

Mechanismen als Alternative zu kausaler Attribution – Frühjahrstagung 2017 des Arbeitskreises „Methoden in der Evaluation“ der DeGEval – Gesellschaft für Evaluation e.V.

Heike Steckhan¹, Sabine Brüntrup-Seidemann¹, Ida Verspohl¹

Was ist ein Mechanismus und welche Rolle spielt er in Evaluierungen? Diese Frage stand im Mittelpunkt der Frühjahrstagung 2017 des Arbeitskreises „Methoden in der Evaluation“ der DeGEval – Gesellschaft für Evaluation e.V. in Bonn. In den Räumlichkeiten des Wissenschaftszentrums Bonn konnte der Arbeitskreis gemeinsam mit der gastgebenden Organisation, dem Deutschen Evaluierungsinstitut der Entwicklungszusammenarbeit (DEval), 90 Teilnehmende begrüßen, die an der Beantwortung dieser Frage interessiert waren. Dr. Sven Harten, stellvertretender Direktor des DEval, kündigte in seinen einleitenden Worten einen Einblick in die Analyse von Mechanismen als einen Schlüssel zur Öffnung der ‚Black Box‘ der Wirkungsweise von Evaluierungsgegenständen an.

Im Verlauf der Tagung wurde immer wieder deutlich, dass eine Vielzahl von unterschiedlichen Definitionen des Begriffs ‚Mechanismus‘ existiert: James Mahoney hat in einer Sammlung – die keinen Anspruch auf Vollständigkeit erhebt – 24 unterschiedliche Definitionen gesammelt (vgl. Mahoney, 2001: 579f.). Auch die Vorträge der fünf hochkarätigen Referentinnen und Referenten beleuchteten unterschiedliche Sichtweisen auf den Begriff. In der abschließenden Podiumsdiskussion glichen sie ihre Verständnisse miteinander ab und arbeiteten Gemeinsamkeiten und Unterschiede heraus.

Den Auftakt der Beiträge machte Prof. Dr. Frans L. Leeuw von der Maastricht University bzw. vom Research and Documentation Centre (WODC) des Ministry of Security and Justice der Niederlande (Den Haag) mit seinem anspruchsvollen und unterhaltsamen Vortrag „Mechanismen und generative Kausalität: Mechanismen in theoriebasierten Evaluierungen“. Er nahm sein Publikum mit auf eine rasante Reise durch die Zeit mit einem bunten Strauß an Beispielen, die zurückgingen bis zum Einfluss des Konsums von Zitronen auf Skorbut bei Matrosen. Er definierte Mechanismen als „cogs and wheels“, also als ineinandergreifende Zahnräder, die zwischen Input und Output liegen. Sie werden von Interventionen initiiert und führen schließ-

¹ Deutsches Evaluierungsinstitut der Entwicklungszusammenarbeit – DEval, Bonn

lich zur Wirkung. Dabei unterschied er – basierend auf dem Mikro-Makro-Modell der soziologischen Erklärung von Coleman (1990), auch genannt die „Coleman’sche Badewanne“ – Makro-Mikro-Mechanismen (situational mechanisms), Mikro-Mikro-Mechanismen (action-formation mechanisms) und Mikro-Makro-Mechanismen (transformational mechanisms). Zur Rolle von Mechanismen in theoriebasierten Evaluierungen führte er die Begriffe der „Unvollendeten“ und der „fake handbags“ ein. Wirkungslogiken in theoriebasierten Evaluierungen wiesen nach seiner Einschätzung häufig Schwachstellen auf: Einige Wirkungslogiken beziehen sich zwar auf Mechanismen oder nutzen sogar den Begriff, z.B. in ihren Annahmen über Wirkungszusammenhänge, vergessen aber wichtige Aspekte oder Dimensionen des Mechanismus oder lassen sie bewusst außen vor („die Unvollendeten“). Andere Wirkungslogiken stellen die Intervention, das Programm oder die Maßnahme wie einen Mechanismus dar, indem z.B. die Aktivitäten in Boxen abgebildet und durch Linien und Pfeile ergänzt werden. So entsteht der Eindruck, es handele sich um eine Kausalkette oder einen Mechanismus. Es ist aber eigentlich eine rein tautologische Darstellung des Evaluierungsgegenstands („fake handbags“), denn: „Boxes and arrows are not causal explanations“. Leeuw verwies auf den Ansatz der ‚Realist Evaluation‘ als eine Herangehensweise, in der CMO-Beziehungen (context-mechanism-outcome) expliziert werden. Zentraler Gedanke des CMO-Konzepts ist es, dass Mechanismen sowie Effekte bzw. Wirkungen auch von Kontextfaktoren abhängig sind. Um diese zu untersuchen, empfahl er die Auswertung von wissenschaftlichen und evaluationsspezifischen Datenbanken, um bereits vorhandenes Wissen einzubeziehen, sowie die Integration von „little quick experiments“ in das Evaluierungsdesign, um Mechanismen beispielhaft und ohne ausufernden Ressourceneinsatz zu identifizieren.

