

Editorial – Evaluation in Zeiten der digitalen Transformation

Im Zentrum der vorliegenden Ausgabe stehen Beiträge und Berichte in aktuell wichtigen Politikfeldern, die neue Möglichkeiten der Datenerhebung und -analyse aufgreifen und reflektieren. Stichworte wie Web-Society, Big Data und künstliche Intelligenz prägen derzeit die Diskussion um die Zukunft der Evaluation. Die vorliegenden Beiträge setzen hier an und zeigen auf, wie in Zeiten der digitalen Transformation und angesichts der Zunahme von viral verbreiteten Falschmeldungen valide, evidenzorientierte Evaluationen möglich sind und wie anspruchsvolle Evaluationen durchzuführen sind. Die erfreuliche Beobachtung der Zunahme an Evaluationen und die zunehmende Kompetenz der Evaluator(inn)en gehen dabei einher mit einem Anspruch, komplexere Methoden einzusetzen, die Theorie und Empirie zunehmend pragmatisch zu verbinden suchen. Informationsportale werden zu Lern- und Wissensressourcen über Evaluationsergebnisse. Dies erfordert Evaluationskompetenz aufzubauen. Die Beiträge belegen den fortschreitenden Aufbau von interner Evaluationskompetenz und die beachtliche Professionalisierung der Evaluator(inn)en. Zunehmend werden Arbeitsschritte entwickelt, die komplexere Analysen mit einer theoriebasierten Empirie erlauben und zeigen, wie valides, evidenzbasiertes Wissen produziert und vermittelt werden kann. Hinter dem Hype um die Digitalisierung stehen dabei zwei traditionell zentrale Fragen der Evaluation, die teilweise neue Antworten erfordern: Welchen Nutzen erbringen die zahlreicher werdenden Evaluationen und welche Verfahren gewährleisten durch die Einhaltung welcher Standards und Kriterien valide Ergebnisse?

Ein Ergebnis der zunehmenden Verfügbarkeit von Daten ist die wachsende Verbreitung von nationalen und internationalen Rankings. Diese dienen vielfach der Rechtfertigung und als Grundlage für Entscheidungen, z.B. für die Wahl eines Standortes, eines Krankenhauses oder eines Studiengangs. Meist liegen ihnen relativ einfache Kennziffern zugrunde. Am Beispiel eines kleinen Programms der Exzellenzinitiative zeigen *Klaus Wohlrabe* vom ifo Institut in München, *Lutz Bornmann* von der Generalverwaltung der Max-Planck-Gesellschaft, *Sabine Gralka* von der Technischen Universität in Dresden und *Felix de Moya Anegón* vom Institute of Public Goods and Policies (IPP) in Madrid, wie es gelingt, statt einfacher Kennziffernarithmetik mehrere Inputs und Outputs für eine Beurteilung der Leistung einzubeziehen. Nach einem Literaturüberblick über Effizienzanalysen an deutschen Universitäten demonstrieren die Autoren den Einsatz einer Data Envelopment Analysis (DEA) sowie ihrer Weiterentwicklung, einer Super-efficient DEA (SDEA), am Beispiel der Evaluation der Zukunftskonzepte von den 14 Universitäten, deren Zukunftskonzepte

zum projektbezogenen Ausbau der universitären Spitzenforschung im Rahmen der Exzellenzinitiative ausgezeichnet wurden. Ihre Vorgehensweise erlaubt eine verlässlichere Beurteilung der Effizienz beim Einsatz der Fördermittel für wissenschaftliche Spitzenleistungen. Die Analyse zeigt auch, dass bei kleinen Datensätzen deren Weiterentwicklung, die Super-efficient DEA (SDEA), eine bessere Differenzierung zwischen den Universitäten ermöglicht. Allerdings ist kritisch anzumerken, dass die DEA nur die relative Effizienz der beteiligten Einrichtungen und nicht deren absolute Effizienz misst. Dennoch dürfte die DEA, wie der informative Überblick der Autor(inn)en zeigt, inzwischen zum Standard für Effizienzanalysen im Rahmen der Evaluation der Produktivität von öffentlichen Gütern und Einrichtungen geworden sein.

