

Rezension zu:

Hennefeld, Vera/Meyer, Wolfgang/Silvestrini, Stefan (Hg.): Nachhaltige Evaluation? Auftragsforschung zwischen Praxis und Wissenschaft. Festschrift zum 60. Geburtstag von Reinhard Stockmann. Münster: Waxmann, 2015. 316 Seiten, 34,90 EUR, ISBN 978-3-8309-3245-1

*Ulrich Degen*¹

Der Festschrift-Sammelband schlägt nicht ganz ohne Absicht im Internationalen Jahr der Evaluation² mit 15 Einzelbeiträgen in vier Kapiteln zwei Seiten der Medaille ‚Evaluation‘ auf: Er dokumentiert einerseits die Entwicklung der deutschen Evaluierungsforschung von den Anfängen bis zu ihrer heutigen Professionalisierung in Lehre und Forschung sowie in ihrer institutionellen Verankerung und mit ihren Kommunikationsinstrumenten, andererseits werden durch die Beiträge die bisherigen direkten und indirekten Leistungen des Jubilars in ihren vielen Facetten des praktischen, theoretischen und wissenschaftsstrategischen Wirkens von Reinhard Stockmann dokumentiert. Durch die Beiträge wird für die Leser(innen) damit unmittelbar deutlich, dass es ohne Reinhard Stockmann die jetzige wissenschaftlich innovative und politikkritische und praxisgeleitete Evaluationsforschung in der deutschen Sozialwissenschaft so nicht geben würde. Dass Stockmann die Evaluationsforschung als angewandte Sozialforschung über die Analyse der Nachhaltigkeit von Projekten der Entwicklungszusammenarbeit elaboriert, ist kein Zufall, stieß er sich doch schon früh an der wenig wirkungsanalytischen Ausrichtung von Bewertungen dieser Projekte und ließ nicht nach in seinem Bemühen, solche Projekte der namhaften Träger von Entwicklungszusammenarbeit systematisch einer fundierten Bewertung, eben Evaluation, zu unterziehen und die Bedeutung der Evaluation in

diesem Feld der Politik überzeugend deutlich zu machen.

Die vier Kapitel ‚Evaluation als Notwendigkeit? Nachhaltige Modernisierung‘, ‚Evaluation als Profession? Nachhaltige Forschung und Lehre‘, ‚Evaluation als Entwicklung? Nachhaltige Nutzung‘ und ‚Evaluation als Erfolg? Nachhaltige Wirkungen‘ klappen vier Domänen auf, die Reinhard Stockmann maßgeblich geprägt hat: Gesellschaftlich oder staatlich geförderte (Groß-)Projekte kommen heute und in Zukunft nicht mehr ohne wissenschaftsbasierte Evaluation aus, die Professionalisierung der Evaluation muss sich stärker als bisher in Forschung und Lehre niederschlagen und bedarf einer universitären Generierung in Curricula (Studiengang). Wenn Politikberatung für Projekte nachhaltig sein soll, geht dies nur über theorie- und wissenschaftsbasierte Evaluationen und es muss zwei nationale Orte für die Dokumentation des Forschungs- und Praxiswissens ‚Evaluation‘ geben: eine nationale Institution, wie z.B. das CEval, und ein Kommunikationsinstrument für den nationalen und internationalen Wissens- und Erfahrungsaustausch der Community und die Fortentwicklung der Evaluierungsforschung, wie die Zeitschrift für Evaluation.

Den interessierten und unvorbelasteten Leser(inne)n wird bei der Lektüre der Kapitel und der Einzelbeiträge des Sammelbandes deutlich, dass sich hinter dem bisherigen und aktuellen Wirken von Reinhard Stockmann ein aus-

1 Bonn

2 Stefanie Krapp weist in ihrem Beitrag „Evaluierungen – so what? Voraussetzungen für die Nützlichkeit von Evaluierungen in der Entwicklungszusammenarbeit“ (S. 227-241) darauf hin, dass die Generalversammlung der Vereinten Nationen (UN) am 19. Dezember 2014 erstmals in ihrer Geschichte eine solche Resolution zur Unterstützung von Evaluierung verabschiedete. Sie stellt dort kurz diese Resolution dar und unterstreicht insbesondere, dass hier eine Berichtspflicht über die Fortschritte derselben verankert wurde.

