EDITORIAL

Zeitschrift für Evaluation | 24. Jahrgang | 2025
Heft 2 | S. 191–197 | https://doi.org/10.31244/zfe.2025.02.01 |
© 2025 Waxmann

Liebe Leserinnen und Leser,

Evaluation und Wirkungsorientierung werden gerade in politisch turbulenten Zeiten neu entdeckt. Die prominent besetzte „Initiative für einen handlungsfähigen Staat" fordert in ihrem jüngst vorgelegten Endbericht (Initiative für einen handlungsfähigen Staat, 2025; Jäkel et al., 2025) an mehreren Stellen verstärkte ‚Wirkungsorientierung' ein. Der Koalitionsvertrag von CDU, CSU und SPD war schon stark vom Zwischenbericht der Initiative beeinflusst und greift diese Forderungen noch deutlicher auf: „Um den Wirkungsgrad von Gesetzen nachprüfbar zu machen, etablieren wir Erfolgsindikatoren, an deren Maßstab der spätere Gesetzesvollzug gemessen werden kann." (CDU, CSU & SPD, 2025, S. 44) Evaluationen werden im Koalitionsvertrag nicht nur abstrakt genannt, sondern in mehreren Fällen konkret gefordert, in so unterschiedlichen Feldern wie dem Bundesteilhabegesetz, den Registrierkassenpflichten, dem Selbstbestimmungsgesetz, der Krankenhausstrukturreform oder dem Angehörigenentlastungsgesetz.

Diesen beiden Ecksteinen der gegenwärtigen Diskussion um die sogenannte ‚Staatsmodernisierung' ging eine Reihe von Dokumenten aus dem politischen Raum voraus, in denen Wirkungsorientierung und Evaluation (beide Begriffe werden hier häufig synonym eingesetzt, aber dazu später mehr) eingefordert werden. Sowohl in Grundlagenpapieren der Parteien wie aus politikberatenden Institutionen heraus erfolgt eine neue Beschäftigung mit Evaluierung und Wirkungsfragen, die für Evaluation und Evaluationsforschung eigentlich hoffnungsvoll stimmen müsste. Scheinbar parteiübergreifend gibt es einen Konsens, Gesetzgebung und Gesetzesvollzug stärker wirkungsorientiert aufzustellen. Aus der CDU heraus gibt es schon länger Forderungen, Evaluation auch stärker im parlamentarischen Raum zu verankern, etwa durch sogenannte ‚Ziel- und Umsetzungswochen' im Bundestag, die sich auf die Diskussion von Umsetzungsberichten und Evaluationen konzentrieren sollen (Brinkhaus, 2024; Heilmann & Schön, 2020).

Auch von beratenden Institutionen kann in den letzten Jahren eine verstärkte Aufmerksamkeit für Evaluation berichtet werden. Anfang des Jahres führte die nationale Akademie der Wissenschaften Leopoldina eine Veranstaltung mit dem durchaus diskussionswürdigen Titel „Zehn Jahre Politikevaluation in Deutschland" durch. Angesichts der über fünfzigjährigen Geschichte von Evaluation und Evaluationsforschung in Deutschland könnte dieser Titel entweder als Provokation oder als Naivität interpretiert werden. Gemeint war allerdings der ungefähr zehnjährige Abstand zum Staatssekretärsbeschluss von 2013, der vorgibt, dass Gesetzgebungsvorhaben mit einem Erfüllungsaufwand von über einer Million Euro evaluiert werden sollen. Dieser 2019 erneuerte Beschluss wird vom Normenkontrollrat begleitet, der in seinen Jahres-

berichten regelmäßig berichtet. Der unglückliche Titel macht aber auch deutlich, dass Evaluation offensichtlich immer wieder neu erfunden wird. Dies wird auch in einem kurzen Papier der Veranstalterinnen deutlich, in dem durchaus unterstützenswerte Forderungen nach einer weitergehenden Institutionalisierung evidenzbasierter Politik vorgetragen werden (Diekmann et al., 2025; Riphahn & Schnitzer, 2025).

