

## Schema F oder Innovationen? Vorletzte Runde und erste Zwischenbilanz einer politikfeldübergreifenden Diskussion zu methodischen Standards der Evaluation

*Simone Mazari*<sup>1</sup>

Die 22. Jahrestagung der DeGEval – Gesellschaft für Evaluation e.V. fand vom 11. bis 13. September 2019 an der Rheinischen Friedrich-Wilhelms-Universität in Bonn statt. Der Arbeitskreis (AK) Methoden in der Evaluation, vertreten durch die Sprecherin Prof. Dr. Alexandra Caspari und den Sprecher Dr. Jan Tobias Polak, organisierte in diesem Rahmen im vierten Jahr in Folge den systematischen Austausch über das eingesetzte Methodenrepertoire. Neben der Bestandsaufnahme politikfeldspezifischer Evaluationsmethoden und methodischer Herangehensweisen bestanden Ziele darin, mögliche, politikfeldübergreifende Qualitätsstandards und Entwicklungen im jeweils eingesetzten Methodenrepertoire zu identifizieren. Beteiligungsorientiert, Synergien und Erwartbarkeit erzeugend, wurde das Format des Fishbowl (Innen- und Außenkreismethode) systematisch weitergeführt. Entwickelt wurde der methodenbezogene, strukturierte Austausch über verschiedene Arbeitskreise hinweg erstmalig für die Jahrestagung 2016 in Salzburg. Neben methodischer Cross-Fertilisation (vgl. Ringhofer 2019) über die Arbeitsfelder hinweg bietet das Format einen Raum für Metareflexion, Erfahrungsaufnahme auch aus dem Plenum und die Chance, perspektivisch die Arbeitskreise einer komparativen Gesamtschau zu unterziehen. Die DeGEval besteht aus sechzehn Arbeitskreisen, so dass mit der vorliegenden Session dieser Zyklus abgeschlossen werden und perspektivisch in eine Gesamtschau überführt werden konnte. Erste Überlegungen und Konturen ließen sich im Laufe der Diskussion bereits feststellen.

Für die Arbeitskreise waren im Innenkreis des Fishbowl vertreten:

- Dr. Oliver Schwab (ehemaliger Sprecher AK Strukturpolitik; IfS Institut für Stadtforschung und Strukturpolitik GmbH Berlin)
- Oliver Rohde (Sprecher AK Wirtschaft; DLR Projektträger Projektgruppe „Analysen, Studien, Evaluation“)

---

1 Freiberufliche Sozialwissenschaftlerin und wissenschaftliche Online-Tutorin Fernuniversität in Hagen

- Dr. Bert Droste-Franke (Mitglied im AK Wirtschaft; Institut für qualifizierende Innovationsforschung und -beratung (IQIB))
  - Dr. Franziska Pfitzner-Eden (Sprecherin AK Berufliche Bildung; Phineo gAG)
- Tobias Polak (Sprecher(innen)-Team AK Methoden) moderierte zu vorab angefragten Dimensionen möglicher, methodischer Standards im jeweiligen AK und Beispielen guter Praxis, ob und welche Mindestanforderungen an gute Evaluation im AK formulierbar seien und ob und welche methodische/methodologische Innovationen sich benennen ließen.

