängen beobachtet werden können. Aus den Verlaufsmessungen zur Studienmotivation werden sinnvolle Schlussfolgerungen zur Weiterentwicklung der Studiengänge abgeleitet. Die Autor(inn)en weisen darauf hin, dass die Aussagekraft dieser Panelstudie bislang dadurch eingeschränkt war, dass lediglich zwei Messzeitpunkte vorlagen, die jedoch künftig vermehrt werden sollen. Daniel Großmann berichtet in seinem Input-Beitrag von einer Studie zur Erfassung der Studienmotivation, wenngleich es sich im Gegensatz zu der Panelstudie

© Waxmann Verlag GmbH. Nur für den privaten Gebrauch.

von Pohlenz und Kolleg(inn)en bei Großmann um eine Querschnittsstudie handelte. Interessant wäre die Frage, ob beide Studien ähnliche Befunde lieferten. Beispielsweise kommen die Studien insofern zu unterschiedlichen Ergebnissen, als Großmann von einer, im Vergleich zu den Fachstudiengängen stärkeren, Unzufriedenheit von Lehramtsstudierenden berichtet, während die Potsdamer Panelstudie signifikante Anstiege bestimmter Motivationsaspekte bei Lehramtsstudierenden nachwies. Die Verlaufsabbildungen zu diesen Panelbefunden auf Seite 312 lassen vermuten, dass bei der Gruppe der Lehramtsstudierenden ein vergleichsweise geringer Anstieg der Studienmotivation vorliegt. Dies käme der Interpretation der Befunde von Großmann wiederum nahe (vgl. S. 161: „Sonderfall Lehramtsstudium"). Weniger in methodischer Hinsicht, sondern im Hinblick auf die Interessen von Leser(inne)n, sei kritisch angemerkt, dass die Leipziger Querschnittsstudie und die Potsdamer Panelstudie sowie ihre Befunde zur Studienmotivation nicht auf eine Vergleichbarkeit hin ausgelegt waren und es daher möglicherweise interessanter gewesen wäre, wenn die Befunde der Panelstudie zu den selbsteingeschätzten Kompetenzen der Studierenden anstatt zu ihrer Studienmotivation berichtet worden wären.

Output

Der erste Beitrag des Output-Abschnitts „Studienabbruchquoten als Evaluationskriterium" stammt von Daniel Klein und Volker Stocké. Dieser Beitrag dürfte für zentrale hochschulpolitische Akteure, wie die Kultusministerkonferenz (KM), den Wissenschaftsrat (WR) sowie die Hochschulrektorenkonferenz (HRK) besonders informativ sein. Die Autoren kritisieren, dass Studienabbruchquoten häufig als Evaluationskriterium und Steuerungsinstrument der Qualitätssicherung genutzt werden, obwohl die Vergleichbarkeit und Validität der Messungen fragwürdig ist. Zudem weisen sie auf mögliche Risiken dieser Praxis im Hochschulbereich hin. Mit dem Monitoring des deutschen Hochschulsystems befassen sich z.B. die Autorengruppe Bildungsberichterstattung, das Statistische Bundesamt sowie das Deutsche Zentrum für Hochschul- und Wissenschaftsforschung in Zusammenarbeit mit der Arbeitsgruppe Hochschulforschung der Universität Konstanz. Diese Institutionen kommen, aufgrund ihrer Berechnungsvarianten, zu deutlich unterschiedlichen Kennzahlen für die Studienabbruchquote. Beispielsweise unterscheidet sich die Berücksichtigung von Bildungsinländer(inne)n und Bildungsausländer(inne)n. Des Weiteren diskutieren die Autoren die empirisch bislang mehr oder weniger einheitlich vorgefundenen Ursachen für Studienabbrüche, von denen manche außerhalb des Einflussbereichs der evaluierten Akteure liegen, z.B. soziodemographische Merkmale. Während die Befundlage zur Lehrqualität als Ursachenfaktor für Studienabbrüche uneinheitlich ist, hat sich die Note der Hochschulreifeprüfung als wichtiger Ursachenfaktor für den Studienerfolg erwiesen. Einen Ansatz zur akademischen und sozialen Integration von Studierenden als Präventionsmaßnahme gegen Studienabbrüche halten die Autoren für schwer umsetzbar und wenig Erfolg versprechend.

Felix Wolter und Jürgen Schiener stellen in ihrem Beitrag zum Thema Kompetenzdiagnostik an der Hochschule „Methoden, Herausforderungen und aktuelle Entwicklungen" (S. 367 ff.) vor. Dieser Beitrag ist für Evaluationsbeauftragte sowie für Forscher(innen) interessant. Es geht um die Frage der Vergleichbarkeit von Noten im Hochschulbereich angesichts u.a. dozentenspezifischer Gepflogenheiten. Die zentralen Konzepte und Prinzipien der Kompetenzdiagnostik, wie sie beispielsweise auch im Rahmen von Schulleistungsstudien angewandt werden, werden gut verständlich erläutert und u.a. bildlich dargestellt, beispielsweise Kompetenzstrukturmodell und Kompetenzniveaumodell, adaptives Testen, IRT-Verfahren usw. Die Bedeutsamkeit des „Differential Item Functioning" (DIF; vgl. S. 377)[2]und seine Implikationen für den Hochschulbereich (vgl. z.B. S. 381) werden erklärt. Es werden Beispiele für deutsche und internationale Projekte zur Kompetenzmessung im Hochschulbereich genannt. Speziell für Deutschland wären dies eine Teilstudie des Nationalen Bildungspanels (NEPS) sowie der vom Bundesministerium für Bildung und Forschung (BMBF) geförderte Forschungsverbund „Kompetenzmodellierung und Kompetenzerfassung im Hochschulsektor". Methodische Anmerkungen sowie einige Ergebnisse des Pilotprojekts werden in aller Kürze vorgestellt und sind sehr interessant. Die Autoren nennen konkrete Anwendungsmöglichkeiten u.a. für die Studiengangsentwicklung.

