

Editorial

Es ist sehr erfreulich, feststellen zu können, dass die Bedeutung von Evaluation und Evaluator(inn)en laufend im Zunehmen begriffen ist. Dies mag einerseits daran liegen, dass Evaluator(inn)en mit steigender Tendenz von Entscheidungsträgern zur Bewältigung komplexer Problemstellungen herangezogen werden und andererseits auch daran, dass eine anwachsende Methodenvielfalt zu verzeichnen ist.

Diesen Tendenzen entsprechend finden sich in Heft 1/2014 der *Zeitschrift für Evaluation* vier Originalbeiträge, nämlich ein literatur- und erfahrungsorientierter Überblicksartikel zum evaluativen Umgang mit Komplexität sowie drei methodisch orientierte Artikel aus dem quantitativen (2) und qualitativen (1) Forschungsparadigma. Im Informations- und Serviceteil finden sich mehrere Praxisberichte, gefolgt von Rezensionen aktueller Publikationen zur Evaluation. Über Aktuelles aus der Gesellschaft für Evaluation e.V. berichtet die Rubrik DeGeval ...Info.

Der erste Originalbeitrag von *Steve Fleischman* befasst sich mit dem Thema „Managing Complexity: How Evaluators Can Help Achieve Better Outcomes in Policy and Practice“. Es handelt sich dabei um die Adaptation eines Vortrages, der 2013 als Keynote-Lecture auf der Jahrestagung der DeGeval in München gehalten wurde. Der Beitrag beleuchtet die Tatsache, dass (politische) Entscheidungsträger immer häufiger auf das Know-how von Evaluator(inn)en zurückgreifen, wenn es darum geht, mit komplex ineinander greifenden Systemen umzugehen. Aus naheliegenden Gründen sind die Anforderungen hierbei so geartet, dass nicht immer Feststellungen getroffen werden können, die allen Ansprüchen der Evaluationsforschung genügen; vielmehr geht es dabei um Anwendungsforschung, die den Entscheidungsträgern direkt umsetzbare Anhaltspunkte und Nachweise liefern soll. Wie sich Evaluator(inn)en nach dem State of the Art in solchen Situationen zur Erzielung angemessener Ergebnisse verhalten sollen, wird von Steve Fleischman in Form von Verhaltensempfehlungen ausgeführt, die sich zum einen auf relevante Fachliteratur, zum anderen aber auch auf die spezifischen Erfahrungen des Autors als Direktor von Education Northwest, Oregon, stützen.

Für den Umgang mit Komplexität werden im Wesentlichen drei Punkte herausgearbeitet: (1) Es ist notwendig, den Sachverhalt zu akzeptieren, dass viele komplexe Phänomene das (wissenschaftliche) Leistungsvermögen des Evaluators/der Evaluatorin (zunächst) übersteigen und sich erst im Laufe der Zeit klarer entwickeln; zudem muss vielfach erst eine gemeinsame Sprache von Wissenschaftler(inne)n und Praktiker(inne)n gefunden werden; erschwerend kommt hinzu, dass Einsichten in

komplexe Systeme auch deshalb nur schwer gelingen, weil Vieles am menschlichen Verhalten nicht rational ist. (2) Es wird empfohlen, mit der gebotenen Umsicht eine systemische Perspektive einzunehmen und die vielfach bereichsübergreifenden Erkenntnisse der aufstrebenden Disziplin ‚complexity science‘ anzuwenden, die durch das Studium von Mustern und Tendenzen die Identifizierung von effektiven Wegen der positiven Beeinflussung der Systeme ermöglicht. Hierbei sollte bei der Nutzenbewertung vom ‚isolated impact‘ Abstand genommen werden zugunsten eines ‚collective impact‘. (3) Nahegelegt wird auch die Suche nach neuen und besseren Wegen der Zusammenarbeit mit Entscheidungsträgern durch Beachtung von ‚improvement science‘ (Erforschung von fortschrittsfördernden bzw. -behindernden Bedingungen) und insbesondere auch durch engere Forschungs-Praxis-Partnerschaften und Forschungsallianzen, um einen geeigneten Umgang mit Komplexität zu erleichtern.