gefeilter und langfristig angelegter Plan zur Etablierung einer Evaluationsforschung und -praxis verbirgt, der Projekte der deutschen Entwicklungszusammenarbeit zum Ausgangspunkt nahm, um Evaluierung zu einer in Politik und Praxis in Deutschland anerkannten sozialwissenschaftlichen Wirkungs- und Anwendungsforschung zu führen. Damit beeinflusste er maßgeblich und entscheidend die Standards der Evaluationsforschung in Deutschland und trug dazu bei, dass Deutschland Anschluss an die Entwicklung der in diesem Feld weiter fortgeschrittenen Nationen Europas und Nordamerikas fand. Dass es ihm ein besonderes Anliegen ist, die Evaluationsforschung insbesondere auf dem südamerikanischen Kontinent zu etablieren und heimisch zu machen, ist auch seinen vielfältigen frühen entwicklungspolitischen Aktivitäten in seiner Forschungskarriere geschuldet.

Alle Autorinnen und Autoren des Sammelbandes sind in der Evaluierung auch internationaler Projekte forschungs- und praxiserfahren und berichten über die Entwicklung der Evaluationsforschung und -praxis aus ihren jeweiligen Untersuchungs- und speziellen Kompetenzfeldern. Das Herausgebertrio (Vera Hennefeld, Wolfgang Meyer und Stefan Silvestrini) leitet den Band mit einem selbstkritischen Blick auf die bisherige Entwicklung der Evaluationsforschung und der Evaluierung als Steuerungsinstrument ein und charakterisiert die vier Buchabschnitte mit ihren Einzelbeiträgen. Sie verdeutlichen das differenzierte Wirken von Reinhard Stockmann bei der Entwicklung und dem Ausbau der Evaluierung in Deutschland in Lehre und Forschung, Aus- und Weiterbildung und Durchführung von Evaluationsdienstleistungen. Dass für Reinhard Stockmann noch nicht das ‚Ende der Fahnenstange‘ erreicht ist, wird in allen Beiträgen deutlich. Er treibt weiterhin das ‚Evaluation Capacity Building‘ voran und damit die weltweite Verbreitung der Evaluation. Dass die ‚Nachhaltigkeit der (im Sammelband, Anmerkung des Rezensenten) beschriebenen Erfolge des Wirkens Reinhard Stockmanns und auch der Evaluation insgesamt als wichtiges Steuerungselement (...) gegenwärtig noch nicht gesichert (ist)‘ (S. 12) merken die Herausgeber hier in der Einleitung an, gehen aber davon aus, dass die dazu nötige Überzeugungsarbeit auch künftig geleistet werden wird – von wem wohl? Dem Jubilar dieses überzeugenden und weiterführenden Sammelbandes. Wenn man sich mit Fragen der Evaluierung zu beschäftigen beginnt, sucht man in der Regel nach einem kompetenten ersten Überblick mit wissenschaftlichem Tiefgang und Anknüpfungsmöglichkeiten an aktu-

elle, drängende und weiterführende Fragen im Feld. Diesen Überblick findet man in diesem Sammelband mit Beiträgen des State of the Art, so dass dieser durchaus auch gut als ‚Einführung in die Evaluierung‘ dienen kann. Reinhard Stockmann wird dies nicht als ‚Konkurrenz‘ zu dem von ihm herausgegebenen Standardwerk ‚Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder‘ (2006, Bd. 1 der Reihe Sozialwissenschaftliche Evaluationsforschung, Münster: Waxmann, 3. Aufl.) ansehen.

Die Autor(inn)en sind langjährige Gefährt(inn)en, Wegbegleiter(innen) und Mitstreiter(innen) in der Sache, auch Mitarbeiter(innen), die etwas von ihrem Geschäft verstehen und ebenfalls beim Aufbau der Evaluierung als Sozialwissenschaft der Moderne beteiligt waren. Sie alle ließen sich – wie auch der Jubilar – im Prinzip davon leiten, wie es Wolfgang Meyer in seinem Beitrag ‚Reflexive Modernisierung – oder doch: Evaluation als Ausrede?‘ (S. 53-72) so treffend ausgedrückt hat, dass ‚für die Modernisierung moderner Gesellschaften (...) die Nutzung steuerungsrelevanten Wissens, welches durch Evaluationen produziert wird, von zentraler Bedeutung (ist)‘ (ebd. S. 53). Und dass man gegenhalten muss gegen die ‚Gegenkräfte zu Monitoring und Evaluation‘ (ebd. S. 69) die sich in vielfacher Weise und Form zusammmentun: ‚als interessengeleitete Lobbyisten, die über Gebühr von staatlichen Investitionen profitieren wollen; als Technokraten, deren Weltbild bestimmte Handlungsalternativen prinzipiell ausschließt und denen kollektives Experimentieren zuwider ist; als gewinnmaximierende Consultants, die Evaluation als neues Geschäftsfeld sehen und primär an guten Einnahmen interessiert sind; als bornierte Manager, für die jede Handlungsempfehlung jenseits ihrer eigenen Präferenzen irrelevant ist, als dauerbedürftige Zielgruppen, welche auf die zusätzlichen Einnahmen aus Transferleistungen nicht verzichten wollen; als skandalisierende Medien, die Rückschläge und Misserfolge personalisieren und zum ‚No-Go‘ stilisieren und schließlich auch als gutgläubige Öffentlichkeit, welche populistischen Argumentationen mehr Aufmerksamkeit und Glauben schenkt als sorgfältiger wissenschaftlicher Forschungsarbeit‘ (ebd. S. 69-70).