Der abschließende Beitrag des Output-Abschnitts sowie des Sammelbandes, „Berufseinstiege von AkademikerInnen mit Migrationshintergrund in Österreich", stammt von Roland Verwiebe und Melek Hacioglu. Sie stellen ei-

2 Ein DIF liegt vor, wenn Test-Items nicht aufgrund der Personenfähigkeit, sondern aufgrund einer Drittvariablen Teilnehmergruppen diskriminieren.

© Waxmann Verlag GmbH. Nur für den privaten Gebrauch.

nen Mixed-Methods-Ansatz vor, der über herkömmliche Absolventenbefragungen hinausgeht und dazu dient, Barrieren und Diskriminierungserfahrungen von Absolvent(inn)en zu identifizieren. An der Universität Wien wurden zeitgleich quantitative Fragebogendaten und qualitative Interviewdaten in einem ‚Convergent Parallel Design' erhoben. Die quantitativ untersuchten Faktoren umfassten beispielsweise die Dauer der Arbeitssuche beim Berufseinstieg, die Anzahl an Bewerbungen sowie die berufliche Stellung der ersten Tätigkeit. Interessanterweise wurden keine Lohnunterschiede zwischen Akademiker(inne)n mit versus ohne Migrationshintergrund bei ihrer ersten beruflichen Tätigkeit festgestellt. Die Probleme, die im Rahmen der Interviews genannt wurden, betrafen beispielsweise den ausländischen Namen oder Stereotype bezüglich des optimalen Aussehens, um für Kundengespräche einsetzbar zu sein. Die Autorin und der Autor fanden Hinweise darauf, dass sich die Benachteiligungen insbesondere bei Akademikerinnen mit Migrationshintergrund kumulieren. Es wäre wichtig, eine ähnliche Studie in Deutschland durchzuführen (als Beispiel für eine frühere Studie: vgl. Farrokhzad 2008). Auch bereits zu früheren Zeitpunkten, beispielsweise im Hinblick auf die Betreuung auf verschiedenen Qualifikationsstufen, sollten Studium und Lehre auf Barrieren und Diskriminierungsgefahren hin evaluiert und geeignete Gegenmaßnahmen entwickelt werden.

Fazit

Abschließend lässt sich feststellen, dass es vielfältige und auch gut ausgereifte Lösungsansätze für die Evaluation von Studium und Lehre gibt. Es existieren beispielsweise differenzierte Instrumente zur Erfassung der Studienmotivation, des Workloads sowie der Qualität von Lehrveranstaltungen. Da die genannten Verfahren bislang überwiegend auf subjektiven Einschätzungen der Studierenden beruhen, liegt die Zukunft der Evaluationsforschung im Hochschulbereich in der Entwicklung objektiverer Maße, beispielsweise im Bereich der Kompetenzmessung. Der Sammelband informiert auch über zahlreiche Rahmenbedingungen der Studienaufnahme, Fachwahl sowie des Studienerfolgs bzw. des Studienabbruchs. Damit die vorgestellten Lösungsansätze nicht bloß isolierte Puzzlestücke bleiben, besteht eine Herausforderung für Akteure im Hochschulbereich darin, diese Ansätze in einer praktikablen Art und Weise zu verknüpfen. Für die Forschung wäre es beispielsweise interessant zu untersuchen, wie sich der Workload auf die Studienmotivation (bzw. auch umgekehrt), auf die Kompetenzen der Studierenden und auf die Studienabbruchquote auswirkt.

Der Sammelband erfüllt sein Ziel, Probleme und mögliche Lösungsansätze aufzuzeigen. Aus der Sicht der Rezensentin werden alle genannten Zielgruppen, d.h. Forschende und Studierende in den Bildungs- und Sozialwissenschaften, Evaluationsbeauftragte, Tätige in der Hochschulforschung sowie Hochschulleitungen und -management, durch die Beiträge angesprochen.

Literatur

Farrokhzad, Schahrzad (2008): Erfahrungen, Strategien und Potenziale von Akademikerinnen mit Migrationshintergrund. In Hentges, Gudrun/Hinnenkamp, Volker/Zwengel, Almut (Hg.): Migrations- und Integrationsforschung in der Diskussion. Biografie, Sprache und Bildung als zentrale Bezugspunkte. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 303-322.