Weiterhin fehle eine klare Verantwortung für Evaluierungen, die Umsetzung in den Bundesministerien sei fragmentiert und es fehle an Qualitätssicherung und Datenzugang. Die Erkenntnisse aus den Evaluierungen würden zudem nicht systematisch und strukturiert genug genutzt. Die postulierten Forderungen nach einer stärkeren Verankerung der Evaluationsfunktion und -kompetenz in den Bundesministerien und einer evaluationstauglichen Gesetzgebung (sprich mittels klarer und messbarer Zielformulierungen) sind nicht neu. Die Nutzungsforschung und auch die Evaluation ziel- und wirkungsorientierter Steuerungsversuche hat allerdings auch systematische Barrieren einer konsequenten Evaluationskultur aufgezeigt, die in der gegenwärtigen Diskussion bereits wieder in Vergessenheit geraten zu sein scheinen. Der Erfahrungsschatz liegt jedoch vor.

In den 1960er Jahren erlebte Evaluation einen ersten Schub im Zuge der ambitionierten Reformprogramme der aktiven Politik. Eine zweite Welle ergab sich aus den in den 1990er Jahren intensiv diskutierten Programmatiken des New Public Managements. In beiden Phasen waren Forderungen nach Evaluation und evidenzbasierter Politik begleitet von Institutionalisierungs- und Professionalisierungsschüben die Folge, aus deren Erfolgen und Scheitern man durchaus lernen könnte.

Eine ähnliche Schlagrichtung nehmen die (für die oben skizzierte Staatsmodernisierungsdiskussion einflussreichen) Diskussionen um Wirkungsorientierung ein. Zwar glaubte man hier über den naiven Status der linearen Wirkungsmodelle aus dem New Public Management der 1990er Jahre hinausgekommen zu sein (siehe kritisch dazu im letzten Heft Haunberger et al., 2025; Hense, 2025). In Gutachten für das Finanzministerium (Deloitte & ZEW, 2024) oder einem bündigen Policy-Papier der Agora Digitale Transformation (Göller, 2025) werden aber wieder extrem simple lineare Wirkungsmodelle präsentiert. Auch wird hier deutlich, dass Evaluationen nicht trennscharf von anderen ‚bewertenden' Verfahren wie Monitoring und insbesondere ziel- und wirkungsorientierter Haushaltsführung unterschieden werden. Auch hier werden die Diskussionen um Programmtheorien und Wirkungsmodelle bislang nicht rezipiert oder man scheint aus Fehlentwicklungen der Vergangenheit nicht gelernt zu haben.

Teilweise hat man also den Eindruck, dass im politischen Diskurs um ‚Staatsmodernisierung' der Versuch unternommen wird, das Rad neu zu erfinden. Das mag ein generationaler Effekt sein, der geleistete Arbeit – gerade, wenn sie in digitaler Form nicht immer schnell vorliegt – schlicht nicht wahrnimmt. Auf der anderen Seite steckt dahinter vermutlich auch eine strategische Abgrenzung gegenüber der ‚alten' Evaluationsforschung und eine Positionierung als neuen Ansatz, der offensichtlich im politischen Raum auf offene Ohren stößt. Das vermeintliche Neuerfinden hat sicherlich auch einen strategischen Hintergrund – z. B. im Konkurrieren um für Evaluationen zur Verfügung stehende Ressourcen. Die in den letzten Jahrzehnten geleistete Arbeit droht hier verloren zu gehen. Angesichts immer offensichtlicher werdender Grenzen der evidenzbasierten Politikgestaltung – gerade in Zeiten politischer Polarisierung

und der durch Desinformation gesteigerten epistemischen Unsicherheit – wäre ein Aufgreifen der bisherigen Erfahrungen notwendig.