Vermittelt der Beitrag einen wesentlichen Fortschritt in der Nutzung innovativer Methoden, so zeigt der Praxisbericht von *Andreas Armbrorst* und *Maria Walsh* vom NZK (Nationales Zentrum Kriminalprävention) beim Bundesministerium des Innern, für Bau und Heimat, wie durch Online-Portale der Transfer von Fachwissen gesteigert werden kann. Wissenschaftliche Erkenntnisse von Evaluationen von Sicherheits- und Präventionsmaßnahmen werden durch das Portal WESPE des NZK vermittelt. Sie umfassen Evaluationsergebnisse zu aktuellen Themen, wie Islamistische Radikalisierung, zu Sexualdelikten, zu Mehrfachtätern und Gewalt im Fußball. Ziel ist es, zu einem besseren Verständnis evidenzbasierter Prävention in Wissenschaft, Verwaltung, Politik und Praxis beizutragen.

Die Zusammenfassung von Studien nach spezifischen methodischen und inhaltlichen Kriterien in Systematic Reviews, Metaanalysen und Forschungssynthesen als Reaktion auf die Flut von Evaluationsstudien ist nicht neu, immer wieder wurde dazu in dieser Zeitschrift berichtet. Datenbanken zur Evidenzbasierung finden sich inzwischen für zahlreiche Bereiche, wie zur Bildung, zum Arbeitsmarkt, zur Gesundheit, der Justiz und Kriminalitätsbekämpfung. Federführend waren u.a. die Cochrane Collaboration und zunehmend das Netzwerk der Campbell Collaboration. Mit der Entwicklung eines Evidenzindexes liefert der Beitrag einen wichtigen Beitrag, wie Portale sinnvoll und pragmatisch Kriterien zur Beurteilung von Evaluationsergebnissen nach Art der Intervention, Methoden und Bewertung der wissenschaftlichen Güte darstellen können, um falschen Informationen entgegenzuwirken.

Wie wichtig und vielversprechend der Aufbau von Evaluationskompetenz ist, wird im Beitrag von *Henrik Brinkmann*, *Christiane Kerlen* und *Katharina Bilaine* veranschaulicht. Sie reflektieren ihre Erfahrungen aus einer Evaluation bei der Bertelsmann Stiftung. Nach einer ausführlichen Diskussion der Vor- und Nachteile interner und externer Evaluationen, wie sie vielfach bereits als Maßnahme der Evaluation Capacity Building (ECB) vorgestellt wurden, plädieren sie für die gegenseitige Ergänzung von interner Selbst- und Inhouse-Evaluation mit einer externen Evaluationsbegleitung. Ihre Studie zeigt die Bedeutung einer internen Evaluation für eine effiziente Nutzung von vorhandenem Evaluationswissen und gleichzeitig die Vorteile und Notwendigkeit einer externen Begleitung. Feldzugang, der Aufbau von Evaluationskultur, die Akzeptanz der Ergebnisse und deren Umsetzung werden in der Kombination eher gewährleistet. Der Aufbau einer internen Evaluationskompetenz und ihre externe Begleitung sichert Qualität und erhöht den Evaluationsnutzen. Er

reduziert auch Kontrolllängste und Risiken unerwünschter negativer Wirkungen auf das für die Governance von NPOs besonders wichtige interne Vertrauenskapital. Die Ergebnisse dürften über den Einzelfall hinaus für Evaluationen im Non-Profit-Sektor insgesamt relevant sein.

Der Frage nach den Einflussgrößen auf die Wirkungen einer externen Evaluation gehen *Kirsten Schweinberger*, *Sara Mahler* und *Carsten Quesel* in einer Evaluation von 25 Schulen in der Schweiz nach. Ihr Beitrag verdeutlicht die große Bedeutung einer theorieorientierten Evaluation. In einem Mixed-Methods-Längsschnitt-Design qualitativer und quantitativer Daten wird die Rolle von Schulgröße und zugestanderener schulischer Entwicklungsautonomie auf den subjektiven Wissensgewinn der Lehrpersonen durch eine Externe Schulevaluation untersucht. Ausgehend von einem Überblick von Studien zur Wirksamkeit einer Externen Schulevaluation wird ein theoretisches Modell zum Zusammenhang von Schulgröße und Wissensgewinn vorgestellt und empirisch überprüft. Es zeigt sich, dass die Schulgröße und der politisch definierte Spielraum für autonomes schulisches Handeln den subjektiv wahrgenommenen Wissensgewinn beeinflussen.