## Typische und innovative Evaluationsmethoden – eine Bestandsaufnahme

Oliver Schwab vom AK Strukturpolitik berichtete in seinem Statement von den Erfahrungen und Ausrichtungen von Evaluationen bundesdeutscher Programme im Kontext der EU-Förderlandschaft – insbesondere Europäischer Regionalfonds und Europäischer Sozialfonds geförderter Projekte in der Förderperiode 2014-2020. Durch eine relativ gute Datenausgangslage im Hinblick auf die Anlage und Verfügbarkeit der Evaluationen in Deutschland konnte er die durchgeführten Evaluierungen vergleichsweise eindeutig den prozess- und anwendungsbezogenen Monitoringverfahren („implementation oriented“) oder Wirkungsanalysen („result and impact oriented“) Verfahren zuordnen. Diese Zuordnung erfolgte auf der Grundlage von durch das Evaluation Helpdesk der Generaldirektion Regionalpolitik und Beschäftigung der Europäischen Kommission entwickelten Kategorien und Kriterien, worauf er erläuternd hinwies. Damit konnte eine internationale Vergleichbarkeit hergestellt werden. Die zur Verfügung gestellten Raster entsprachen nicht zwangsläufig den fachlichen Erfahrungen von Akteurinnen und Akteuren im Feld, erlaubten aber dennoch die (internationale) Vergleichbarkeit. Diese Metaperspektive auf die EU-Förderlandschaft und darin durchgeführte Evaluierungen förderten zu Tage, dass nur kleinere Anteile an Mitteln zur Verfügung standen, die für methodische Innovationen verwandt wurden. Auf der Grundlage von 78 ausgewählten Evaluationen aus Regional- und Sozialfonds stellte Oliver Schwab fest, dass sich 72 der 78 Untersuchungen auf die Prozess- und Anwendungsseite beziehen (implementation oriented) und ein Begleitmonitoring durchführen. Wirkungsanalysen (impact oriented) stellten ‚nur‘ 12 von 78 Studien dar. Die verwandten Ansätze ließen sich kaum einem klaren Design zuordnen. Grundlage für die Unterscheidung zwischen theorie- oder kontrafaktischem Design bildeten auch hier sogenannte Guidance-Papiere der Europäischen Kommission. In der Mehrzahl der Studien wurde auf Mixed Methods zurückgegriffen. Ein Schema F wäre demnach auf der Basis der Auswertung im Bereich der Strukturpolitik mit Umsetzungs- und Zielerreichungsverfahren auf der Grundlage von Monitoring gegeben. Vergleichsweise selten eingesetzt wurden eigene Erhebungen oder Studien, Wirkungsanalysen oder kontrafaktische Designs, resümierte Schwab.

Für den AK Wirtschaft berichtete Oliver Rohde: Er konstatierte einen eher weiten Evaluationsbegriff in wirtschaftsbezogenen Kontexten, beispielsweise *über* re-

gelmäßige Maßnahmen, die in Unternehmen durchgeführt werden. Es handele sich mehr um begleitende Evaluation und weniger Ex-post-Messungen. Im Vordergrund stünde die lernende Dimension für das Unternehmen. Themen im AK seien aktuell, wie u.a. CO<sub>2</sub>-Fußabdrücke ‚gemonitort‘ werden könnten oder Qualitätsmanagement von Monitoring unterlegt werde. Ex-ante-Untersuchungen kämen auch eher kaum zum Einsatz wie möglicherweise die agentenbasierte Modellierung vor einer Wirkungsanalyse. Die Verfeinerung von Prognostik oder der Einsatz zukunftsbezogener Methoden waren ein Thema im AK.

Bert Droste-Franke, Mitglied des AK Wirtschaft, beschäftigte sich mit Modellierung im Bereich von Innovationssystemen und erläuterte das methodische Beispiel der agentenbasierten Modellierung. Dabei wurden Unternehmen als Agenten konfiguriert, die zunächst forschen und entsprechende Forschungsprojekte mit beantragen, dann Kooperationen und Netzwerke aufsetzen, über den Einkauf von externer Expertise oder Input-Produkte erstellen. Diese dann vertriebenen Produkte wurden vermarktet und abschließend die angewandte Strategie ex-post untersucht. Dieser Kreislauf sei typisch für ein agentenbasiertes Modell, das sowohl substantiell als auch vom Verfahren her ein innovatives Vorgehen darstelle, so Droste-Franke. Branchen, in denen dieses Vorgehen genutzt wurde, sind beispielsweise sogenannte Schlüsseltechnologien im Energiebereich, für den er die exemplarische Vorgehensweise erläuterte. Wichtig war der Versuch, die dabei ablaufenden Prozesse, so präzise wie möglich abzubilden, sei es graphisch oder im Modell, Netzwerkmaße zu entwickeln, um daraus möglicherweise entstehende Patentnetzwerke idealtypisch abzubilden und systematisch zu begreifen. Bereits die Phase der Forschungsantragserstellung war wichtig für das Gesamtmodell. Auf der Mikro- und ergo Handlungsebene stellten sich die Herausforderung der Rekonstruktion und die Frage der Validität der Darstellung.

