

## Rezension zu:

Patton, Michael Quinn: *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*. New York: Guilford Press, 2010. 375 Seiten. 35,99 EUR (Paperback), ISBN: 978-1-60623-872-1

Wolfgang Beywl<sup>1</sup>

Michael Quinn Patton zählt zu den Bestsellerautoren in der Evaluation. Die Auflage seiner Bücher geht auf die halbe Million zu. Sein Standardwerk „Utilization-Focused Evaluation“ (UFE) ist 2008 in der vierten Auflage erschienen. Dessen Kernaussage lautet: Der Evaluationsansatz muss je passend auf Zweck und Kontext gewählt werden, in einer Bandbreite zwischen qualitativer Einzelfallstudie und zufallskontrolliertem Experiment.

Das UFE-Auswahlmenü „Alternative Ways of Focusing Evaluations“ enthält 79 Ansätze. Ist die *Developmental Evaluation* (DE) einer von vielen? Oder öffnet sie eine neue Dimension, um die Evaluationswelt neu zu kartographieren?

Bis dato gilt ein zweidimensionales Koordinatensystem: Michael Scriven hat vor gut fünfzig Jahren – oft missverstanden – zwischen *formativer* und *summativer* Evaluationsrolle unterschieden. Diese grundlegende Dichotomie hat er in den letzten Jahren um die *askriptive* (erkenntnisfokussierte) Rolle ergänzt. Patton – sein ewiger Counterpart – arbeitet schon seit den 1980er Jahren mit der DE. Nun widmet er dieser zu Scrivens Dreier-Set zusätzlichen Rolle ein Lehrbuch. Ich möchte diese Rolle mit der rhetorischen Figur des Homoioteleutons (wiederholte Wortendungen) *evolutiv* nennen. Robert Stakes eingängige Analogie aufnehmend, hieße dies: *Formativ* ist, wenn der Koch die Suppe probiert; *summativ*, wenn es der Gast tut; *askriptiv*, wenn die Kulturhistorikerin zu ‚Suppe und Gesellschaft‘ arbeitet; – und *evolutiv*, wenn der in die Fremde verschlagene Koch sich rasch nach neuen Ideen für Suppen oder überhaupt neuen Spei-

sen umschaut, und nach Gästen, die mit ihm zusammen kochen und probieren (S. 22f.).

Die nachfolgenden Textpassagen akzentuieren die Besonderheiten der DE. Diese ist speziell auf innovative Programme zugeschnitten. Und: Sie fordert der Evaluatorin ein grundlegend verändertes Rollenverständnis ab:

„Das unterscheidende Merkmal entwickelnder Evaluation ist, dass sie zu etwas beiträgt, das entwickelt wird. Das ist der Zweck: Entwicklung. Die während der Evaluation gesammelten Daten ermöglichen eine schnelle glaubwürdige Rückmeldung für eine adaptive und responsive Entwicklung. Aber die entwickelnde Evaluatorin sammelt nicht nur Daten, wertet sie aus und schreibt einen Bericht. Entwickelnde Evaluation ist ein Prozess des Engagements. Die entwickelnde Evaluatorin ko-kreiert eine Innovation zusammen mit sozialen Innovatoren, indem sie das Wesen und die Folgen dieser Innovation untersucht“ (S. 305; wie weitere Zitate aus Patton vom Verf. übersetzt).

„Der aktive Part (...) zu sein ist, dass wir unsere eigenen Anliegen, Themen und Werte auf den Tisch legen. Die Evaluatorin ist also auch ein Stakeholder – nicht ein primärer Stakeholder. Aber: *In jeder Evaluation stehen ihre Reputation, Glaubwürdigkeit, Integrität und Überzeugungen auf dem Spiel*“ (S. 314).

Durch dieses Plädoyer für *aktives* Einmischen mögen manche in der Evaluationsgemeinschaft einen vermeintlichen Kern der Evaluation bedroht sehen, nämlich dass sie sich *immer* unabhängig und objektiv positionieren muss. Wer sich mit der evolutiven Evaluation auseinandersetzen will, sei es, weil er sein methodologisches

1 Institut Weiterbildung und Beratung, Pädagogische Hochschule FHNW, Aarau; Univation, Köln

Repertoire erweitern, sei es, weil er sich informiert distanzieren will, der findet in diesem Lehrbuch ein *pièce de résistance*.

Mit der besprochenen Publikation bereitet Patton seine Erfahrungen mit entwickelnder Evaluation auf. Er will sie erlernbar machen. Das Buch umfasst *zehn Kapitel*, differenzierte Personen- und Sachregister und ist auf „Google Bücher“ komplett durchsuchbar.