Also genug zu tun in einem solchen Spannungsfeld für alle, die sich auch in Zukunft einer unabhängigen Evaluationsforschung widmen wollen. Darauf machen die Beiträge des Sammelbandes ebenso aufmerksam und unterfüttern, dass es sich mittel- und langfristig lohnt, professionelle Evaluierung in den zentralen Arbeitsfeldern, in denen Evaluationsforschung stattfindend

kann, zu betreiben: „dem Gegenstandsbereich von Evaluationen, Datenerhebungsmethoden, Bewertungskriterien, Evaluationsdesigns, Qualifikationen, Nutzen und Nutzung von Evaluationsergebnissen und der Evaluationspraxis“ (Silvestrini, Stefan: Der Beitrag Reinhard Stockmanns zur Evaluationsforschung – Eine (Zwischen-)Bilanz, S. 279). Dass hierbei Evaluationen grundlegende Qualitätskriterien erfüllen müssen, muss jedem Auftraggebenden und jeder/jedem Evaluierenden klar sein, denn sonst „ist ihr Nutzen gering, im schlimmsten Fall können sie zu Fehlentscheidungen führen“ (Caspari, Alexandra: Well done? Who knows ... Ein Plädoyer für Meta-Evaluationen, S. 161). Wenn Stockmann bei einer Evaluation der Entwicklungszusammenarbeit dabei ist, braucht einem hier nicht bange zu sein, da er „als Mäeutiker (...) die Kunst der Fragestellung (beherrscht), wobei er geduldig in stets engeren Kreisen um ein Problem herumgeht, bis er dessen harten Kern erreicht; als Methodologe kann er sich auf seine klare Struktur, seine große Stringenz und Prägnanz sowie seine hohe Anschaulichkeit berufen; als Vermittler kann er auseinandergelagerte Meinungen mit seiner Spontaneität und Witz sowie mit der Lockerheit seiner freien Rede zusammenführen“ (de Crombrughe, Dominique: Evaluierungssysteme in der Entwicklungszusammenarbeit. Oder was Evaluation mit der Völklinger Hütte zu tun hat. Einige Überlegungen aus Beobachtungen Reinhard Stockmann gewidmet, S. 225). Im Kern geht es darum, dass man eine Entwicklung beherrschen muss, insbesondere auch bei Entwicklungsmaßnahmen, bei denen „das Evaluandum (...) nicht nur komplexer (wird) als zuvor, (sondern) die Evaluation (...) auch besser als bisher die Komplexität von Entwicklungsmaßnahmen wiedergeben (soll)“ (ebd. S. 224). Dabei müssen sich wissenschaftlich durchgeführte Evaluationen wegen ihrer Reliabilität und Validität immer an sechs ‚Muss-Kriterien‘ orientieren und messen lassen:

- Sie müssen „auf einen klar definierten Gegenstand bezogen“ sein,
- für die „Informationsgewinnung (müssen) objektivierende empirische Datenerhebungsmethoden eingesetzt werden“,
- die Bewertung muss „anhand präzise festgelegter und offengelegter Kriterien“ erfolgen,
- die Evaluation muss „mit Hilfe systematisch vergleichender Verfahren vorgenommen werden“,
- in der Regel ist sie „von dafür besonders befähigten Personen (Evaluatoren)“ durchzuführen,

- „um auf den Evaluationsgegenstand bezogene Entscheidungen zu treffen“ (Silvestrini, S. 277).