Franziska Pfitzner-Eden (Sprecherin AK Berufliche Bildung; Phineo gAG) berichtete einfühend von ihren eigenen Erfahrungen im AK (seit 2010 Mitglied, seit 2016 Sprecherin) und schlussfolgerte aus der Rückschau und dem Vorgespräch mit ihrer Co-Sprecherin, dass im AK Berufliche Bildung ein breites Methodenrepertoire eingesetzt wurde und zwar die gesamte Bandbreite an sozialwissenschaftlichen Methoden. Dies läge am Gegenstand selbst, der sich eher breit über fast den gesamten Lebenslauf erstreckte, kontextualisierte Pfitzner-Eden. Mixed Methods überwogen auch hier. Methoden wurden in Abhängigkeit vom Evaluationsgegenstand sinnhaft miteinander verbunden. Eingesetzt wurden qualitative Fallstudien, wissenschaftliche Begleitforschung, die oft als Evaluation begriffen wurde, aber nicht umstandslos so gesehen werden könne, bemerkte sie kritisch. In den Feldern der Berufsorientierung, der Qualität der Ausbildung, begleitet von beruflichen Schulen und in Betrieben, fanden verstärkt sog. Ausbildungsexporte statt, d.h. Angebote aus der dualen Ausbildung, die in anderen EU-Ländern implementiert werden sollten. Themen seien auch die Evaluation von Ausbildungsordnungen, Diskussionen über die Ausbildungsreife und der Umgang mit unterschiedlichen Zielgruppen wie Migrantinnen und Migranten oder Personen mit Fluchterfahrung. Das liegt daran, dass es eine Vielzahl von Akteurinnen und Akteuren in dem Feld gibt wie berufliche Schulen, Betriebe unterschiedlicher Größen und Branchen, Jugendagenturen und viele mehr, die un-

terschiedliche Erwartungen im Kontext von Evaluationen formulierten – entsprechend stünde beispielsweise eher die Messung von Kompetenzen im Vordergrund oder eher die Wirkung von Maßnahmen. Deshalb gab es durchaus auch Prä- und Ex-post-Designs in der Praxis der Evaluation. Nichtsdestotrotz seien methodische Verfeinerungen der Ex-Post-Messungen durchaus wünschenswert, resümiert sie und verweist auf zu lösende, methodische Probleme wie die Veränderungen des Antwortverhaltens nach der Durchführung einer Maßnahme im Vergleich zur vorgelagerten Befragung (response shift bias). Unterschiede, die sich durch unterschiedliche Befragungszeitpunkte ergäben, würden noch nicht hinreichend auf einer intraindividuellen Ebene analysiert. Die eingesetzten Gruppenmittelwerte wären dann konstant und die darin eingelagerte Streuung könne nicht erfasst werden, womit sie sich für eine stärkere individuelle Betrachtung auf methodischer Ebene einsetzte.