*Kapitel 1* definiert DE, identifiziert deren Position und Reichweite, erläutert das didaktische Konzept des Lehrwerkes und enthält Zusammenfassungen der folgenden Kapitel. Didaktisch setzt der Autor auf bis zu zehn Seiten lange Fallbeispiele. Sie berichten über eigene DE und Erfahrungen, die er mit konkreten Personen, den „Innovatoren“, gemacht hat. Dies soll die Entstehung des Ansatzes der DE nachvollziehbar machen und veranschaulichen, wie sie in verschiedenen Kontexten und Ausgangssituationen angewendet werden kann. Auf diese Geschichten kommt der Text immer wieder zurück. Er nutzt sie, um theoretische Annahmen und Erkenntnisse bei den Lesenden kognitiv zu vernetzen. Die Fallbeschreibungen müssen gelesen werden, will man dem Text folgen.

*Kapitel 2* „Entwickelnde Evaluation – als eigener Zweck und als Nische“ behandelt die beiden für die DE reservierten Nischen. Zum einen sei dies – temporär – die *präformative* Phase, in der Lösungen für ein noch nicht existentes Programm gesucht werden. Das Ergebnis könne, müsse jedoch nicht sein, dass sich ein Programmkonzept herauschält, welches nachfolgend (formativ) verbessert oder (summativ) bilanziert werden könnte. Zum anderen sei die DE – dauerhaft – für jene dynamischen Situationen bevorrechtigt, in denen die wichtigen Beteiligten davon ausgingen, oder sich ausdrücklich wünschten, dass das Programm auf unabsehbare Zeit richtungsverändernde Entwicklungen durchlaufen soll. Patton behauptet, dass hier schon formative und erst recht summative Evaluationen ins Leere laufen, d.h. nutzlos bleiben können. Dies geschehe z.B. bei Modellprojekten, die sich auf unbekanntes Terrain vorwagten. Hier existierten vorab geklärte und definitiv unplanbare Prozesse nebeneinander. Das intendierte Vorgehen erweise sich evtl. als gar nicht realisierbar und müsse kurzfristig, ggf. grundlegend, abgeändert werden. Patton illustriert dies an zwei DE innovativer Projekte zur Qualifizierung und Beschäftigung armer Menschen. Da diese sozial und psychisch vielfach belastet seien, erfordere bereits deren Gewinnung für das Programm krea-

tive Versuche. Bei Umschlagen sozialer Lebenslagen müsse das Programm immer wieder neu justiert werden. Dies ermögliche die DE mit ihrem empirischen Vorgehen.

*Kapitel 3* „Außerhalb des Denkens in Evaluationskästchen“ enthält das wichtigste Fallbeispiel des Buches. Dieses zeichnet die ‚Geburt‘ der DE in den 1980er Jahren nach: Patton war Kodirektor eines landwirtschaftlichen Bildungsprogramms in neun Karibikstaaten, das er zugleich evaluiert hat. Es wird herausgearbeitet, dass die DE zum Bestandteil der Programmsteuerung und zur treibenden Kraft der Programmentwicklung werden kann.

*Kapitel 4* „Systemdenken und Komplexitätskonzepte“ grenzt systematisch ein, für welche Evaluationsgegenstände die DE besonders geeignet ist, bisweilen sogar die einzige Wahl bleibt. Brenda Zimmermann folgend werden „einfache“, „komplizierte“ und „komplexe“ Situationen unterschieden. Diese Kategorisierung ergibt sich aus einer zweidimensionalen Betrachtung: erstens nach dem Grad der *Unsicherheit*, was in einer bestimmten Situation zur Lösung des Problems beitragen kann; zweitens nach Grad des *Konflikts* zwischen den Stakeholdern darüber, wie wünschenswert eine Intervention oder ein Programm ist. Seien beide Merkmale gering ausgeprägt, liege eine einfache, nach standardisierten Rezepten bearbeitbare Situation vor. Sei lediglich eine Dimension hoch ausgeprägt, existiere entweder ein „technisch“ oder ein „sozial“ kompliziertes Problem. Manifestierten sich beide Merkmale stark, liege eine komplexe Situation vor. In dieser sei die DE besonders tauglich. Als Beispiel aus dem westlichen Kulturkreis nennt Patton „Erziehung von Kindern“. Entscheidend sei, dass in komplexen Situationen der Ursache-Wirkungs-Mechanismus von Interventionen auf Resultate nicht nur unbekannt sei, sondern auch gar nicht gewusst werden könne, bis zu dem Zeitpunkt, an dem sich die Wirkung eingestellt habe.