Wie eine theoretisch und methodisch anspruchsvolle Wirkungsanalyse mit einem quasi-experimentellen Forschungsdesign durchgeführt werden kann, zeigt *Andrea Löther* von der GESIS in Köln am Beispiel einer „Rigorosen“ Wirkungsevaluation des seit 2008 bestehenden Professorinnenprogramms. Pre-Post-Messungen von teilnehmenden und nicht teilnehmenden Hochschulen sowie eine Zeitreihenanalyse der Personaldaten erlauben eine erste Einschätzung der Wirkungen, auch wenn viele Fragen der hochschulinternen Wirkungen noch offenbleiben.

Die Verbindung von Theorie und Empirie ist auch das Anliegen von *Christian Diller*. Am Beispiel der Wirkung von Instrumenten der Stadterneuerungspolitik auf Gentrifizierungsprozesse diskutiert er bisherige Erfahrungen und notwendige Arbeitsschritte einer wirkungsorientierten, theoriegeleiteten Evaluation. Da die Evaluation in der Stadterneuerungspolitik eine lange Tradition besitzt, werden die unterschiedlichen Typen von Evaluationen (Kontextanalysen, Verlaufsanalysen, Ergebnisanalysen, Wirkungsanalysen) auf ihre Leistungen betrachtet, um dann, von den theoretischen Überlegungen der Gentrification-Forschung ausgehend, die Wirkungen der Instrumente der Stadterneuerung auf die unterschiedlichen Formen der Verdrängung (physisch, kulturell, ökonomisch) zu untersuchen. Hierzu werden Arbeitsschritte vorgestellt, wie diese Integration von Ansätzen zur Messung von Gentrification-Effekten in Evaluationen umzusetzen ist.

Ein wichtiger Teil jeder Ausgabe sind Berichte aus der Praxis und über die Evaluationsarbeit von Institutionen. Den Einsatz neuer Instrumente für die Evaluationsforschung durch die Digitalisierung diskutieren *Christoph E. Müller* von der InGFA (Institut für Gesetzesfolgenabschätzung und Evaluation) und *Maria Albrecht* von der HTW in Saarbrücken. Sie untersuchen die Möglichkeiten, Potentiale und Grenzen eines Crowdsourcing als einer neuen Möglichkeit der Datenerhebung bei Evaluationen. Crowdworker übernehmen bereits zahlreiche Dienstleistungen, wie die Prüfung der Funktionalität von Webseiten oder die Ausführung von Recherchetätigkeiten, die Erstellung und Korrektur von Texten oder die Anreicherung und Pflege von Datenbanken. In ihrem Beitrag geben die Autorin und der Autor einen Überblick über un-

terschiedliche Möglichkeiten, wie Crowdsourcing für die Datenerhebung eingesetzt werden kann, sei es für die Überprüfung von Erhebungsinstrumenten, für die Bildung von Vergleichsgruppen, bei der Durchführung von Pilotstudien oder zur Forschung über Evaluation, z.B. zur Überprüfung der Validität von Evaluationen. Allerdings besteht noch großer Forschungsbedarf für eine wirkliche Beurteilung der Aussagekraft, Validität und Reliabilität angesichts der meist heterogenen Stichprobenszusammensetzung.

In der Rubrik ‚Evaluationsinstitute stellen sich vor‘, zeigt der informative Bericht von *Andreas Armbrorst* am Beispiel des Nationalen Zentrums für Kriminalprävention, Bonn (NZK), wie wichtig und gleichzeitig wie schwierig die institutionelle Absicherung von Evaluationen im politischen Raum in der Praxis ist. Mit dem Ziel, Erkenntnisse über die Wirksamkeit von Ansätzen der Kriminal- und Extremismusprävention zu vermitteln, trägt das Institut zur Forschung und Verbreitung evidenzbasierter Prävention von Kriminalität und Extremismus bei. Wünschenswert wäre, noch mehr vergleichbare Berichte über die Institutionalisierung und von Ansätzen zum Ausbau der Evaluationskompetenzen zu erhalten.