Rindermann, Heiner/Baumeister, Antonia E. (2015): Validating the Interpretations of PISA and TIMSS Tasks: A Rating Study. In: International Journal of Testing, 15 (1), S. 1-22. doi:10.1080/15305058.2014.966911

© Waxmann Verlag GmbH. Nur für den privaten Gebrauch.

Rezension zu:

Griese, Christiane/Marburger, Helga/Müller, Thomas (Hg.): Bildungs- und Bildungsorganisationsevaluation. Ein Lehrbuch. Berlin/Boston: De Gruyter, 2016. 412 Seiten, 34,95 EUR, ISBN 978-3-11-040039-7

Franziska Heinze,[1] Frank König,[1] Stefanie Reiter[1]

Auftrag und Aufbau des Lehrbuchs

Das von Christiane Griese, Helga Marburger und Thomas Müller vorgelegte Lehrbuch für „Evaluation von Bildung und Bildungsorganisationen" möchte die von den Herausgebenden konstatierte Lücke im Feld der Veröffentlichungen über Evaluationen im Bildungsbereich schließen, die zwischen Werken besteht, die ausschließlich spezifische Handlungsfelder im Bereich von Bildung betrachten, und Publikationen, die sich vorrangig auf konkrete Evaluationsprojekte konzentrieren. Griese, Marburger und Müller formulieren als Anspruch, „Evaluation in Bildungsorganisationen in einer umfassenden und zugleich kompakten Weise darzustellen, die sich auf die Bandbreite bestehender Bildungsorganisationen wie auf das Spektrum verschiedener Inhaltsdimensionen von Bildung bezieht" (Vorwort). Das Lehrbuch richtet sich an Studierende und Lehrende an Universitäten und (Fach-)Hochschulen, an Lehrende und Lernende in Fort- und Weiterbildungen sowie an mit Evaluation befasste Mitarbeitende aus Organisationen im Bildungsbereich und soll zu deren „Employability" durch die Vermittlung von Forschungsexpertise in unterschiedlichen Aufgabenfeldern beitragen.

Das Lehrbuch ist in vier Teile gegliedert. Die beiden ersten bieten einen Überblick zu Entstehungsbedingungen, Funktionen und Bedeutungen von Evaluation im Bildungsbereich und stellen ausgewählte Formate, Methoden und Instrumente von Evaluation vor. Die beiden hinteren Teile skizzieren das Feld der Evaluation im Hinblick auf ausgewählte Bildungsorganisationen und inhaltliche Dimensionen von Bildung. Alle Parts sind in mehrere, eigenständige Einzelbeiträge untergliedert, an deren Ende sich jeweils Vertiefungsaufgaben und -fragen befinden. Die Autorinnen und Autoren stammen laut Vorwort aus dem Bereich der Hochschulen, des Bildungs- und Kulturmanagements sowie der Bildungsforschung.

Teil I: Kontexte, Theorien, Debatten

Helga Marburger eröffnet das Lehrbuch mit einem Überblick zu Entwicklungslinien und Funktionsbestimmungen von Evaluation in unterschiedlichen Bildungskontexten. Sie stellt in historischer Perspektive gesellschaftliche und bildungspolitische Entstehungskontexte, erste theoretische Konzeptualisierungen und Ansätze in den USA sowie darauf folgende Entwicklungslinien in Deutschland dar, welche die aktuellen Ausprägungen von und Debatten über Evaluation im deutschsprachigen Bereich wesentlich beeinflusst haben. Skizziert wird der Bedeutungszuwachs von Evaluation in modernen, demokratischen Gesellschaften, damit korrespondierende Funktionserweiterungen und zentrale Formate sowie entsprechende Objekte im Bildungsbereich. Der historische Rückblick wird für einen Blick auf aktuelle Diskussionen zur institutionellen und wissenschaftlichen Verortung von Evaluation herangezogen, wobei der Entwicklung von Standards besondere Beachtung beigemessen wird. Marburger prognostiziert eine Tendenz zu flexiblen und gegenstandsangemessenen Konzepten, die zwischen transdisziplinären und sektoralen Verortungen dynamisch changieren. Dieser Trend zeichnet sich auch in entsprechenden Studiengängen ab (Evaluation als eigenständiges Studienfach und als Modul in Einzeldisziplinen).

Christiane Griese konturiert in ihrem Beitrag das dem Lehrbuch zugrundeliegende Ver-

1 Deutsches Jugendinstitut e.V., Halle

© Waxmann Verlag GmbH. Nur für den privaten Gebrauch.

ständnis von Bildungsorganisationen, um den Lesenden die Spezifika dieses Bereichs verständlich zu machen. Dies umfasst auch Evaluation als Steuerungsaufgabe und als immanenter Bestandteil von Organisationsentwicklungsprozessen und -konzepten. Aufbauend auf einer Darstellung von Merkmalen des zu betrachtenden Objektes Bildung, welche die Evaluierbarkeit wesentlich beeinträchtigen (u.a. komplexe Prozessqualitätsbedingungen, Ergebnisoffenheit pädagogischer Prozesse), wird das Konzept der Bildungsqualität eingeführt. Dieses ermöglicht eine Operationalisierung des Gegenstandsbereichs und damit den Zugriff von Evaluation. Die Autorin schließt mit einem Ausblick, in welchem sie (zukünftige) Mitarbeitende in Bildungsorganisationen auffordert, die Durchführung evaluativer Maßnahmen in diesen Organisationen selbst zu professionalisieren.