*Kapitel 5* thematisiert „Systemdenken und Komplexitätskonzepte“ als Schlüssel, um komplexe Situationen zu analysieren. Veranschaulicht werden diese am Beispiel eines – vordergründig einfachen – Programms. Dieses hat zum Ziel, Polizisten einer amerikanischen Großstadt einen besseren Schlaf zu ermöglichen. Als für die DE relevante Komplexitätskonstrukte werden behandelt: „Nicht-Linearität“ (kleinste Handlungen können riesengroße Reaktionen auslösen); „Emergenz“ (das nicht vorhersehbare Interagieren von Individuen bis hin zu Systemen führt zu

Unvorhersehbares); „Adaptierung“ (Anpassung der Akteure an Umweltbedingungen); die bereits erwähnte „Unsicherheit“ über Ursache-Wirkungs-Zusammenhänge; „dynamischer Systemwandel“ (unvorhersehbare Muster des Wandels); „Koevolution“ (in der die Handelnden, zum Beispiel Innovatoren und Evaluatorinnen, schrittweise und im Wechselspiel handeln).

*Kapitel 6* „Wie die Welt verändert wird“ setzt sich kritisch sowohl mit dem Top-Down-als auch mit dem Bottom-Up-Ansatz für Innovationen und für ihre Evaluationen auseinander. Höchste Handlungsfähigkeit der Akteure wird in der Mitte verortet, „wo Wissen und Interessen aufeinander prallen, sich überkreuzen und sich verfangen“ (S. 152). Patton kritisiert einerseits den Top-Down-Ansatz der „Besten Praktiken“. Dieser unterstelle, dass Modellprojekte nach evaluativ nachgewiesener lokaler Bewährung national oder gar international vervielfältigt werden könnten. Dies müsse für komplexe Evaluationsgegenstände scheitern. Es könne nicht um „Skalierung“ oder „Multiplikation“ von Pilotprojekten gehen, sondern ausschließlich um lokale Übertragungen. Andererseits greife die auf die „Weisheit von der Straße“ vertrauende Bottom-Up-Gegenbewegung ebenfalls zu kurz. Dabei sei diese ebenso wenig verzichtbar wie das generalisierte Forschungswissen. Entwickelnde Evaluatorinnen müssten in beide Richtungen kritisch fragen.

*Kapitel 7* führt den „adaptiven Zyklus“ ein, der komplexe Evaluationsgegenstände kennzeichnet. Als Analogie dient dem Autor das ökologische System Wald. Es passt sich zyklisch an Umweltbedingungen (Feuer, Krankheiten, Trockenperioden) an. Der adaptive Zyklus enthält die vier Phasen „Freisetzung“, „Reorganisation“, „Erschließung“, „Erhaltung“. Soziale Systeme durchlaufen ähnliche Zyklen. Patton veranschaulicht den heuristischen Nutzen des adaptiven Zyklus an einem pädagogischen Programm, um Familien zu stärken und Kindesmissbrauch zu verhindern. Die besondere Stärke der DE liege in der explorativen Phase der Reorganisation, d.h. bei der Programmentwicklung.

*Kapitel 8* will klären, wie man die relevanten Fragestellungen für eine DE identifiziert. Es werden elf „Untersuchungsrahmen“ für folgende häufige Settings vorgestellt: krisengeschüttelte Umgebungen; Beginnen mit Evaluationsneulingen; Drängen auf eine steile Lernkurve; ausgeprägt positives Denken; nachdrücklicher Fokus auf Systemwandel; Kollaboration im Innovationsnetzwerk; wertgetriebene Innovation; an-

spruchsvolle Innovatoren; Insistieren auf Monitoring bzw. Vergleiche.

*Kapitel 9* enthält eine Sammlung von methodischen Zugängen für die DE. Patton verdeutlicht mit der Metapher des „Bastelns“ (*bricolage*), dass systematisches und geplantes Untersuchen schnell an Grenzen stößt und kreatives Kombinieren ausschlaggebend ist.

*Kapitel 10* zeigt resümierend die Verbindung zwischen UFE und DE auf. Patton betont – dem flexibel-adaptiven Denkgrundsatz treu bleibend, dass auch die DE keinerlei Präferenzen für irgendwelche Untersuchungsansätze habe. Ihr Herzstück sei das anspruchsvolle „logische Nachdenken“ (*reasoning*).