Der nächste Vortrag „Soziale Mechanismen, Mediation und Phantomvariablen“ von Prof. em. Dr. Peter Schmidt von der Justus-Liebig-Universität Gießen bewegte sich in der statistischen Welt der Hypothesen und Variablen. Er definierte soziale Mechanismen als „Erklärung von Hypothesen durch andere Hypothesen“ z.B. mittels ‚Phantomvariablen‘ in latenten Variablenmodellen. Dabei unterschied er zwischen der ‚Aktionshypothese‘ und der ‚Mediationshypothese‘: der Mediationshypothese zufolge tritt eine Wirkung (indirekt) über Mediation ein, nach der Aktionshypothese kann von direkteren Wirkungen ausgegangen werden. Die Mediationshypothese geht bei Unterstellung von ‚full mediation‘ davon aus, dass Interventionen oder Programme keine direkte Veränderung hinsichtlich Outcomes (z.B. Verhaltensänderung) bewirken, sondern dies ausschließlich indirekt, z.B. über sich verändernde Werte/Haltungen/Einstellungen geschieht. Es wird wiederum davon ausgegangen, dass sich Werte nur verändern, wenn eine Intervention auf Offenheit (‚openness to change‘) trifft. Der Mechanismus ‚openness to change‘ ist hier also eine zugrunde liegende Hypothese, welche erst ermöglicht, dass die darüberliegende Hypothese „die Intervention bewirkt eine Verhaltensänderung“ greift. Auch Schmidt wies in diesem Zusammenhang auf die Bedeutung von Kontextfaktoren und deren Einbezug z.B. durch Multi-Level Regressions hin. Als Analysemethoden empfahl er quasi-experimentelle Designs, aber auch – wie sein Vorredner – die Integration von Experimenten in Evaluationsdesigns sowie Meta-Analysen.

Zudem verwies er auf die Möglichkeiten der Bayes'schen Statistik sowie der Faktorenanalyse, aber auch auf die Vorzüge von kognitiven Interviews, Längsschnittuntersuchungen (z.B. Wachstumsmodellen) und nicht zuletzt auf neue Möglichkeiten der Verhaltensmessung über Smartphones und andere mobile Endgeräte.

Im nächsten Vortrag mit dem Titel „Mechanismen in Evaluierungen: Mechanismen als kausaler Prozess“ beleuchtete Dr. Johannes Schmitt vom Deutschen Evaluierungsinstitut der Entwicklungszusammenarbeit (DEval) Mechanismen aus einem weiteren Blickwinkel. Schmitt bezog sich auf eine prozessuale Definition von Mechanismen als System, das kausale Kräfte von der Intervention zum Outcome transferiert. Er widersprach seinem Vorredner mit der Aussage, Mechanismen seien keine intervenierenden Variablen, sondern ineinandergreifende ‚Teile‘ bestehend aus ‚Einheiten‘ und ‚Aktivitäten‘, die gemeinsam die Kausalkraft übertragen. Sein Vortrag konzentrierte sich auf die Methode des ‚Process Tracing‘ als innovative Möglichkeit und geeignete Analysemethode, die so definierten Mechanismen zu untersuchen. Die Methode ‚Process Tracing‘ kann in folgende fünf Schritte untergliedert werden: Zuerst wird ein hypothetischer Kausalmechanismus entwickelt, der in einem zweiten Schritt operationalisiert wird. Dabei macht das Evaluierungsteam Vorhersagen bezüglich der zu erwarteten Evidenzen. Es gilt also im Vorfeld zu überlegen, welche empirisch zu erfassenden Spuren oder ‚Fingerabdrücke‘ jeder Teil des Kausalmechanismus hinterlassen haben könnte. Gleichzeitig wird die Aussagekraft der Einzelevidenz bewertet, wobei insbesondere auch untersucht wird, ob es plausible alternative Erklärungen gibt. Nach der Datenerhebung wird in einem vierten Schritt die Aussagekraft der Gesamtevidenz geprüft, um dann in einem letzten fünften Schritt die Ergebnisse für die einzelnen Teile des Kausalmechanismus zusammenzuführen und zu bewerten, ob der hypothetisierte Kausalmechanismus im untersuchten Fall funktioniert hat oder nicht. Process Tracing kann sowohl theoriegestützt, als auch theoriebildend oder Outcome-erklärend angewandt werden. Die Methode ist laut Schmitt vor allem zur Beantwortung von Ob- und Wie-Fragen in Einzelfallstudien geeignet. Der Vorteil besteht in der hohen internen Validität. Eine Generalisierung bzw. Übertragung auf andere Kontexte (externe Validität) wird allerdings erst durch die Kombination mit anderen Methoden wie z.B. vergleichenden Fallstudien, QCA oder statistischen Verfahren möglich. Als zentrale Herausforderungen nannte er die Erfordernis eines umfangreichen Kontextwissens, eines guten Datenzugangs bereits in der Vorbereitung und die zeitintensive Vorbereitung von Feldstudien. Außerdem testet die Methode nur hypothetisierte Kausalmechanismen, alternative Erklärungen müssen vorab mitgedacht werden.