Thomas Müller geht auf die Bedeutung von Evaluation im Spannungsfeld divergierender Interessen im Bildungsbereich ein. Er erörtert unterschiedliche Akteurskonstellationen und deren Interessen an Evaluationen für die Bereiche Bildungspolitik, -verwaltung, -forschung sowie pädagogische Praxis. Dabei werden für die jeweiligen Instanzen zentrale historische Entwicklungslinien sowie aktuelle Trends (z.B. evidenzbasierte Steuerung) nachgezeichnet und mit Beispielen versehen. Der Beitrag schließt mit einer Zusammenfassung von Herausforderungen, die durch das dynamische Beziehungsgefüge der dargestellten Instanzen entstehen, und mit daraus resultierenden Gelingensbedingungen von Bildungsevaluation. Letztere sieht der Autor vor allem in der produktiven Auseinandersetzung mit Konflikten bei der Planung, Durchführung und Auswertung von Evaluationen.

Teil II: Formen, Methoden und Instrumente

Der zweite Teil beginnt mit einem Überblick von Jörg Nicht über sozialwissenschaftliche Methoden, die gegenstandsabhängig bei Evaluationen im Bildungsbereich eingesetzt werden. Der Autor geht auf quantitative und qualitative Methoden der Datengewinnung und -analyse sowie auf Möglichkeiten der Triangulation ein und diskutiert in Grundzügen die damit verbundenen spezifischen Schwierigkeiten vor dem Hintergrund von Evaluationstrends (Abkehr vom quasi-experimentellen Design als Golden Standard in den 1960er Jahren hin zum Einsatz vielfältiger Methoden). Dabei berücksichtigt er aktuelle methodologische Positionen in der Diskussion über Standards der Evaluation. Er plädiert für eine Transparenz und Begründung jeweils gegenstandsangemessener Methoden.

Karsten Speck führt mit seinem Beitrag zentrale Evaluationsbegrifflichkeiten im Bildungsbereich ein. Er geht auf Diskursstränge zur Fremd- und Selbstevaluation ein und erläutert die Verwendung der Begriffe ‚Programm-, Prozess- und Produktevaluation'. Darauf aufbauend liefert er einen Überblick über Ansätze und Modelle der Fremd- und Selbstevaluation und stellt das CIPP-Evaluationsmodell (Daniel L. Stuffelbeam) als ein für den Bildungsbereich oftmals angemessenes Modell vor. Anwendungsbezogene Hinweise auf etablierte Standards für Fremdevaluationen, eine Aufzählung von möglichen Arbeitsschritten einer Selbstevaluation und Checklisten ergänzen den Überblick. Ausgewählte, kompakt skizzierte Beispiele für Programmevaluationen und eine Selbstevaluation geben einen groben Einblick in Anwendungsmöglichkeiten in unterschiedlichen Bildungsfeldern. Der Beitrag endet mit der Darstellung zentraler Trends (Anspruch der Nutzung von Fremdevaluationen zur Verbesserung von Programmen, zunehmende Anforderungen an pädagogische Fachkräfte bzgl. Selbstevaluation, Bedeutungszuwachs IT-gestützter Formen von Evaluation). Letztendlich spricht sich der Autor dafür aus, derzeit bestehende Forschungslücken hinsichtlich des Nutzens und der Verwendung von Evaluationen im Bildungsbereich durch „komplexe, theoriegeleitete Wirkungsmodelle von Evaluationen sowie eine quantitative und qualitativ ausgerichtete Evaluationsforschung" (S. 101) zu schließen.

Der Text von Susanne Giel führt in die Potenziale von Programmevaluationen ein, in denen theoriebasierte Konzepte genutzt werden. Die Autorin erläutert den Begriff ‚Programmtheorie' – als zentralen Bestandteil dieser Konzepte – und mögliche Einsatzgebiete. Diese sieht sie in der Erfassung von Wirkungen sowie in der Verbesserung von Programmen und Maßnahmen. Sie erörtert zentrale Begriffsverständnisse und Anmerkungen zum Design und stellt allgemeine Schritte der Rekonstruktion und Überprüfung von Programmtheorien dar. Anschließend illustriert sie an zwei Beispielen, welche Konsequenzen ein theoriebasiertes Vorgehen bei der Durchführung von Evaluationen hat. Zusammenfassend werden dessen Voraussetzungen (u.a. zusätzliche Ressourcen) und Potenziale (z.B. Erhöhung von Transparenz für Programmbeteiligte) dargestellt.