*Diskussion:* Die besprochene Publikation ist wichtig, besonders für die beiden „Nischen“ der befristet präformativen Evaluationsrolle und der dauerhaft emergenten Evaluationsgegenstände. Letztere sind typisch für das große Feld der „grundsätzlich vertrackten (*wicked*) Bildungsprogramme“. Ihnen geht es darum, Wissen, Urteilsvermögen usw. von Menschen in Kooperation mit ihnen zu fördern. Diese Aufgabe ist angesichts der Globalisierung des Wissens und des Lebenslangen Lernens allgegenwärtig. Die Handlungsbedingungen sind hier chronisch „komplex“: Emotionen und Werte sind ausschlaggebend für Entscheidungen auf allen Systemebenen; Unsicherheit über Ursache-Wirkungs-Beziehungen ist konstitutiv; Bildungsagenten und Lernende *müssen* interagieren. Viele Evaluationsansätze stoßen hier auf harte Grenzen: z.B. Hypothesenprüfen, kontrafaktische Methoden, evidenzbasierte Praktiken oder der „Ceteris-Paribus-Schabernack“ (S. 197). Bewertungskriterien wie Umsetzungstreue (*compliance*), Skalierung und Multiplizierbarkeit von Programmen sind oft unangemessen. Auch systematische Bedarfsanalysen, Operationalisierung von Zielen, Logische Modelle sowie Outcome-Measurement – allesamt Kernelemente der aktuellen Evaluationsmethodologie – passen häufig nicht. Verbreitete Evaluationsansätze griffen nicht nur im Bildungsbereich, sondern auch in Feldern wie Gesundheit, Soziales, Ökologie oder Entwicklungszusammenarbeit vielfach ins Leere. Es mag einerseits verwirren, dass Patton selbst – intensiv in UFE, aber auch in diesem Buch (wenn angemessen) – Logische Modelle nutzt, Programmannahmen expliziert, eine „theory of change“ formuliert und systematisch Forschungsergebnisse heranzieht. Andererseits ist genau dies Beleg für die konsequente Grundhal-

tung, jede Evaluation auf Kontext, Zweck usw. maßzuschneidern.

Das multidisziplinär angelegte Buch sprengt den Ring der Sozialwissenschaften: Es greift auf Managementlehren, Physik und Umweltwissenschaften zurück. Eklektisch greift der Autor über die eigenen Wissensgrenzen hinaus. Pattons Einschub „nicht alle sind bereit für diese Aufgabe“ (S. 79) kann als präventiv in Bezug auf die Fähigkeit, oder als resignativ in Bezug auf die Bereitschaft der Evaluationsgemeinschaft gelesen werden, die DE für den ausgewiesenen Geltungsbereich einzusetzen. Patton ist sich bewusst, dass das Buch „Anfeindungen“ bei denen auslösen kann, die auf das klassische Forschungsparadigma verpflichtet sind (S. 16).

Die möglicherweise größte Schwäche des Buches liegt darin, dass es erst ansatzweise darlegt, wie evolutiv Evaluierende ihre Gegenstände professionell untersuchen können. Es bietet im Unterschied zu UFE und zu Pattons meistverkauftem Buch (*Qualitative Research and Evaluation Methods*) wenig Anleitung für die gleichwohl essentielle empirische Untersuchungsarbeit („reality testing“, „real-time data“, „to track how such innovation occurs“; S. 35).

Der Stil des Buches wird manche begeistern, andere nicht. In seinem didaktischen Konzept ist das Narrative zentral. Die Fallgeschichten sind teils lang, und von manchen gibt es mehrere Versionen. Argumente werden wiederholt. Die Lesenden sollen offenbar in Text und Geschichten eintauchen.

Beim Publikum werden die Meinungen über das Buch je nach disziplinärem Hintergrund und erkenntnistheoretischem Grundverständnis auseinandergehen. Vielleicht eröffnet es einen Dialog über die Eignung oft mechanistisch eingesetzter Methoden und Designs. Im deutschsprachigen Raum gilt es, Anknüpfungspunkte herzustellen, etwa zur soziologischen Systemtheorie (z.B. betreffend Umwelt-System-Differenz) oder zur kritischen Handlungsfor-schung.

Aus meiner Sicht markiert Pattons *Developmental Evaluation* den Beginn einer weiteren Wende im Weltbild der Evaluation, das sich in den vergangenen 50 Jahren mehrfach stark gewandelt hat. Schon oft bin ich bei komplexen Evaluationsgegenständen vom geraden Pfad des empirischen Untersuchens abgewichen und habe Schleifen wie Organisationsentwicklung oder lösungsorientierte Beratung eingezogen. Nach der Lektüre des Buches kann ich besser begründen, dass dies keine willkürliche Vermischung von strengem Untersuchen und reflektierter Einflussnahme im System des Evaluationsgegenstandes ist. Es kann die dem Zweck – Entwicklung – einzig angemessene evaluative Vorgehensweise sein. Das Buch zeigt, dass DE nicht ein Kompromiss und nicht minderwertig gegenüber ‚wirklicher‘ Wissenschaft ist, sondern oft die bessere Wissenschaft.