## Politikfeldübergreifende Diskussion zu methodischen Standards

Tobias Polak eröffnete die Diskussion und fragte nach Einschätzungen zur Qualität der eingesetzten Methoden und Erhebungsinstrumente im jeweiligen Handlungsfeld. Oliver Rohde vom AK Wirtschaft verwies auf das eher breite Set an Methoden und stärker prozessbegleitende Evaluationen, weniger Wirkungsforschung, die eingesetzt wurden. Er betonte auch, dass Evaluationen in der Wirtschaft ein besonderes Handlungsfeld darstellten, da der Evaluationsbegriff sehr weit gefasst sei. Bei der Auswertung von Expert(inn)eninterviews könne zudem davon ausgegangen werden, dass diese nicht von einer gründlichen Inhaltsanalyse begleitet sind, so ein Beispiel von Rohde. Franziska Pfitzner-Eden hob als ein Thema der Qualität hervor, dass beispielsweise auch freie Institute aus Ressourcengründen eher keine Transkriptionen von qualitativ erhobenem Material im Standardgeschäft durchführen, die wiederum an Hochschulen durchaus geleistet werden könnten. Damit schloss sie auch an die politikfeldübergreifende Diskussion aus dem letzten Jahr an, bei der das Thema ‚Transkription‘ prominent behandelt wurde (vgl. Ringhofer 2019). Auch Oliver Schwab beobachtete Uneinheitlichkeit in Bezug auf die Qualität der eingesetzten Methoden. Bei einfachen Expert(inn)eninterviews würde beispielsweise nicht systematisch transkribiert und auch keine inhaltsanalytische Auswertung angeschlossen. Er mahnte an, die Methodenqualität künftig verstärkt zu berücksichtigen, so dass auch bei Fallstudien das Sampling oder die Strukturbeschreibung für Außenstehende besser rekonstruiert werden könne und plädierte für mehr Methodenreflexion insgesamt.

Im Laufe der Diskussion und der Öffnung des Fishbowls um den sog. Außenkreis, zeigte sich eine tendenziell als fachlich ‚selbstkritisch‘ zu beschreibende Haltung der Teilnehmerinnen und Teilnehmer. Diese wird im Laufe der Diskussion polemisch zugespitzt auf die Formel ‚Methodenschlampigkeit‘, die in eigener Praxis und im Feld verstärkt in den Blick zu nehmen sei: Eine Teilnehmerin berichtete aus der Arbeitsmarkt- und Berufsbildungsforschung, dass bereits die Auftraggeber nur in geringem Maße an Methoden interessiert seien, was in der Durchführung Me-

thodenschlampigkeit befördern könne, und fragte, wie mehr methodische Komponenten, unter anderem auch gegenderte Ansätze, in die Auftragsforschung Eingang finden könnten. Franziska Pfitzner-Eden pflichtete bei und ergänzte, dass teilweise die Designs bereits in den Leistungsbeschreibungen vorgegeben seien und ob nicht möglicherweise eine Art Vorabschulung von Auftraggeber(inne)n initiiert werden könne, um diese zu mehr methodischer Expertise zu befähigen. Die Notwendigkeit von methodischen Qualitätsdiskussionen wurde auch seitens des AK Strukturpolitik von Oliver Schwab bestätigt. Diese könne bei den Designs und Methodeneinsatz beginnen und wie diese gestaltet sein sollten, wenn nicht umstandslos ein Mehrwert an Erkenntnis gegeben scheint. Das Capacity-Building für die Auftraggeber(innen) könne demnach auch beinhalten, nicht zielführende Erwartungen argumentativ zu verändern. Ansätze der Bearbeitung könnten auch – bei längerfristiger Zusammenarbeit mit Auftraggeberinnen und Auftraggebern – darin bestehen, gemeinsam über die Anlage der Arbeit zu diskutieren. In diesem Zusammenhang ergäbe sich dann vermutlich wiederum ein Problem der Unabhängigkeit der Forschung. Dieses Dilemma stoße allerdings auf ein gestiegenes Interesse bei Auftraggeber(inne)n, Leistungsbeschreibungen detaillierter zu gestalten.