Der zweite Originalbeitrag „Makroevaluation bei heterogener Programmumsetzung mittels mehrdimensionaler Typisierung“ von *Tim Stegmann* thematisiert das bei regionalisierten und dezentralisierten Arbeitsmarktprogrammen auftretende Problem, dass wegen zu geringer Fallzahlen nicht alle lokalen Strukturen und Strategien evaluiert werden können. Um dennoch zu geeigneten Fallzahlen zu gelangen, wird ein mehrdimensionaler clusteranalytischer Ansatz vorgeschlagen, bei dem jene Jobcenter, die unter verschiedenen Betrachtungsdimensionen eine hinreichende Ähnlichkeit ihrer Programmindikatoren aufweisen, zusammengefasst werden. Dementsprechend können die zugehörigen Stichproben so gepoolt werden, dass die Fallzahlen für eine Makroevaluation ausreichen. Konkret behandelt wird das Thema am Beispiel einer wissenschaftlichen Begleituntersuchung zur Evaluation des Bundesprogramms ‚Perspektive 50plus‘. Dem Ansatz der Politikfeldbezogenen Verwaltungsanalyse folgend wurden für die Typisierung der Jobcenter geschäftspolitische Ziele und Adressaten, arbeitgeber- bzw. teilnehmerseitige operative und instrumentelle Elemente, verwaltungsstrukturelle und prozedurale Elemente sowie beteiligte Umsetzer unterschieden, wobei sich die Jobcenter gemäß den Elementen der verschiedenen Dimensionen zu unterschiedlich vielen (zwei bis fünf) Clustern zusammenfassen ließen. Schließlich wurde auf der Basis von 10 Indikatoren in den drei Bereichen qualifikations- und tätigkeitsbezogene Komponenten, marktbezogene Komponenten sowie soziale Stabilität ein ‚Beschäftigungsfähigkeitsindex‘ gebildet, der zur Einschätzung von Fortschritten bei der Aktivierung von Teilnehmenden am Arbeitsmarktprogramm Verwendung fand, indem die individuellen Indexwerte über die Zugehörigkeit zu einem bestimmten Jobcenter mit den Typisierungen zusammengespielt wurden. Bei den aufgezeigten Effekten handelt es sich um Durchschnittseffekte für das jeweilige Cluster, welche zwar keine Bewertung der Arbeit der einzelnen Jobcenter, wohl aber eine makroevaluative Bewertung des Zusammenhanges von Umsetzungsvarianten und Programmindikatoren ermöglichen.

Anlass für den dritten Originalbeitrag „Zur Evaluation wissenschaftlicher Publikationsleistungen in der Sportwissenschaft“ von *Benedict Dessauer, Eike Emrich, Markus Klein* und *Christian Pierdzioch* gab das aktuelle CHE-Ranking in der Sportwissenschaft. Am Beginn ihrer Ausführungen nehmen die Autoren eine übersichtliche und detailreiche Betrachtung spezifischer Effekte des Einflusses von Rankings

auf den Wissenschaftsmarkt und insbesondere des CHE-Rankings auf die Sportwissenschaft vor, wobei auch einzelne ungeklärte Schwachstellen der Bewertungsprozedur aufgedeckt werden.

Im empirischen Teil untersuchen die Autoren in einer weitgehend CHE analogen Totalerhebung alle 42 sportwissenschaftlichen Institute an deutschen Universitäten und fokussieren vor allem den Bereich bibliometrischer Daten. Um die Validität und Sinnhaftigkeit des CHE-Rankings zu hinterfragen, werden verschiedene themenspezifische Hypothesen überprüft. So werden z.B. Vergleiche vorgenommen zwischen den Rangplätzen der Institute gemäß ungewichteter Kumulation des Publikationsoutputs, den gemäß CHE-Gewichtung transformierten Rangplätzen sowie solchen Rangplätzen, die auf einer weniger stark spreizenden, wenngleich ordinal unveränderten Gewichtung basieren. Die Autoren können zeigen, dass diese drei Herangehensweisen zu sehr unterschiedlichen Reihungen zwischen den einzelnen Instituten führen (bis zu 12 Rangplatzverschiebungen) und problematisieren, ob die CHE-Gewichtungen zu Recht jenen starken Einfluss auf das Publikationsverhalten der Sportwissenschaftler(innen) haben sollten, den sie zur Zeit faktisch ausüben.

Im vierten Originalbeitrag „Qualität für das Lernen? Konzeption einer Wirkungsstudie zum Qualitätsmanagement nach Q2E an beruflichen Schulen des Landes Bremen“ von *Helmut Ittner* und *Annikka Zurwehme* wird zunächst konstatiert, dass die Forschung zur Wirkung von Qualitätsmanagement im schulischen Kontext bisher nur wenig gesicherte Aussagen liefert. Deshalb propagieren Ittner und Zurwehme eine andere Methodik des empirischen Herangehens mit einer veränderten Betrachtungsrichtung, bei der als Ergebnis pädagogischer Lernprozesse nicht das Resultat fokussiert wird. Vielmehr soll – ausgehend von individuellen Begründungen – verstanden werden, welche typischen Begründungen dem Handeln und Denken der Beteiligten zugrunde liegen; das Ziel des Qualitätsmanagements sei also in einer Kontrolle des Begreifens und Verstehens der Intentionen der Beteiligten und nicht im Beschreiben der Abhängigkeiten zwischen Bedingungen und resultierendem Handeln zu sehen.