Reinhard Stockmanns Beitrag zur Evaluationsforschung wird durch die Aufsätze des Sammelbandes methodisch, theoretisch und strategisch gewürdigt und die Beiträge spiegeln zugleich eine Beurteilung und Darstellung des Standes der Evaluationsforschung in Deutschland wider und zeigen auf, wohin die Reise in Lehre und Forschung und Institutionalisierung für eine stärkere Akzeptanz gehen kann. Explizit würdigt Stefan Silvestrini in seiner Zwischenbilanz am Schluss zusammenfassend Reinhard Stockmanns Beitrag zur Evaluationsforschung. Reinhard Stockmann kann man dabei kein größeres Kompliment als dieses machen, dass „bekanntermaßen (...) produktive Forscherinnen und Forscher mehr Raum (hinterlassen) für weiterführende Forschungsarbeiten als sie vorgefunden haben“ (Silvestrini, S. 285). Alle, die sich mit Evaluierung beschäftigen oder gedenken, dies künftig zu tun, sollten zu dieser Festschrift greifen, denn sie bietet einen schnellen, differenzierten und wissenschaftlich fundierten Zugang zur Evaluation; der Band sollte darum bei keiner/keinem tätigen oder angehenden Forscher(in) der Evaluation im analogen Bücherschrank fehlen!

Rezension zu:

Brunner, Andreas/Michel, Iris/von Mandach, Julie/Graf, Martin: Externe Schulevaluation im Kontext des schulischen Qualitätsmanagements. Bern: hep Verlag, 2015. 96 Seiten, 15,00 EUR, ISBN 978-3-0355-0247-3

*Katharina Klockgether*¹

Überblick

Das Buch informiert über das Qualitätsmanagement von Schweizer Schulen in vier Kantonen (Bern, Aargau, Luzern und Zürich) und die damit verbundene externe Schulevaluation. Entstanden ist es aus einem Pilotprojekt der Pädagogischen Hochschule Bern (PHBern). Die Hochschule entwickelte ein Konzept für externe Schulevaluation und testete dieses zwischen 2011 und 2013 an vier Berner Volksschulen.

Als Zielsetzungen des Buches werden angegeben:

- „die Diskussion der Bedeutung von externer Schulevaluation für Bildungspolitik und Schulentwicklung,
- die Darstellung der PHBern als Dienstleisterin für externe Schulevaluation sowie
- die Dokumentation der praktischen Verfahren der externen Schulevaluation.“ (S. 8)

Die Arbeit erscheint in der Reihe „Beiträge für die Praxis“ des hep-Verlags; ein Hinweis auf ihre praktische Ausrichtung. Es werden nun die auf die Einleitung folgenden Kapitel des Buches kurz vorgestellt:

Den Lesenden wird in Kapitel 2 der Wirkungsbereich der externen Schulevaluation anhand eines Modells von Landwehr (2011) verdeutlicht, das vier Evaluationszwecke unterscheidet (Wissensgewinnung, Rechenschaftslegung, Schulentwicklung und Normendurchsetzung). Dieses Modell wird später – auf der Grundlage der durch das Pilotprojekt gewonnenen Erfahrungen – weiterentwickelt. Daraufhin wird die aktuelle Praxis der externen Schulevaluation in den vier Kantonen veranschaulicht. Als Ausgangspunkt dient das „Logische Programmmodell: Externe Schulevaluation in der Volksschule“ von Ehren und Visscher

(2006). Im Modell werden als Programmelemente aufgeführt: *Kontext, Input, Incomes* und *Struktur*, das *Konzept* sowie das passende *Zielsystem*. Hinzu kommen einerseits die *Aktivitäten* und die daraus resultierenden *Outputs* und *Outcomes* der einzelnen Schulen, andererseits die Aktivitäten und Resultate auf Kantonenebene. Übergeordnet sind die *Impacts* und separat zu betrachten die *nicht intendierten Resultate*. Alle Elemente werden gemäß Eval Wiki: Glossar der Evaluation (vgl. Beywl/Niestroj 2009) definiert, analog zu dem von Univation entwickelten Programmbaum (vgl. Bartsch/Beywl/Niestroj 2015). Es folgt eine nähere Beschreibung der für das jeweilige Element relevanten Aspekte. Beispielsweise werden bei *Struktur* die institutionelle Verankerung, die Einbettung ins kantonale Qualitätsmanagement und in Netzwerke in Bezug auf externe Schulevaluation der vier Kantone beleuchtet. Als Datengrundlage dienen neben Literatur- und Dokumentenstudium auch Expertenaussagen. In der Schlussfolgerung werden die Charakteristika der Schulevaluationskonzepte der vier Kantone tabellarisch und grafisch in einem „Analyseraster“ und einem „Funktionsprofil“ dargestellt. Die Ausprägungen der vier adaptierten Zwecke von Landwehr werden eingeschätzt. So wird beispielsweise Normenkommunikation mittels klar definierter Qualitätsanforderungen im Kanton Bern eingeschränkt praktiziert. Im Folgenden wird diskutiert, welches Profil die meiste Wirkung in Bezug auf die Schulentwicklung entfalten kann. Dem Engagement der beteiligten Schulleitungen und Lehrpersonen dafür wird große Bedeutung zugemessen. Es wird erörtert, Ergebnisse der externen Schulevaluation