Auch in unserer Redaktions- und Herausgebertätigkeit wird immer häufiger deutlich, dass Autorinnen und Autoren zunehmend den Stand der Diskussion der letzten Jahrzehnte nicht mehr kennen. Einer der häufigsten editorialen Hinweise bei ‚desk rejects' oder ‚major revisions' ist es, die Papiere besser in den Stand der Evaluationsforschung einzuordnen. Wir als Zeitschrift sehen uns durchaus in der Pflicht, hier ‚pädagogische' Nachhilfe zu leisten. Dies ist in Feldern verständlich, die Evaluation gerade neu für sich entdecken – siehe z. B. den über mehrere Hefte verteilten Mikroschwerpunkt zu Evaluation in kirchlichen Arbeitsfeldern (Zimmer & Lämmlin, 2025), wird aber in etablierten Feldern problematisch.

Richtig ist, dass der Ausbau notwendiger Infrastrukturen deutlich defizitär ist. Zwar wurde bereits in der letzten Legislaturperiode mit einer Kompetenzstelle im statistischen Bundesamt eine kleine Einheit geschaffen, die die Bundesministerien beraten soll (Behre et al., 2023). Zudem wurden in einzelnen Politikbereichen wie der Entwicklungszusammenarbeit mit dem DEval schlagkräftige Einheiten geschaffen. Für ambitionierte Programme evidenzbasierter Politik müssten aber nicht nur die Mittel für durchzuführende Evaluationen bereitgestellt werden. Vielleicht wichtiger noch wäre eine Kompetenzentwicklung sowohl auf Auftraggebenden- wie auf Auftragnehmendenseite. Umso bedenklicher ist die Einstellung des letzten, explizit auf Evaluation ausgerichteten Ausbildungsangebots an der Universität des Saarlandes, welches sich in der Vergangenheit große Verdienste um eine Professionalisierung von Evaluation erworben hatte (Schopper, 2025). Die diesjährige Jahrestagung der DeGEval zum Thema Institutionalisierung von Evaluation ist leider eher ein Abschied denn ein Neubeginn einer akademischen Verankerung von Evaluation und Evaluationsforschung. Im kommenden Jahr wird die Zeitschrift für Evaluation 25. Inzwischen ist eine große ‚Backlist' von Artikeln elektronisch beim Waxmann Verlag zugänglich. Zum Jubiläum arbeiten wir daran, dass auch die frühen, aber immer noch aktuellen Hefte aus der Vor-Waxmann-Zeit elektronisch verfügbar werden, so dass es hinsichtlich der Verfügbarkeit von Evaluationswissen zumindest keine technischen Engpässe gibt.

Auch in den Beiträgen dieses Hefts geht es häufig um das Wiederentdecken und das Einordnen neuer Erkenntnisse und Verfahren in Evaluation und Evaluationsforschung.

Besonders deutlich wird dies im Beitrag zum Verhältnis von Organisation und Evaluation von *Moritz Ansmann* und *Markus Seyfried,* der die Reihe der Originalbeiträge eröffnet. Obwohl Organisationen sowohl als Evaluationsgegenstand, als Auftraggebende und Durchführende von Evaluationen alltäglich das Evaluationsgeschehen bestimmen, werden organisationstheoretische Ansätze in der Evaluationsforschung selten rezipiert. Die Autoren schließen diese Lücke durch eine Sichtung der einschlägigen Literatur und der Chancen, die eine Beschäftigung mit Organisationen und Organisationstheorie für das Evaluationsfeld bieten. Ausgehend von der naheliegenden These, dass alle Evaluationen auch durch ihre organisationalen Felder beeinflusst werden, plädieren sie für eine eingehendere Beschäftigung von Evaluator:innen mit Organisationstheorien. Eine Nicht-Berücksichtigung von Organisationstheorien könne bestimmte Erkenntnislücken zur Folge haben, deren Ergebnis ein verkürztes Verständnis der Funktionen von Evaluation in organisationalen Kontexten ist. Mit der Gegenüber-

stellung von instrumentellen und institutionellen Organisationstheorien werden die klassischen Evaluationsfunktionen neu interpretiert. Moritz Ansmann und Markus Seyfried liefern damit auch eine Brücke zu stärkerem interdisziplinärem Dialog und einer stärkeren Selbstreflexion.