Die Konzeption der vorgestellten Wirkungsstudie ist dem qualitativen Forschungsparadigma zuzuordnen und geht von der These aus, dass ein lernförderndes Qualitätsmanagement intersubjektive Verständigungsprozesse zwischen Lehrenden und Lernenden wie auch zwischen Lehrenden und Leitungsverantwortlichen der Institution begünstigen müsse. Ein entsprechend konzipiertes und als ‚Q2E‘ (seit 2013 ‚QEE‘) bezeichnetes Qualitätsmanagement kommt seit 2002 an den beruflichen Schulen im Bundesland Bremen zum Einsatz. Nähere Daten zur Studie werden nicht mitgeteilt, so dass bis auf Weiteres offen bleibt, wie weit die subjektwissenschaftliche Sichtweise des Qualitätsmanagements Q2E geeignet ist, Wirkungen nachzuweisen, zumal das Verfahren, wie Ittner und Zurwehme konkludierend feststellen, weder im wissenschaftlichen noch im Praxisumfeld etabliert ist; es erscheint ihnen aber durchaus realistisch, an die Erweiterung schulischer Qualitätsmanagementsysteme in Hinblick auf eine Kontrolle des Begreifens und Verstehens der Intentionen der Beteiligten qualitätsverbessernde Erwartungen zu knüpfen.

Zusätzlich zu den vier Originalbeiträgen liefert das vorliegende Heft im Informations- und Serviceteil sowie in der DeGEval ...Info wiederum eine Fülle von Beiträgen.

Zunächst beschäftigen sich zwei Praxisberichte mit dem Einsatz qualitativer Evaluationsmethoden. *Stefan Mann* setzt sich mit der „Gesprächsführung in der Politikevaluation“ unter Anwendung der Methode der objektiven Hermeneutik auseinander. *Angela Bauer* beleuchtet die Vorzüge einer qualitativen Vorgehensweise bei der Evaluation eines pädagogischen Entwicklungsprojekts im Bereich der sozialen Arbeit. Der Tagungsbericht von *P. Linh Nguyen* gibt dann einen gemeinsamen Überblick über die Leitthemen der zentralen Konferenzen 2013 der International Development Evaluation Association (IDEAS) und der American Evaluation Association (AEA). Abgerundet wird der Serviceteil schließlich durch vier von *Christian Erzberger*, *Philipp Gonon*, *Jörn Gründer*, *Christoph Emanuel Müller* sowie *Wolfgang Meyer* verfasste Rezensionen aktueller Publikationen zur Evaluation.

Auch die DeGEval ...Info bietet wieder viele interessante Informationen, beginnend mit einem Bericht zur 16. Jahrestagung der DeGEval – Gesellschaft für Evaluation e.V. 2013 in München. Darüber hinaus sind die Laudationes zu den dort verliehenen Nachwuchs- und Medienpreisen der DeGEval wiedergegeben sowie Kurzfassungen der beiden mit Nachwuchspreisen ausgezeichneten Dissertationen, eine Übersicht über die Vorstandsmitglieder der DeGEval seit der Mitgliederversammlung 2013, das Positionspapier des Vorstandes der DeGEval zum Thema „Komplexität und Evaluation“, eine Kurzfassung des Arbeitspapiers Nr. 4 des Arbeitskreises Entwicklungspolitik in der DeGEval mit dem Titel „Systemische Ansätze in der Evaluation“ sowie abschließend zwei Beiträge des Arbeitskreises Evaluation in der Wirtschaft zum „Selbstverständnis des Arbeitskreises und Nutzen von Evaluation für die Wirtschaft“ und zu „Unterschiede[n] und Gemeinsamkeiten der Evaluation in Unternehmen gegenüber anderen Bewertungsverfahren“.

Explizit hingewiesen sei auch auf die Ankündigung der 17. Jahrestagung der DeGEval, die vom 10. bis 12. September 2014 in Zürich stattfindet, sowie die Ausschreibung für den Nachwuchspreis der DeGEval, der in diesem Rahmen verliehen wird.

Viel Vergnügen beim Lesen und viele interessante Anregungen!

Helfried Moosbrugger

Für die Herausgeberinnen und Herausgeber