1 Univation – Institut für Evaluation Dr. Beywl & Associates GmbH, Köln

und Schülerleistungsdaten miteinander zu verknüpfen.

Kapitel 3 ist das ‚Marketingkapitel‘ der PHBern. Hier werden die Schulevaluationsangebote des Instituts für Forschung, Entwicklung und Evaluation erläutert. Neben einer Beschreibung der Vorteile von externer Schulevaluation erfolgt eine Reflexion zur Erreichung der Pilotprojektziele. U.a. wird beschrieben, aus welchen Gründen die „ursprünglich geplante Überführung in eine ordentliche Betriebsphase“ nicht realisiert werden konnte.

Der Praxisbericht zu den vier Pilotevaluationen findet sich in Kapitel 4. Das Design und die Umsetzung der externen Schulevaluation werden zu den fünf Teilprozessen Auftragsklärung, Durchführung der Evaluation, Berichtfassung, Ergebnisvalidierung und Ergebnispräsentation dokumentiert. Im abschließenden Teil „Möglichkeiten und Grenzen des bedarfsorientierten Verfahrens“ wird vorerst auf den großen Aufwand der im Pilotprojekt erprobten externen Evaluation sowie auf die fehlenden Vorgaben des Qualitätscontrollings der Erziehungsdirektion im Kanton Bern eingegangen. Es folgt eine Diskussion der Vorteile von externer Schulevaluation.

In Kapitel 5 wird von Andreas Brunner Bilanz gezogen. Insbesondere geht er dabei auf die Lage der Schulen im Kanton Bern in Bezug auf Qualitätsmanagement ein. Er zeigt kurz die Hürden und die Früchte des Pilotprojekts auf und verweist abschließend auf das nach wie vor bestehende Dienstleistungsangebot der PHBern.

Diskussion

Struktur und Layout: Das Buch ist gut strukturiert, an den passenden Stellen durch Abbildungen oder Tabellen ergänzt und dadurch übersichtlich, auch das Layout betreffend. Einzig bei Querverweisen zwischen den Kapiteln ist es mir schwergefallen, direkt die richtige Seite zu finden. Zudem könnte das Inhaltsverzeichnis übersichtlicher gestaltet werden.

Gegebenenfalls würde besonders Novizen im Thema und sehr praktisch Interessierten der Einstieg leichter fallen, wenn der Praxisbericht vor der Beschreibung der Programmelemente stünde. Denn die Darstellung der Programmelemente basiert u.a. auf den Ergebnissen und Erfahrungen der Praxisphase.

Inhalt: Die oben angeführten Ziele des Buches werden durchweg erreicht. Der Prozess des Pilotprojekts wird transparent und die gezogenen Schlüsse nachvollziehbar dargestellt. Die theoretischen Diskussionen sind mit externen

Quellen unterfüttert und gewinnen dadurch an Aussagekraft.

Erfrischend und professionell habe ich zudem empfunden, dass auch die Hürden rund um das Projekt und die Schwierigkeiten der externen Schulevaluation deutlich herausgearbeitet sind. Die Vorteile der externen Schulevaluation werden hingegen nicht immer klar. Beispielsweise wird von „verwertbaren“ Ergebnissen gesprochen; wie jedoch diese Verwertung an den Schulen aussieht, bleibt ungewiss. Darüber hinaus fehlt mir an einigen Stellen des Praxisberichts die Tiefe. Ich habe mir Fragen gestellt wie: Ist das Ziel der Ergebnisvalidierung im Evaluationsprozess – „Akzeptanz und Einsicht der Entscheidungsträger hinsichtlich der Evaluationsergebnisse zu gewinnen“ – bei den fünf teilnehmenden Schulen erreicht worden? Wie fielen die Ergebnisse der Zufriedenheitsbefragung der teilnehmenden Schulen in Bezug auf das Pilotprojekt aus? Wie wurden die Evaluationsergebnisse von den Schulen verwendet? Wie genau kann man sich die Beratung der Schulen im Nachgang zur Evaluation vorstellen? Zu den Abbildungen ist einschränkend zu erwähnen, dass an der ein oder anderen Stelle Informationen fehlen: Linien (S. 19), Sternchen (S. 74) oder Abkürzungen (S. 83) werden nicht erklärt.