Im zweiten Originalbeitrag von *Alexandra Totter* wird ein mehrperspektivischer und modularer Ansatz einer formativen Lehrmittelevaluation vorgestellt. Damit wird insbesondere ein Beitrag zur Einordnung der formativen Evaluation geleistet. Bereits während der Entwicklung eines Lehrmittels sollen forschungsbasierte empirische Ergebnisse zur Konzeption, Qualität und Akzeptanz untersucht werden, um deren Ergebnisse unmittelbar in den Entwicklungsprozess wiedereinspeisen zu können. Im Ergebnis sollen so Nutzen und Nutzung von Evaluationsergebnissen erhöht werden. Aus dieser Nutzenperspektive heraus kann auch eine weitere Institutionalisierung und Professionalisierung vorangetrieben werden. Die Erfahrungen der Entwicklungsteams mit formativen Evaluationen wurden aufgrund von datenbasierten Erkenntnissen so positiv wahrgenommen, dass eine Ausweitung dieses Vorgehens auch auf andere Bereiche erfolgte.

Dieses Heft enthält gleich zwei Originalbeiträge zur Rubrik *Methoden in der Evaluation.* Nachdem die ersten Beiträge zu dieser Rubrik vornehmlich quantitative Methoden in den Fokus nahmen, beschäftigen sich beide Beiträge mit der Weiterentwicklung *qualitativer* bzw. *fallbasierter Ansätze.*

Felix Hörisch und *Kai Brumm* führen in die Qualitative Comparative Analysis (QCA) ein, einem Ansatz der vergleichenden Fallanalyse, der die Brücke zwischen quantitativer und qualitativer Analyse zu schlagen sucht. Er eignet sich vor allem für die vergleichende Analyse kleiner und mittlerer Fallzahlen. Während QCA sich in anderen sozialwissenschaftlichen Teildisziplinen bereits stark etabliert hat, fristet sie in der Evaluationsforschung noch ein Schattendasein. Umso wichtiger ist es, dass die Autoren den Nutzen des Verfahrens für die Evaluationsforschung herausarbeiten, ohne in den Fachjargon zu verfallen, der oft eine große Einstiegshürde zur Beschäftigung mit QCA darstellt. Dabei stellen sie insbesondere den kausalanalytischen Beitrag heraus, den QCA in komplexeren Wirkungszusammenhängen leisten kann.

Um die Aufklärung von Wirkungszusammenhängen geht es auch im Beitrag von *Anna von Werthern* und *Elisa Knödler*. Im Mittelpunkt stehen theoriebasierte Herangehensweisen bei der Konstruktion von Programmtheorien im Rahmen theoriebasierter Evaluationen. Die Autorinnen stellen mit der *Success Case Method* (SCM) nach Brinkerhoff und dem *Semi Structured Interview Protocol* (SSIP) nach Gugiu und Rodriguez-Campos zwei qualitative Interviewmethoden vor, welche sowohl zur Generierung einer Programmtheorie als auch für die Erhebung von Daten im Evaluierungsprozess selbst zur Anwendung gebracht werden können. An einem Beispiel aus dem Hochschulkontext wird deren Applikation illustriert und Vor- und Nachteile werden diskutiert.

In den Praxisbeiträgen des Heftes wird ebenfalls der Zusammenhang zwischen Methoden und der Nutzung von Evaluationen thematisiert.

Daniel Deimel berichtet über eine Untersuchung der Wirkung von individuellen Schulrückmeldungen aus der International Civic and Citizenship Education Study (ICCS) 2022 in Deutschland. Die ergänzende Rezeptionsstudie evaluierte die Verständlichkeit und Nutzung der Schulrückmeldungen in der Schule. Entscheidend sei

die verständliche und praxisrelevante Aufbereitung der Inhalte. Klar strukturierte Informationen tragen dazu bei, dass die Rückmeldungen als hilfreich und unterstützend in der schulischen Arbeit bewertet wurden.