Alexandra Caspari hakte nach, wie es um Metaevaluations bestellt sei, in deren Rahmen insbesondere die Qualität der einer Evaluation zugrunde liegenden Methoden bewertet wird, und ob diese in den jeweiligen Handlungsfeldern durchgeführt würden. Eine Teilnehmerin berichtete, dass Metaevaluations aufgrund der eigenen Positionierung im Feld kollegial schwierig sein könnten. Außerdem würden die zur Verfügung stehenden Texte mehr oder weniger intensiv methodische Fragen behandeln. Sie regte an, im Sinne eines Peer-Review-Verfahrens, fachlich einschlägige deutschsprachige Wissenschaftler(innen) aus den Nachbarländern für Metaevaluations zu bemühen. Oliver Schwab sah über das Evaluation Helpdesk der Generaldirektion Ansätze für Metaevaluations im Politikfeld Strukturpolitik gegeben.

Während im AK Wirtschaft keine Metaevaluations durchgeführt wurden, griff Franziska Pfitzner-Eden ihr Anliegen erneut auf, für eine methodische Qualitätserhöhung Auftragsevaluation zu reduzieren und die jeweiligen Verwaltungen/Ministerien selbst fachlich zu befähigen, methodisch kompetent zu evaluieren. Als Beispiel nannte sie dafür Neuseeland: Evaluator(inn)en seien hier fest in den Fachabteilungen aller Ministerien verankert.

Eine Teilnehmerin unterstützte Oliver Rohde und bestätigte, dass die Sensibilisierung für die Notwendigkeit von Evaluationen in der Wirtschaft bereits schwierig sei. In diesem Feld würden in der Praxis auch Unternehmensberatungen in beträchtlichem Umfang Angebote unterbreiten, die zunächst ähnliche Methoden wie Befragungen einsetzen, aber nicht als Evaluation gewertet werden konnten.

Tobias Polak erkundigte sich nach den Erfahrungen mit kontrafaktischen und theoriebasierten Designs und deren Qualität. Wurden beispielsweise randomisierte Kontrollstudien eingesetzt (fully fledged RCTs) und wie wurde mit Kausalität umgegangen? Oliver Schwab berichtete von den zwei Beispielen für kontrafaktische Designs aus dem Handlungsfeld Strukturpolitik und den Hürden. Randomisierte Verfahren könnten aus rechtlichen Gründen in Deutschland nicht sehr einfach kon-

struiert werden. Das Thema der Kausalität würde im Wesentlichen in theoriebasierten Untersuchungsdesigns angegangen. Dazu verwies er auch auf die Guidance-Papiere der Europäischen Kommission. Diese methodische Entwicklung würde demnach von außen und eher top-down angestoßen. Er vermutete eine erhebliche Bandbreite in der Qualität der durchgeführten Untersuchungen und schlussfolgerte, dass eine theoriebasierte Evaluierung ein weites, künftiges Lernfeld für die Evaluation darstelle. Tobias Polak hakte nach, ob und welche Faktoren im jeweiligen Handlungsfeld für eine gute theoriebasierte Evaluierung zu identifizieren seien. Oliver Schwab verwies auf eine geringe Anzahl an Evaluator(inn)en, die das Feld überhaupt in einem theoriebasierten Zugang überblickten. Franziska Pfitzner-Eden ergänzte für die Berufliche Bildung, dass tendenziell größere Programme eher höhere methodische Ansprüche erfüllen könnten als kleinere Auftragsarbeiten in eng gestricktem Kostenrahmen. Oliver Rohde sah den theoriebasierten Zugang für den AK Wirtschaft tendenziell auch eher als wenig relevant an für die Praxis.