Fazit

In diesem Band aus der Reihe „Beiträge für die Praxis“ der PHBern geht es um den spezifischen Umgang mit externer Schulevaluation einzelner Schweizer Kantone. Lesende erhalten einen Einblick in das schulische Qualitätsmanagement der vier Kantone sowie in die Beschreibung und kritische Reflektion des Pilotprojekts im Kanton Bern. Etwas unklar bleibt dabei die Zielgruppe des Buches. Nimmt man das Ziel der *Darstellung der PHBern als Dienstleisterin* in den Fokus, steht eine Leserschaft aus einerseits politischen Kreisen, andererseits von Schulleitungsebene im Vordergrund. Gegebenenfalls würde für diese Zielgruppe eine Broschüre ausreichen. Auf der anderen Seite stellt das Buch einen Beitrag für die Evaluationspraxis dar und spricht somit Evaluierende im Schulbereich an. Sicherlich können auch Anregungen für die Praxis außerhalb der vier Kantone und auch außerhalb der Schweiz gewonnen werden. Das kurze und übersichtliche Werk leistet neben seinem Beitrag für die Praxis außerdem einen Beitrag für die Professionalisierung der Evaluation, indem die Programmelemente systematisch definiert und leicht verständlich veranschaulicht werden.

Literatur

- Bartsch, Samera/Beywl, Wolfgang/Niestroj, Melanie (2015): Der Programmbaum als Evaluationsinstrument. In: Giel, Susanne/Klockgether, Katharina/Mäder, Susanne (Hg.): Evaluationspraxis. Professionalisierung – Ansätze – Methoden. Münster: Waxmann, S. 87-109.
- Beywl, Wolfgang/Niestroj, Melanie (2009): Der Programmbaum – Landmarke wirkungsorientierter Evaluation. In: Beywl, Wolfgang/Niestroj, Melanie (Hg.): Das ABC der wirkungsorientierten Evaluation. Glossar – Deutsch/Englisch – der wirkungsorientierten Evaluation. Köln: Univation, S. 137-149.
- Ehren, Melanie C. M./Visscher, Adrie J. (2006): Towards a Theory on the Impact of School Inspections. In: British Journal of Educational Studies, 54 (1), S. 51-72.
- Landwehr, Norbert (2011): Thesen zur Wirkung und Wirksamkeit der externen Schulevaluation. In: Quesel, Carsten/Husfeldt, Vera/Landwehr, Norbert/Steiner, Peter (Hg.): Wirkungen und Wirksamkeit der externen Schulevaluation. Bern: hep, S. 35-69.

Rezension zu:

Chen, Huey T.: *Practical Program Evaluation: Theory-Driven Evaluation and the Integrated Evaluation Perspective*. Thousand Oaks: SAGE Publications, 2015 (2. Auflage). 464 Seiten, 40,99 £, ISBN 978-1-4129-9230-5

*Uwe Schmidt*¹

Das Buch „Practical Program Evaluation“ von Huey T. Chen gibt eine umfassende Einführung in Grundlagen und insbesondere auch in den Transfer von Evaluationsmodellen in die Praxis. Die nachfolgende Besprechung des mehr als 400 Seiten umfassenden Werkes muss selbstredend fokussieren und ist darum bemüht, die spezifische Perspektive Chens zulasten einer weniger breiten Darstellung der Arbeit zu skizzieren. Gegliedert ist das Buch in fünf größere Abschnitte, die wiederum in insgesamt 15 Kapitel unterteilt sind. Die einzelnen Kapitel schließen jeweils mit Wiederholungsfragen zu den zentralen Inhalten ab, was auf die Intention eines Lehrbuchs zur modellgeleiteten Programmevaluation hinweist. Entsprechend sind die Adressaten des Buchs neben denjenigen, die in der Evaluationspraxis stehen, insbesondere auch Studierende.