Den letzten Beitrag des zweiten Teils widmet Bernd Benikowski der Planung und Durchführung von Evaluation zur systematischen Bewertung von Bildungsmaßnahmen und -institutionen. Dabei unterscheidet er zwei grund-

© Waxmann Verlag GmbH. Nur für den privaten Gebrauch.

legende Phasen: Die erste dient der Moderation eines Klärungs- und Planungsprozesses, in welchem vor allem divergierende Ziele und Interessen an der Evaluation herausgearbeitet und die Rahmenbedingungen und Ziele konkretisiert werden müssen. Darauf folgt die Phase des wissenschaftlichen Arbeitens im engeren Sinn, welche mit der Entwicklung von Bewertungskriterien beginnt und sich bis zur Ergebnispräsentation erstreckt. Im Beitrag werden beide Phasen am Beispiel eines Lernprogramms dargestellt und diskutiert. Der Artikel schließt mit Aspekten, die bei der Planung und Durchführung von Evaluationen berücksichtigt werden sollten und spricht sich für einen systematischen Einsatz bei Bildungsmaßnahmen und -programmen aus.

Teil III: Evaluation in ausgewählten Bildungsorganisationen

Der dritte Teil beginnt mit einem Beitrag von Ergin Focali zu frühkindlichen Bildungsorganisationen. Der Autor ordnet Evaluation in Kindertagesstätten in gesellschafts-, fach- und verwaltungspolitische Debatten ein und thematisiert sie als Instrument der Qualitätsmessung. Ausgehend von einer gesetzlich vorgeschriebenen Evaluationspflicht diskutiert Focali die Schwierigkeiten der Operationalisierbarkeit von pädagogischer Qualität und beleuchtet deren Dimensionen in komplexen Dienstleistungssettings. Er benennt etablierte Evaluationsverfahren zur Qualitätsmessung und -sicherung. Anhand der Implementierung eines Bildungsprogramms im Kitabereich erläutert er den Ablauf einer Evaluierungsmaßnahme. Abschließend hinterfragt er ökonomische Aspekte von und in Evaluationen (u.a. Ressourcen) vor dem Hintergrund der generellen Bedeutung von frühkindlicher Bildung sowie Entwicklung und fordert eine Orientierung am Maßstab des Kindeswohls.

Christiane Griese thematisiert die Bedeutung und Verfahren von Evaluation in der Schule. Veränderte Steuerungslogiken (vor allem Output-Orientierung, Bildungsstandards) sowie schon länger währende Diskussionen um Qualitätssicherung und Schulautonomie werden als Entwicklungskontexte skizziert. Neben strukturellen Rahmenbedingungen benennt Griese die empirische Bildungsforschung als wichtigen Bezugspunkt, gegenüber der sich Evaluation durch eine stärkere Nutzenorientierung und die Generierung von Aussagen für konkrete Verwendungszusammenhänge auszeichne. Die Autorin fokussiert die Unterrichtsentwicklung als zentrales Anliegen von Evaluation und bezieht sowohl interne Evaluationen, externe Schulinspektionen und Verfahren der Selbstevaluation pädagogischer Qualität in der Schule auf dieses Anliegen. Nach der Systematisierung unterschiedlicher Adressatinnen und Adressaten von Evaluation stellt sie Verfahren, Methoden und Instrumente der internen, der externen und der Selbstevaluation dar und skizziert den Ablauf einer Evaluierungsmaßnahme. Die Aufzählung von Gelingensbedingungen und Problemlagen verweisen darauf, dass Evaluationen im Kontext Schule nach Einschätzung von Griese erfolgreich und nutzbringend sein können, aber bisher kein fester Bestandteil von schulischer Organisationskultur sind.

Evaluationen im Hochschulbereich behandelt Claudia Gómez Tutor. In diesem jungen Feld waren vor allem Demokratisierungsbestrebungen seit den 1970er Jahren sowie Fragen der Qualität und des zunehmenden Wettbewerbs von Hochschulen (Hochschulrankings) maßgeblich für die Etablierung von Evaluationen. Sie erläutert deren Bedeutung für die Qualitätsentwicklung in Lehre und Forschung, grenzt sie von anderen Konzepten ab und skizziert die Antinomien, innerhalb derer Evaluation an Hochschulen sich bewegt. Der Beitrag benennt Adressatinnen und Adressaten sowie Gegenstände, Verfahren, Methoden und Instrumente und konkretisiert diese am Beispiel des Ablaufs einer (primär auf quantitativen Verfahren basierenden) Evaluierung einer Lehrmaßnahme. Eine Darstellung zum Umgang mit Ergebnissen, zu förderlichen und hinderlichen Bedingungen sowie eine Diskussion zu veränderten Rahmenbedingungen der Hochschulen, verstärkten Wettbewerbsanforderungen und antinomischen Handlungslogiken runden den Beitrag ab.