Ein Teilnehmer fragte beim AK Strukturpolitik nach dem Stellenwert der Unterscheidung zwischen theoriebasierter und kontrafaktischer Anlage einer Evaluation, da ihm die Abgrenzung aus der Praxis eher weniger geläufig schien. Er erkundigte sich, inwiefern die agentenbasierte Modellierung sozialpsychologisch fundiert sei und nach interdisziplinärer Modellentwicklung und ob es Empfehlungen gäbe. Da das Modell noch im Prozess der Durchführung stecke, seien noch keine Aussagen zur Nutzbarkeit möglich, so die Replik. In der Vergangenheit gebaute Modelle hätten auch auf der Mikro- und Verhaltensebene Aussagen darüber erlaubt, welches Wissen beispielsweise in welchen Formaten ausgetauscht wurde und wie dies nutzbar wurde für Unternehmen und in die Gestaltung neuer Produkte überführt werden konnte. Eine Validierung fand vor allem auf der Makroebene statt. Das heißt – ausgehend von der Makroebene und vergangener Handlung – wurden Annahmen über das Verhalten auf der Mikroebene getroffen. Auch Oliver Schwab sah keinen Widerspruch zwischen theoriebasierter und kontrafaktischer Herangehensweise. Er begründete dies mit dem Umstand, dass in jedem kontrafaktischen Design ein Wirkungsmodell enthalten sei, das in eine Evaluation eingelassen werden könne. Ebenso könnten innerhalb eines Theoriemodells für Teilbereiche kontrafaktische Analysen durchgeführt werden, so Schwab. Ursprünge für die Gegenüberstellung von theoriebasierter versus kontrafaktischer Analyse lägen in der Interpretation eines Guidance-Papieres im EU-Förderkontext, das dichotomisierend gelesen werden könne, auch wenn dieses Nebeneinander inhaltlich nicht gerechtfertigt sei.

Im Rahmen einer abschließenden Themenrunde fragte Tobias Polak nach methodischen Innovationen im jeweiligen Handlungsfeld. Im AK Wirtschaft sah Oliver Rohde neben bereits genutzten Methoden, die kontextbezogene, agentenbasierte Modellierung als neueren Ansatz. Franziska Pfitzner-Eden identifizierte mögliche Entwicklungspfade weg von Modellversuchen hin zu experimentellen Designs. Sie überlegte, ob nicht auch Megatrends wie die Digitalisierung zu methodischen Innovationen führen könnten (Beispiel: virtuell vermittelte Augmented Reality-Experimente). Über den Einsatz neuer Technologien seien auch Verfeinerungen oder Veränderungen von Evaluationsmethoden gefragt. Oliver Schwab wies nochmals auf die enge Rahmensetzung hin, die durch Regional- und Sozialfonds in methodischer

Hinsicht gegeben sei. Er sieht eher Multiinstrumentenevaluierung als einzelne Innovationen, die seiner Meinung nach nicht als Trend gewertet werden könnten. Eine Teilnehmerin ergänzte für die ländliche Entwicklung, dass ein institutionell angebundenes Setting für Evaluatord(inn)en begünstigend wirkte. Dann seien Modellentwicklungen wie agentenbasierte oder umweltbezogene Verfahren eher aufzusetzen, wie sie beschrieb. Sie beobachtete eine Tendenz bei Auftraggeber(inne)n, Methoden dann verstärkt zu diskutieren, wenn auch Ergebnisse zuvor kontrovers aufgenommen wurden.

Alexandra Caspari leitete zu den Schlussstatements über, in denen die Qualität von Methoden der Evaluation in den Handlungsfeldern, Innovationen und Perspektiven resümiert wurden.

## Ansätze, Qualität und Methoden der Evaluation in den Handlungsfeldern im Ausblick

Wie Heinze und Reiter (2018: 195) im Rückblick auf die Diskussionen der Frage nach typischen und innovativen Methoden in den Arbeitskreisen proklamieren, muss die Frage „immer auch im Kontext dieser Subbereiche gestellt und diskutiert werden.“ Für den AK Wirtschaft hielt Oliver Rohde fest, dass in einem weiten, eher unbestimmten Feld eher nicht institutionalisierte oder kanonisierte Methoden eingesetzt würden. Trotz neuerer Ansätze sei hier Kontinuität weiterhin gegeben. Oliver Schwab entwarf eine Vision, in der – in zu bestimmenden Formen – stärkere Diskussionen um Methodenqualität mit Auftraggeber(inne)n oder auch vorgelagert intern im AK Strukturpolitik stattfinden könnten. Ein Fokus läge dann auf präziseren Designs, um zielgerichteter zu arbeiten. Franziska Pfitzner-Eden schlug für den AK Berufliche Bildung vor, die breite Qualitätsdiskussion in einem weiten Feld perspektivisch an die Entwicklung von Mindeststandards an methodische Qualität zu knüpfen.