Im ersten Abschnitt befasst sich Chen im Sinne einer Einführung mit den Grundlagen der Programmevaluation und rekurriert im Hinblick auf die Deskription von Programmen zunächst auf ein Input-Output-Modell, auf Interventionen und damit verbundene Transformationen sowie auf die Umweltbezogenheit von Programmen. Dies impliziert, dass Programmevaluationen nicht nur auf die Maßnahme als solche und ihre unmittelbaren Wirkungen, sondern in gleicher Weise auf die Einbettung des Programms und damit auf unterschiedliche Akteure Bezug nehmen sollten. Dies ist prägend für die Perspektive Chens und die von ihm entwickelten Charakteristika von Programmevaluation, die zwischen Theorie und Anwendung, zwischen Evaluatord(inn)en und Praxisakteuren auf unterschiedlicher Ebene vermittelnd wirkt. Im Hinblick auf die Ansätze der Pro-

grammevaluation unterscheidet Chen zunächst zwischen basalen Evaluationstypen, die sich zum einen im Sinne einer Matrix nach ihrer Orientierung auf Prozesse oder Outcomes, zum anderen nach einer konstruktiv-begleitenden oder einer konkludierend-bewertenden Perspektive differenzieren lassen. Er befasst sich anschließend mit den Schritten, die eine Evaluation durchläuft, und mit Evaluationsdesigns, wobei er auch hier stets die Einbettung von Evaluation in ein komplexes soziales System und die Nützlichkeit von Evaluation einfordert. Ein integrierter Ansatz bedeutet mithin, dass Evaluatord(inn)en Theorien und Ansätze entwickeln, die die Perspektive der Stakeholder und deren Praxis hinreichend integrieren.

Die zentrale Herausforderung ist hierbei, so Chen, Evaluation nicht nur als Methode zu verstehen, sondern als theoriegeleitet. Traditionelle Evaluationsansätze konzentrierten sich primär auf spezifische methodische Ansätze, wie experimentelle Designs, Fallstudien, Längsschnittuntersuchungen usw. Dies greife aber zu kurz: „However, because the focus of method-driven evaluation is mainly on methodological issues, it often does not capably address stakeholders’ views and needs. The theory-driven evaluation approach argues that while research methods are important elements of an evaluation, evaluation should not be dictated or driven by one particular method“ (S. 25).

Im Fortgang des ersten Abschnittes nimmt Chen entlang der einzelnen Stadien der Entwicklung und Durchführung von Programmen – Chen unterscheidet hier zwischen ‚Program-Planning Stage‘, ‚Initial Implementation Stage‘, ‚Mature Implementation Stage‘ und ‚Outcome Stage‘ – eine Zuordnung von möglichen Evalua-

1 Universität Mainz, Zentrum für Qualitätssicherung und -entwicklung

tionsansätzen und Evaluationsstrategien vor, die leitend für seine weiteren Überlegungen sind. Er greift hierbei mit Blick auf die Evaluationsansätze die Differenzierung nach den zuvor genannten basalen Evaluationstypen auf und ordnet diesen jeweils spezifische Strategien – wie Bewertung, Aufklärung, Problemlösung u.a. – zu. Unterhalb der Ebene der basalen Evaluationsmodelle differenziert er nochmals nach spezifischen Evaluationsansätzen und widmet im Folgenden vor allem dem ‚Logic Model‘ und mit Bezug zur Programmtheorie dem ‚Action Model‘ bzw. ‚Change Model‘ besondere Aufmerksamkeit.

Im zweiten Teil geht Chen näher auf die erste Phase, das Verhältnis von Evaluation und Programmsteuerung, ein und damit auf die Rolle der Evaluation im Kontext der Klärung und Konkretisierung der Programmplanung. Hiermit adressiert er eine Herausforderung, die im Rahmen nahezu jeder Evaluation von Bedeutung ist, indem die in der Regel in politischen Kontexten entstandenen allgemeinen Zielsetzungen konkretisiert und in Form eines ‚conceptual frameworks‘ eingeordnet werden. Chen unterscheidet hier zwischen ‚program scope‘ im Sinne eines Veränderungsmodells, das zudem Antwort darauf gibt, welches Problem und für wen dieses gelöst werden soll, und dem ‚action plan‘, der die Maßnahmen und deren Umsetzung konkretisiert. Er erläutert an Beispielen unterschiedliche Evaluationsansätze (formative research, conceptualization facilitation, pilot testing, advisory), die Stakeholder unterstützen können, ihre Projekte soweit zu konkretisieren, dass sie umsetzbar und evaluierbar werden.