Dennis John, Joachim König, Dietmar Maschke und *Sebastian Ottmann* stellen den „Wirkungsradar" als partizipatives Methodenpaket für eine wirkungsorientierte Praxis dar. Die Autoren zeigen, dass Wirkungsorientierung nicht nur eine bloße Forderung ist, sondern vielfach etablierte Praxis. Am Beispiel kirchlicher Entwicklungsprozesse wird deutlich, wie Wirkungsorientierung im organisationalen Alltag gelebt werden kann.

Wie die Institutionalisierung von Evaluation auch erfolgen kann, erhellt oft ein Blick über die Grenzen. Am Beispiel der Schweiz werden häufig die Defizite des deutschen Evaluationssystems deutlich – sofern man hier überhaupt von einem System sprechen kann. In zwei Beiträgen beleuchtet *Reinhard Zweidler* neuere Entwicklungen in der Schweiz: Zunächst stellt er zusammen mit *Patrick Zenhäusern* das „Vademecum Ausschreibungen" der SEVAL vor. Gerade vor dem Hintergrund der in der DeGEval geführten Debatten zur Novellierung der „Empfehlungen für Auftraggebende" (Nowack et al., 2024) macht dieser erhellende Beitrag deutlich, dass gerade hinsichtlich der Ausschreibungs- und Vergabepraxis noch deutlicher Aufklärungsbedarf besteht. Mit der Vorstellung der SEVAL-Beratungsstelle wird ein weiterer Baustein zu einer besseren Gestaltung des Verhältnisses von Auftraggebenden und Auftragnehmenden in Evaluationen geliefert, der auch für die DeGEval interessant sein könnte. Sowohl präventiv wie im Konfliktfall werden hier Beratungsangebote bereitgestellt, die zu einer besseren Durchführbarkeit von Evaluationen beitragen können.

Eine Rezension des Buches von Baumgartner und Haunberger zu Wirkungsevaluationen in der Sozialen Arbeit von *Sebastian Ottmann* sowie eine Reihe von Beiträgen aus den vielfältigen Arbeitskreisen der DeGEval runden das aktuelle Heft ab. Auch in diesen spielt das Motiv des Wiederentdeckens älterer Ansätze eine Rolle. So findet sich im Beitrag des AK Demokratie die Einsicht, dass Selbstevaluation eigentlich ein altes Thema sei, welches regelmäßig wiederentdeckt wird und auch der AK FTI kommt bei der Weiterentwicklung von Indikatoren zur Wirkungsmessung immer wieder auf ähnliche Lernprozesse zurück. In eigener Sache sei insbesondere auf den Beitrag von *Evelyn Funk* und *Jessica Prigge* zum ersten Lesedialog des AK Professionalisierung mit und über die ZfEv verwiesen. Unter dem Motto ‚Lesen, diskutieren, voneinander lernen' fand dieser erste Dialog im Mai statt. Wir bedanken uns für diese großartige Idee. Der Termin zum Lesedialog über dieses Heft steht bereits fest und wir freuen uns über neue Gesichter und spannende Diskussionen! Bis dahin wünschen wir Ihnen eine anregende und nutzenstiftende Lektüre.

Stephan Grohs
Für die Herausgeberinnen und Herausgeber

Literatur

Behre, L., Gerls, F., & Kühnhenrich, D. (2023). Die Kompetenzstelle der Bundesregierung für Evaluierung im Statistischen Bundesamt. *Zeitschrift für Evaluation*, *22*(1), 172–181. https://doi.org/10.31244/zfe.2023.01.10

Brinkhaus, R. (2024). *100 Vorschläge für den Neustaat. Arbeitspapier*. https://ralph-brinkhaus.de/wp-content/uploads/2024/12/100-Vorschlaege-fuer-den-Neustaat_Ralph-Brinkhaus-MdB.pdf

CDU, CSU, & SPD. (2025). *Verantwortung für Deutschland. Koalitionsvertrag zwischen CDU, CSU und SPD. 21. Legislaturperiode des Deutschen Bundstags*. CDU-Bundesgeschäftsstelle. https://www.spd.de/fileadmin/Dokumente/Koalitionsvertrag2025_bf.pdf