Zum Abschluss des ersten Tages zogen sich die Teilnehmenden in Kleingruppen zurück. Diese Gruppen wurden nach der Funktion der Teilnehmenden als Auftraggeber(in), Auftragnehmer(in) und Wissenschaftler(in) eingeteilt. In ihnen wurden die Vorträge diskutiert und Fragen für die Podiumsdiskussion der Referentinnen und Referenten am nächsten Tag herausgearbeitet.

Der zweite Tagungstag begann mit dem Vortrag „Wirkmechanismen auf der Spur ... Theoretische Überlegungen und praktische Anwendungen am Beispiel der Realistic Evaluation“ von Dr. Sigrid Haunberger von der Fachhochschule Nordwestschweiz, Hochschule für Soziale Arbeit. Auch sie, wie Leeuw am Vortag, sah im

Ansatz der Realistic Evaluation (synonym verwendet zur Realist Evaluation) eine geeignete Möglichkeit, Mechanismen zu identifizieren. Die zentrale Frage der Realistic Evaluation „Wie wirkt eine Intervention für wen und unter welchen Bedingungen in welchem Umfang?“ deckte Mechanismen auf. Haunberger definierte Mechanismen als verbindende, kontextabhängige Elemente zwischen Input und Outcome.

Sie stellte eine Studie vor, in der sie der Frage nachging, wie der Ansatz der Realistic Evaluation als Methode der Wirkungsevaluation in der Sozialen Arbeit sinnvoll angewendet werden kann. Dafür wurde zunächst ein Systematic Review von Realistic Evaluations in diesem Bereich im deutschsprachigen Raum durchgeführt. Der Review zeigte zum einen, dass unterschiedliche methodische Herangehensweisen genutzt werden, um Realistic Evaluations durchzuführen. Insbesondere wurde auf qualitative Ansätze und Mixed Methods zurückgegriffen. Zum anderen wurde deutlich, dass die zentralen Begriffe ‚Kontext‘, ‚Mechanismus‘ und ‚Outcome‘ sehr unterschiedlich verstanden werden. Als nächsten Schritt führte Haunberger in der Studie eine Qualitative Comparative Analysis (QCA) einer betrieblichen Sozialberatung durch. Sie sah in der QCA Parallelen zur Realistic Evaluation und schätzte sie als geeignet ein, Wirkmechanismen aufzudecken. In der QCA stieß sie auf Zusammenhänge, die dem in Realistic Evaluations genutzten Muster des CMO widersprachen, wie CO-, MO-, CCO-, MMO-Konfigurationen. In ihrem Fazit beschrieb sie es als Herausforderung, Kontexte, Mechanismen und Outcomes „eindeutig“ zu identifizieren. Zudem werde die Beziehung zwischen Kontext, Mechanismus und Outcome eher interpretativ überprüft als systematisch getestet. Das führe zu dem Problem der Generalisierbarkeit der Ergebnisse. CMO-Beziehungen seien daher „höchstens Theorien mittlerer Reichweite“ und Realistic Evaluation keine exakte wissenschaftliche Methode, sondern eher ein heuristisches Rahmenwerk bzw. eine Methode, Wissenschaft für die Öffentlichkeit zugänglich zu machen.