Lesenswerte Ergänzungen zu den hochaktuellen Themen dieser Ausgabe finden sich in den Berichten von den lebhaften Frühjahrstagungen 2018 der Arbeitskreise ‚Methoden‘ und ‚Soziale Dienstleistungen‘. *Simone Ulrich* berichtet ausführlich über die Frühjahrstagung 2018 des Arbeitskreises ‚Methoden in der Evaluation‘ der DeGEval zum Thema Potenziale und Grenzen von Big Data in Evaluationen und *Stefanie Reiter*, *Frank Buchheit* und *Rainer Strobl* berichten über die „Evaluation deradikalisierender Maßnahmen – Wirkungen der Prävention gegen gewaltbereiten Extremismus“ des Arbeitskreises Soziale Dienstleistungen in der DeGEval, die in Kooperation mit dem Kompetenzzentrum zur Koordinierung des Präventionsnetzwerks gegen Extremismus in Baden-Württemberg erfolgte.

Die Themen dieser Ausgabe komplementieren ausgewählte Rezensionen wichtiger Veröffentlichungen. *Stefanie Krapp* diskutiert den von *Stockmann* und *Meyer* herausgegebenen bedeutenden Sammelband zur „Zukunft der Evaluation“, *Elisabeth Nindl* referiert kritisch die Veröffentlichung von *Petersson* und *Breul* (Hg.) „Cyber Society, Big Data, and Evaluation“ und *Wolfgang Meyer* diskutiert die gewichtige Habilitationsschrift von *Margit Seckelmann* zum Thema „Evaluation und Recht“.

Aufzufordern ist, sich an der DeGEval-Jahrestagung 2019 in Bonn mit dem Thema „Evaluation und Nachhaltigkeit“ zu beteiligen und Bewerber(innen) für den Nachwuchspreis zu benennen. Für den Nachwuchspreis sollten neben wissenschaftlichen Arbeiten auch vielversprechende und innovative Qualifikations- und Praxisarbeiten, wie Berichte und Aufsätze zu aktuellen Themen eingereicht werden.

Die Liste der zahlreichen Gutachter(innen) vermittelt einen Eindruck von der wachsenden Vernetzung der Evaluations-Community. Ihnen ist für die oft intensive, aufwändige, ehrenamtliche Unterstützung zur Sicherung der Qualität der Zeitschrift zu danken.

Für die Herausgeber(innen) ist es ein besonderes Anliegen, Professor em. *Helfried Moosbrugger*, dem von Beginn an über Jahrzehnte tätigen Mitherausgeber der Zeitschrift, als einem engagierten Unterstützer und Förderer zu danken. Mit seinem weiten Netzwerk von Kontakten, seinem breiten Methodenwissen und seinen Er-

fahrungen in der Evaluation hat er, stets reflektierend und sensibel argumentierend, mit Grazer Noblesse die Zeitschrift mitgeprägt. Am Deutschen Institut für Internationale Pädagogische Forschung in Frankfurt am Main, als Professor für Forschungsmethoden und Evaluation an der Johann Wolfgang Goethe-Universität, als Träger des Wissenschaftspreis der Deutschen Gesellschaft für Psychologie für Computergestützte Diagnostik, als Rezensent führender wissenschaftlicher Zeitschriften hat er wesentlich zur Verbesserung der Qualität von Evaluationen beigetragen. Wichtig für die Evaluation waren seine Arbeiten zur Statistik, zur Testtheorie und zur Fragebogenkonstruktion, aber auch zu inhaltlichen Themen, zu nennen sind Beiträge zur Eignungsbeurteilung oder zur Evaluation der Studierendenauswahl. Dabei hat er früh auf heute immer wichtigere Themen, wie die Bedeutung von Achtsamkeit als einer speziellen Persönlichkeitseigenschaft, hingewiesen. Wie wichtig seine Arbeiten zur Methodenentwicklung waren, zeigen die Beiträge in diesem Heft. Wir hoffen, dass wir einige seiner Vorstellungen und Ideen in dieser Zeitschrift weiter fortsetzen können.

Gerd-Michael Hellstern

Für die Herausgeberinnen und Herausgeber