Tobias Polak zeigte sich als Moderator sehr angetan von der differenzierten Diskussion und hielt eine Kernthese fest für eine Qualitätsdiskussion: Die AkteursEbene, in diesem Fall Auftragnehmer(innen) und Auftraggeber(innen), seien für eine solche Diskussion mit zu berücksichtigen, da das Set-Up bereits als strukturgebendes Moment Einfluss auf die Qualität von Evaluation nehmen könne. Er dankte den Diskutantinnen und Diskutanten und dem Plenum und verwies abschließend noch auf die Gegenstandsabhängigkeit der Methodenqualität, sog. Wicked Problems sowie noch ausbaubare, komplexere Verfahren, wie sie auch im Zuge von technologisch und soziologisch neueren Entwicklungen wie der Digitalisierung entstünden. Als systematischen Ort für eine solche Meta-Diskussion schlug auch er den veranstaltenden Arbeitskreis Methoden in der Evaluation vor, in dem, wie die Veranstaltung selbst zeigte, die Diskussionen systematisch über Jahre fokussiert werden konnten.

Alexandra Caspari wies abschließend darauf hin, dass mit dieser Diskussionsrunde der Zyklus des systematischen Austauschs zwischen einzelnen Handlungsfeldern abgeschlossen sei und perspektivisch erstmalig in eine analytische Gesamt-

schau überführt werden könne. Sie regte an, für einen Austausch über die Qualität der Methoden auch die Standards der Evaluation hinzuzuziehen und verwies auf den Beitrag „Die ‚Standards für Evaluation‘ aus methodischer Perspektive“ im just erschienenen Sammelband von Hense u.a. „Evaluation: Standards in unterschiedlichen Handlungsfeldern. Einheitliche Qualitätsansprüche trotz heterogener Praxis?“ (vgl. Caspari/Erzberger/Kelle/Polak 2019).

Insgesamt zeigten die Diskussionen der letzten Jahre eine große Bereitschaft zur Beteiligung, ein hohes Maß an Selbstreflexion und Ansatzpunkte sowie eine mögliche Verständigung über methodische Mindeststandards, die Integration subjekt- oder akteursbezogener Dimensionen und den Umgang mit anstehenden, komplexen Herausforderungen wie technischem Wandel, gesellschaftlichen Herausforderungen oder umweltbezogenen Fragestellungen.

## Literatur

- Caspari, Alexandra/Erzberger, Christian/Kelle, Udo/Polak, Jan T. (2019): Die ‚Standards für Evaluation‘ aus methodischer Perspektive. In: Hense, Jan U./Böttcher, Wolfgang/Kalman, Michael/Meyer, Wolfgang (Hg): Evaluation: Standards in unterschiedlichen Handlungsfeldern. Einheitliche Qualitätsansprüche trotz heterogener Praxis? Münster/New York: Waxmann, S. 53-71.
- Heinze, Franziska/Reiter, Stefanie (2018): Methodische Standards der Evaluation zwischen Schema F und Innovation – Weiterführung einer politikfeldübergreifenden Diskussion. In: Zeitschrift für Evaluation, 18 (1), S. 188-197.
- Ringhofer, Lisa (2019): Zwischen Schema F und Innovation – Weiterführung einer politikfeldübergreifenden Diskussion zu methodischen Standards der Evaluation. In: Zeitschrift für Evaluation, 19 (2), S. 314-321. <https://doi.org/10.31244/zfe.2019.02.10>