Der Bedeutung, den Rahmenbedingungen und den Herangehensweisen an Evaluation in der Volkshochschule widmet sich der Beitrag von Evelyn Dahme. Sie konturiert zunächst das Handlungsfeld und organisationale Spezifika von Volkshochschulen als Institutionen der Erwachsenen- und Weiterbildung sowie die damit verbundene Bedeutung von Evaluation. Hierbei sind ökonomische Aspekte (im Unterschied zu mehrheitlich öffentlich finanzierten Kindertageseinrichtungen, Schulen und Hochschulen) von besonderer Bedeutung, nicht nur im Hinblick auf die Bewertung von Evaluationsergebnissen, sondern auch als Rahmenbedingungen der Institution und der Evaluierung. Sie spiegeln sich u.a. in Zielgruppen, Verfahren, Methoden und Instrumenten von Evaluation im Bereich der Volkshochschule wider. Anhand einer Fortbildung für ältere Menschen gibt Dahme einen Einblick in den konkreten Ablauf einer Evaluierungsmaßnahme und formuliert methodische Herausforderungen hinsichtlich spezifischer Adressatinnen

© Waxmann Verlag GmbH. Nur für den privaten Gebrauch.

und Adressaten. Abschließend reflektiert sie den Umgang mit Evaluationsergebnissen und zeigt unausgeschöpfte Potenziale von Evaluation im Bereich der Volkshochschulen auf.

Im Fokus von Nora Wegner steht Evaluation in bzw. von Museen und Ausstellungen. Sie beschreibt ein Spannungsfeld von künstlerisch-fachlicher Qualität und zunehmender Besucherorientierung, Legitimierung (Nachfrage, Konkurrenzdruck, Standorterhaltung) sowie daraus resultierender Skepsis gegenüber Evaluation als kennzeichnend für diesen Bereich. Neben verschiedenen Adressatinnen und Adressaten unterscheidet Wegner Verfahren von Evaluation hinsichtlich des Zeitpunktes ihres Einsatzes sowie verschiedene Methoden und beleuchtet schlaglichtartig Möglichkeiten sowie Risiken digitaler Untersuchungs- und Auswertungsmethoden. Anhand mehrerer Beispiele zu unterschiedlichen Evaluationsverfahren im Bereich von Museen und Ausstellungen, stellt sie den Ablauf von Evaluierungsmaßnahmen dar und gibt einen diesbezüglichen Einblick in den Bereich der Museen. Sie diskutiert Voraussetzungen für gelingende Evaluationen, fokussiert hierbei die Umsetzung von Ergebnissen und schließt mit einem Ausblick zu Perspektiven.

Unter dem Titel „Evaluationstheater? Möglichkeiten und Grenzen von Evaluierung im Bereich der darstellenden Künste“ nimmt Steffen Höhne Theater als Organisation kultureller Bildung in den Blick. Einleitend umreißt er die Beziehung von Kunst/Kultur zu ökonomisch präformierten Evaluationsverständnissen („Legitimationszumutung“, S. 261) und fordert für den Bereich der nicht kommerziellen darstellenden Kunst eine strikte Abgrenzung von betriebswirtschaftlicher Logik und Semantik in Evaluation. Ausführlich stellt er Spezifika des Kulturbereichs dar und erläutert die Relevanz von Evaluation bezogen auf deren verschiedene Zeitpunkte, Verfahren und Kriterien. Abschließend geht er beispielhaft auf eine bestimmte Organisationsform von darstellender Kunst/Theater ein und benennt exemplarische Ansätze der Evaluation und Qualitätssicherung.

Teil IV: Inhaltsdimensionen von Bildungsevaluation

Rebecca Lazarides und Sonja Mohr verantworten den ersten Beitrag im vierten Teil und führen in die Kompetenzmessung zur Evaluation von Schule und Unterricht im Rahmen internationaler und nationaler Vergleichsstudien ein. Sie zeigen am Beispiel der Mathematik, wie Kompetenzstufenmodelle aufgebaut sind und erläutern die Bedeutung psychometrischer Modelle für die Kompetenzmessung. Sie stellen den exemplarischen Ablauf einer Evaluierungsmaßnahme vor und benennen zentrale Bedingungen, die den Nutzen von Evaluationserkenntnissen für Lehrpersonen erhöhen können. In ihrem Ausblick verdeutlichen sie Fallstricke quantitativer Kompetenzmessung, verweisen auf deren Interpretationsbedürftigkeit sowie die begrenzte, einzelfallbezogene Aussagekraft und mahnen eine Untersetzung durch qualitative Forschung an.

Ausgehend von der These, dass sich (inter-)religiöse Kompetenz kaum standardisierbar erheben und messen lässt, schlägt Henning Schluß vor, qualitative Forschungsmethoden in der Kompetenzfeststellung (inter-)religiöser Bildung zu nutzen und auf Gruppendiskussionen, (Video-)Vignetten oder fiktionale Religionen (wie z.B. in der TV-Serie ‚Game of Thrones‘) zurückzugreifen. Letztere scheinen vor allem dazu geeignet, die für (inter-)religiöse Kompetenz zentrale Fähigkeit der Perspektivenübernahme weitgehend unabhängig von individuellen motivationalen und volitionalen Einflüssen evaluieren zu können. Mit Blick darauf, wie unterschiedlich religiös-kulturelle Praxen sein können, ist für Schluß unter religionsvergleichender Perspektive offen, inwieweit sich diese Praxen überhaupt in vergleichbare Kompetenzmodelle übersetzen lassen.