Der Vortrag des nächsten Referenten, Frank Willemsen, wie Leeuw vom Research and Documentation Centre des Ministry of Security and Justice der Niederlande (Den Haag), vertrat eine ganz andere Sichtweise als seine Vorrednerinnen und -redner. Er begann seinen Vortrag mit den Worten: „We don't need theory. We look at our data and know everything“. Anschließend eröffnete er seinem Publikum einen Blick in eine digitale Welt der Zukunft: Willemsen beschäftigt sich mit Big Data als Quelle zur Untersuchung von Mechanismen. Dabei nutzt er die Datenmengen, die bei Google und Social-Media-Anbietern, wie Twitter oder Facebook, vorliegen und kombiniert sie mit anderen Methoden (z.B. Experteninterviews), um Vorhersagen über zukünftige Entwicklungen wie z.B. Wirkungen von Politikmaßnahmen zu treffen. Dabei wies er explizit darauf hin, dass eine Validierung mit anderen Methoden nötig sei, da Google seine Suchergebnisse manipuliere und nicht nachvollzogen werden könne, wie dies geschehe. Als Beispiele für überprüfte Vorhersagen nannte er unter anderem die Vorhersage des Verlaufs einer Grippeepidemie anhand von häufig genutzten Suchbegriffen bei Google oder der Route von Flüchtlingen anhand der Auswertung von Anfragen an eine arabischsprachige Suchmaschine. Big Data wird von ihm auch in Evaluationen genutzt, z.B. um Reaktionen der Bevölkerung auf geänderte Gesetzeslagen zu analysieren. Durch eine Gegenüberstellung seiner

Ergebnisse mit amtlichen Aufzeichnungen konnte er die Treffsicherheit seiner Analysen veranschaulichen.

In der anschließenden Podiumsdiskussion, an der alle Referentinnen und Referenten teilnahmen, entspann sich ein lebhafter Austausch über Unterschiede und Gemeinsamkeiten zwischen den Verständnissen von Mechanismen. Vor allem Leeuw und Schmidt sahen viele Gemeinsamkeiten in ihren Sichtweisen: Soziale Mechanismen sind ihrer Ansicht nach verhaltenswissenschaftliche Phänomene bzw. intervenierende Variablen, die nicht beobachtbar sind. Johannes Schmitt hingegen sieht Mechanismen aus der Warte des Process Tracing als beobachtbare Teile an, die in der Summe den kausalen Prozesses bilden. Für ihn sind aber Parallelen zur Kontextabhängigkeit von Outcomes und zu Strukturgleichungsmodellen erkennbar, auf die sich Peter Schmidt bezieht. Im Process Tracing würde man allerdings weniger in der Tiefe untersuchen als vielmehr den gesamten Prozess des Mechanismus betrachten. Schlussendlich wurden sich die Referentinnen und Referenten aber über einige Gemeinsamkeiten einig, die Dr. Tobias Polak als Sprecher des Arbeitskreises Methoden in seiner Zusammenfassung zum Abschluss der Frühjahrstagung aufgriff:

- 1) Eine Theorie oder Wirkungslogik sollte auf Mechanismen Bezug nehmen.
- 2) Von Beginn der Evaluierung an sollten „richtige Handtaschen“ und vollendete Wirkungslogiken erstellt werden. Die üblichen genutzten Pfeildiagramme können einen Ausgangspunkt dafür bieten, wenn sie weiter ausdifferenziert und mit schlüssigen Konzepten hinterlegt werden.
- 3) Bereits vorliegendes Wissen aus Evaluierungen und wissenschaftlichen Studien sollte in Theorien/Wirkungslogiken integriert werden.
- 4) Es existieren unterschiedliche Möglichkeiten, Mechanismen empirisch zu untersuchen, z.B. Realist/Realistic Evaluation, experimentelle Designs oder Process Tracing.
- 5) Einen Mechanismus kann man als eine „Erklärung von Zusammenhängen“ definieren. Die Beschäftigung mit Mechanismen trägt dazu bei, das Bewusstsein für die Bedeutung dieser Zusammenhänge zu schärfen.

Mit dieser Kondensierung der komplexen Diskussionen der Tagung erhielten die Teilnehmenden handhabbare ‚take home messages‘, auf die sie in ihrer weiteren Arbeit zurückgreifen können.

Literatur

- Coleman, James S. (1990): *Foundations of Social Theory*. Cambridge, MA: Belknap Press of Harvard University Press.
- Mahoney, James (2001): *Beyond Correlational Analysis: Recent Innovations in Theory and Method*. In: *Sociological Forum*, 16 (3), pp. 575-593.