Deloitte & ZEW. (2024). *Abschlussbericht für das Bundesminsterium der Finanzen. Erstellung eines Konzepts zur Einführung einer ziel- und wirkungsorientierten Haushaltsführung*. https://ftp.zew.de/pub/zew-docs/gutachten/240206_Abschlussbericht_Forschungsvorhaben_fe5-22-zwoH_Final.pdf

Diekmann, L., Riphahn, R.T., & Schnitzer, M. (2025). Evidenzbasierte Politik – Lehren aus deutschen und internationalen Erfahrungen. *Wirtschaftsdienst*, *105*(4), 288–291. https://doi.org/10.2478/wd-2025-0074

Göller, B. (2025). *Wirkungsorientierung als Weg zu echten Digitalstrategien der Bundesregierung*. Agora Digitale Transformation. https://doi.org/10.5281/zenodo.14987873

Haunberger, S., Braun, M., Hemmann, M., & Rehse, A. (2025). Wirkmechanismen auf der Spur. Zwei Anwendungsbeispiele der Realist Evaluation zur qualitativen Identifikation von Wirkmechanismen. *Zeitschrift für Evaluation, 24*(1), 9–39. https://doi.org/10.31244/zfe.2025.01.02

Heilmann, T. & Schön, N. (2020). *NEUSTAAT: Politik und Staat müssen sich ändern. 64 Abgeordnete & Experten fangen bei sich selbst an – mit 103 Vorschlägen* (3. Aufl.). Finanzbuch Verlag.

Hense, J. (2025). Nicht der Plan wirkt, sondern seine Umsetzung. Ein feldübergreifendes Kriteriensystem der Implementationstreue. *Zeitschrift für Evaluation, 24*(1), 40–69. https://doi.org/10.31244/zfe.2025.01.03

Initiative für einen handlungsfähigen Staat. (2025). *Zwischenbericht*. https://www.ghst.de/fileadmin/images/01_Bilddatenbank_Website/Demokratie_staerken/Initiative_f%C3%BCr_einen_handlungsf%C3%A4higen_Staat/20250311_Zwischenbericht_interaktiv.pdf

Jäkel, J., Maizière, T. de, Steinbrück, P., & Voßkuhle, A. (Hrsg.). (2025). *Initiative für einen handlungsfähigen Staat: Abschlussbericht* (Juli 2025). Herder.

Nowack, C., Jessing, D., Zweidler, R., & Gutknecht-Gmeiner, M. (2024). Ausschreibung von Evaluationsaufträgen. Bericht zur gemeinsamen Frühjahrstagung 2024 der DeGEval-Arbeitskreise Umwelt und Verwaltung unter Beteiligung des Arbeitskreises Gender Mainstreaming und der Ad-hoc-Gruppe Nachhaltigkeit. *Zeitschrift für Evaluation*, *2024*(2), 299–309. https://doi.org/10.31244/zfe.2024.02.12

Riphahn, R.T. & Schnitzer, M. (2025). *Effektives staatliches Handeln durch Lernen aus Evaluierungen. Leopoldina Fokus Nr. 4*. https://www.leopoldina.org/fileadmin/redaktion/Publikationen/Nationale_Empfehlungen/2025_Leopoldina_Fokus_Politikevaluierung_web.pdf

Schopper, S. (2025). Der Masterstudiengang Evaluation (MEval) in Saarbrücken. *Zeitschrift für Evaluation, 24*(1), 134–139. https://doi.org/10.31244/zfe.2025.01.09

Zimmer, M. & Lämmlin, G. (2025). Wirkt der Geist messbar? Evaluation in und für Kirche. Einführung in den Themenfokus Kirche. *Zeitschrift für Evaluation, 24*(1), 89–99. https://doi.org/10.31244/zfe.2025.01.05

Prof. Dr. Stephan Grohs | Deutsche Universität für Verwaltungswissenschaften Speyer | Lehrstuhl für Politikwissenschaft und Deutsches Forschungsinstitut für öffentliche Verwaltung | Freiherr-vom-Stein-Str. 2, D-67346 Speyer | E-Mail: grohs@uni-speyer.de