Gabriele Weiß setzt sich in ihrem Beitrag mit der Evaluation individueller Prozesse der kulturellen und ästhetischen Bildung auseinander. Kulturelle Bildung versteht sie als „Bildung in ästhetischen Dimensionen“ (S. 313), die ihrem Wesen nach nicht messbar, unkontrolliert und zweckfrei verläuft. Bestrebungen, die Qualität solcher Prozesse oder deren Outputs zu messen, sieht sie als Ausfluss der Ökonomisierung und Indienstnahme kultureller Bildung, die Suche nach Wirkungsnachweisen als Versuch der Herstellung von Gewissheiten in einer ungewissen Welt. Ausgangspunkt für Evaluation kann die Bestimmung von Merkmalen qualitativ guter ästhetischer Bildung sein. Dabei können die (Umsetzungs-)Qualität von Bildungsmaßnahmen oder die Qualität von Bildungsprozessen selbst im Mittelpunkt stehen. Die Autorin plädiert für ein partizipatives Vorgehen unter Rückgriff auf rekonstruktive qualitative Verfahren und stellt dies am Beispiel einer Museumsevaluierung vor. Mit Blick auf verschiedene Funktionen von Evaluation sensibilisiert sie für einen bedachten Umgang mit deren Ergebnissen und macht auf deren, dem performativen Charakter ästhetisch-kultureller Bildung geschuldete, kurze Halbwertzeit aufmerksam.

© Waxmann Verlag GmbH. Nur für den privaten Gebrauch.

Die Evaluation von berufsbezogener, zumeist betrieblicher, Weiterbildung thematisiert Ekkehard Nuissl. Er zeigt deren Besonderheiten auf, z.B. dass die Lernenden „Prosumenten“ sind, also (Ko-)Produzenten des Lernergebnisses sowie Konsumenten des jeweiligen Angebots und dass (betriebliche) Weiterbildung eine Kombination aus institutionalisiertem (formalem) und informellem Lernen ist. Der Autor erläutert dann die jeweils zu evaluierenden Bereiche und Merkmalsträger und schlägt vor, dass sich Evaluation auf die Untersuchung verschiedener Aspekte von Weiterbildungsqualität konzentriert. Dabei kann sie sich an gängigen Qualitätssicherungsverfahren orientieren und diese um eine Überprüfung der Passung der Maßnahme(n) zu den beruflichen Kontexten ergänzen. Im Sinne einer Zielerreichungskontrolle sind sowohl die Ziele der Teilnehmenden als auch die betrieblichen Kompetenzziele Gegenstand der Evaluation. Bevor er exemplarisch einen Evaluierungsverlauf in der berufsbezogenen Weiterbildung vorstellt, beleuchtet der Autor kurz die Unterschiede zwischen Evaluation sowie Controlling und stellt wesentliche Merkmale formativer und summativer Evaluation dar. Der Beitrag schließt mit dem Appell, im Evaluationsprozess über die Fokussierung auf Strukturen und Verfahren nicht den Menschen zu vernachlässigen.

Eine kurze Beschreibung der Eigenheiten und Ziele politischer Bildung (an Schulen) unternimmt Carl Deichmann. Politische Bildung begreift er als Teil eines gesamtgesellschaftlichen Kommunikationsprozesses, der auf politische Kultur einwirkt und selbst vielfältigen Einflussfaktoren ausgesetzt ist. Dementsprechend sieht der Autor zwei Ansätze für Evaluation: die Political-Culture-Forschung, mit ihren groß angelegten Einstellungsuntersuchungen, sowie wissenssoziologisch inspirierte, „‚interpretative‘ politische Kulturforschung“ (S. 361). Evaluation soll sich dabei sowohl auf die „Außenseite“ (Handeln) als auch die „Innenseite“ (Deutungsmuster) politischer Identitäten richten (S. 362). Sie könne sich an Lernzielebenen oder an Kompetenzen orientieren. Als Methoden stellt er wissenschaftliche Evaluation der Selbstevaluation von Lernenden und Lehrkräften gegenüber. Die entsprechenden Beispiele widmen sich schulorientiert zum einen qualitativer Unterrichtsforschung und zum anderen dem Verfahren der prozessualen Leistungsbewertung.

Anja Henningsen und Uwe Sielert beschäftigen sich mit der Evaluation sexualpädagogischer Bildung, insbesondere in Kontexten Sozialer Arbeit. Sie führen in die Funktion sowie in Kontexte sexueller Bildung bzw. sexualmoralischer Erziehung ein und umreißen die arbeitsfeldbezogenen Anforderungen an pädagogische Fachkräfte. Nach einem Exkurs zur (Nicht-)Geschichte der Erforschung sexueller Bildung, mit Fokus auf deren Wirkungen, konstatieren sie, dass es sich um ein empirisch bisher weitgehend unerforschtes Feld handelt. Vor diesem Hintergrund stellen sie einige laufende (Praxis-)Forschungsprojekte vor. Zugleich vermerken sie, dass die (Fremd-)Evaluation sexueller Bildung als Bemündigungspraxis unter den gegebenen gesellschaftlichen Bedingungen kaum absehbar ist. Daran anschließend beschreiben und erklären sie ausführlich das Vorgehen für Selbstevaluation, geben den Lesenden Indikatoren sowie Checklisten an die Hand und informieren über Fallstricke.