Im daran anschließenden dritten Abschnitt widmet sich Chen der zweiten Phase, der Implementierung des Programms, und stellt auf den formativen, entwickelnden Charakter von Evaluation ab. Interessant ist hierbei seine deutliche Abgrenzung der formativen Evaluation als ausschließlich entwickelnde und nicht Qualität messende Methode: „It is a strictly developmentally oriented approach, and its results should be used only for timely fine-tuning purposes. Results of formative evaluation should never be used to describe an implementation’s quality (...)“ (S. 164). Während dies insbesondere in einer frühen Implementierungsphase relevant sei, gewinne mit zunehmendem Verlauf der bewertende Aspekt von Evaluation an Bedeutung. Hierbei lasse sich, so Chen, der besondere Mehrwert einer theoriegeleiteten Evaluation im Hinblick darauf beobachten, dass entgegen eines in erster Linie methodengeleiteten Ansatzes nicht intendierte Effekte von Programmen besser in

den Blick genommen werden können, da theoriegeleitete Ansätze eine rein projektimmanente Perspektive überwinden helfen.

Im vierten Abschnitt geht Chen auf die Bewertung des Outcomes ein und differenziert zwischen den Optionen eines Program-Monitorings und einer Outcome-Evaluation. Für letztere nimmt er entsprechend seiner eingangs skizzierten Typisierung von Evaluation eine weitere Differenzierung vor und erläutert in den Kapiteln 9 bis 11 die Ansätze der ‚Constructive Outcome Evaluation‘, der ‚Experimentation Evaluation‘ und der ‚Holistic Effectually Evaluation‘, um abschließend einen theoriegeleiteten Ansatz zur Evaluation des Outcomes zu erarbeiten. Er betont hierbei nochmals die besonderen Vorzüge einer theoriegeleiteten Vorgehensweise, die nicht nur den Outcome, sondern auch die kausalen Prozesse, die diesen bedingen, berücksichtige.

Der fünfte Teil des Buches schließlich befasst sich mit weiterführenden Fragen und Problemen der Programmevaluation. So erörtert Chen die Frage, wie zu verfahren ist, wenn logische Erklärungsmodelle nicht wie erwartet greifen (Kapitel 13), welchen Stärken und Grenzen eher formale oder Stakeholder-orientierte Ansätze unterliegen (Kapitel 14) und in welcher Weise die Ergebnisse von Evaluationen, sollten sie den Erfolg einer Maßnahme messen, in die Praxis transferiert bzw. auf weitere Handlungsbereiche übertragen werden können.

Zusammenfassend ist das Buch „Practical Program Evaluation: Theory-Driven Evaluation and the Integrated Evaluation Perspective“ von Huey T. Chen als äußerst lesenswert einzustufen. Dies vor allem, weil es die Zwischentöne der Evaluation in vielen Facetten herausarbeitet und sowohl die Nutzenperspektive von Evaluation als auch die Aufgabe von Evaluation, sich mit Wirkungen von Programminterventionen in systematischer Weise auseinanderzusetzen, fokussiert. Insofern gibt es vielfältige Anhaltspunkte für Evaluator(inn)en, die eigene Evaluationspraxis zu hinterfragen und weiterzuentwickeln. Wenngleich meines Erachtens der Ansatz einer ‚theoriegeleiteten Evaluation‘ eher dem eines modellgeleiteten Ansatzes entspricht, leistet Chen einen wichtigen Beitrag, Evaluation nicht nur auf eine programmimmanente und an Methoden orientierte Sichtweise zu reduzieren.

Kritisch zu betrachten ist gegebenenfalls die vorgesehene Adressierung des Buchs sowohl an Evaluator(inn)en als auch an Studierende. Zumindest für nicht bereits fortgeschrittene und mit Fragen der Evaluation partiell vertraute Studierende könnte sich die Lektüre trotz der resümierenden Fragen zu Ende der Kapitel als

etwas schwierig gestalten, da einige Bezüge und Abgrenzungen beispielsweise zu spezifischen methodischen Ansätzen, aber auch Querverweise auf Evaluationsansätze, durchaus voraussetzungsvoll sind. Die Arbeit Chens erhebt auch nicht den Anspruch, eine Einführung in die Evaluation zu leisten, womit der Bezug zu anderen

Evaluationsansätzen häufig kursorisch und wenig systematisch ist. Dies ist für mit Evaluationsansätzen und -praxis vertraute Leser(innen) eher bereichernd und anregend, für diejenigen, die sich erstmals Fragen der Evaluation zuwenden, vielleicht nicht in jedem Fall (einfach) zugänglich.