Diskussion, Fazit und abschließende Bemerkungen

Das Lehrbuch liefert einen guten, umfangreichen und zugleich kompakten Überblick zu Kontexten, Theorien, Debatten sowie ausgewählten Formen, Methoden und Instrumenten der Evaluation im Bildungsbereich sowie zu einer bemerkenswerten Vielzahl von Bildungsorganisationen und inhaltlichen Dimensionen von Bildung. Während sich Teil I intensiv und spezifisch mit historischen Entwicklungsverläufen, Kontexten und aktuellen Debatten aus dem Bildungsbereich beschäftigt, wird der Bezug zu Bildung in einigen Beiträgen des zweiten Teils vor allem durch die Vorstellung von Anwendungsbeispielen gesetzt. Insgesamt werden in Teil I und II zentrale Trends und für den Bildungsbereich spezifische Herausforderungen beschrieben (z.B. das Spannungsfeld der divergierenden Handlungslogiken von Evaluierenden und pädagogischen Fachkräften, Herausforderungen der Wirkungsfeststellung aufgrund der Offenheit von Bildungsprozessen und der Spezifik von Ko-Produktionsprozessen in Bildungssettings). Die Auswahl der in Teil III und IV dargestellten Bildungsorganisationen und -dimensionen wird von den Herausgebenden nicht expliziert; einige relevante Felder (z.B. außerschulische Jugendbildung) werden nicht systematisch abgedeckt. Zudem ist die Sortierung nach Bildungsorganisationen und inhaltlichen Dimensionen teilweise überschnittig, wie am Beispiel der ästhetisch-kulturellen Bildung deutlich wird.

Die Publikation wird mit ihrem einführenden Charakter einem Lehrbuch gerecht und liefert Anknüpfungspunkte für eine weitere, fundierte Auseinandersetzung mit Evaluationsmethoden, da die Beiträge diesbezüglich unterschiedlich voraussetzungsreich gestaltet sind.

© Waxmann Verlag GmbH. Nur für den privaten Gebrauch.

Grundsätzlich ist das Werk für Studierende sozial- und bildungswissenschaftlicher Bachelor- und Masterstudiengänge, für Lehrende und Lernende in der Fort- und Weiterbildung und für mit Evaluation befasste Mitarbeitende aus Bildungsorganisationen, die über Grundkenntnisse der empirischen Sozialforschung verfügen, geeignet. Die Umsetzung in Form von weitgehend in sich geschlossenen Einzelbeiträgen, die – je nach Interesse und Bedarf – auch für sich gelesen werden können, hat eher Sammelbandcharakter. Dies zeigt sich u.a. darin, dass es stellenweise zu Wiederholungen kommt (z.B. zu historischen Entwicklungslinien) und Verweise zwischen den einzelnen Beiträgen größtenteils fehlen. Zusammenfassende und bilanzierende Abschnitte zu den vier Teilen sowie Querverweise wären jedoch, insbesondere für Studierende mit wenig Vorwissen oder für die Zielgruppe der mit Evaluation befassten Mitarbeitenden von Bildungsorganisationen, hilfreich, um eine Einordnung – auch vor dem Hintergrund einer heterogenen Autorschaft, die teilweise mit unterschiedlichen Evaluationsverständnissen, Prämissen und methodischem Background argumentiert – zu erleichtern. Der didaktische Charakter der Publikation wird insbesondere in den Aufgabenstellungen und Fragen, die jeden Beitrag abschließen und zu einer vertieften Beschäftigung mit den jeweiligen Themen anregen, sichtbar. Die Aufgaben geben dabei den Auftrag zur gedanklichen Auseinandersetzung mit dem Text und teilweise auch zu einer umfangreichen, weiterführenden Recherche und fördern die Anwendung der Inhalte auf eigene Fragestellungen aus dem Bereich der Konzeptualisierung und Durchführung von Evaluationsprojekten. Dies entspricht dem selbstgesetzten Auftrag, einen Beitrag zur Förderung der Forschungsexpertise in unterschiedlichen Bildungsbereichen zu leisten.

UNSERE BUCHEMPFEHLUNG

2016, 124 Seiten, br., 19,90 €,
ISBN 978-3-8309-3514-8
E-Book: 18,99 €,
ISBN 978-3-8309-8514-3

Andreas Langer, Fabian Frei
Hände für Kinder – Kupferhof gGmbH (Hrsg.)

Kurzzeitwohnen für Kinder und Jugendliche mit Behinderung und deren Familien

Eine wissenschaftliche Evaluation im Neuen Kupferhof

Das Kurzzeitwohnen im Neuen Kupferhof in Hamburg ist ein innovativer und familienorientierter Ansatz, um Kinder, Jugendliche und junge Erwachsene mit Behinderung gezielt zu fördern. Ziel ist es, Identität und Persönlichkeit der Kinder und Jugendlichen zu stärken und gleichzeitig die Teilhabe am familiären und gesellschaftlichen Leben zu ermöglichen. Der vorliegende Band evaluiert Strukturen, Prozesse und Ergebnisse der Pionierphase dieses Angebots und liefert so einen Beitrag zum Übergang der Einrichtung in den Regelbetrieb.

www.waxmann.com

© Waxmann Verlag GmbH. Nur für den privaten Gebrauch.