

von Simon Müller (Deutsches Zentrum für Integrations- und Migrationsforschung)

KAUSALANALYSEN IN DER EVALUATION RICHTIG ANWENDEN

Methodenvielfalt wahrnehmen, gegenstandsangemessen arbeiten und Kapazitäten stärken

Diese Handreichung gewährt Einblicke in den aktuellen Erkenntnisstand zu Kausalanalysen im Kontext von Evaluationen. Damit richtet sie sich vornehmlich an Evaluierende, Praktizierende und Interventionsplanende. Hierzu werden ein Überblick über verschiedene Formen und Anwendungsstrategien von Kausalanalysen gegeben und konkrete Empfehlungen für die Extremismusprävention, Demokratieförderung und politische Bildung in Deutschland formuliert.

1. Methodische Vielfalt stärker wahrnehmen und anwenden

Experimentelle Forschungsdesigns wie die prominent diskutierte *Randomized Controlled Trials* (RCTs) haben sich in komplexen Evaluationssettings als nur bedingt nützlich erwiesen. Sie sind mit hohem Ressourcenaufwand und ethischen Herausforderungen verbunden und weisen sowohl methodische als auch evaluationspraktische Limitationen auf.

Um belastbare Ergebnisse in Kausalanalysen zu generieren, eröffnen sich neben RCTs eine Bandbreite alternativer Methoden. Hier eignen sich insbesondere quasi-experimentelle Ansätze, dezidiert qualitative Ansätze und Mixed-Methods-Designs. Die Methodenwahl hängt dabei jeweils vom spezifischen Kontext und Setting ab und muss immer individuell abgewogen werden.

Insbesondere im Kontext sensibler und komplexer Politikfelder erweisen sich quasi-experimentelle Designs als vielversprechende Alternative. Während diese wie RCTs einer experimentellen Logik und Vergleichen zwischen Behandlungs- und Kontrollgruppe folgen, verzichten sie jedoch auf eine kontrollierte zufällige Zuordnung (Randomisierung). Auf diese Weise können ethische Herausforderungen, die sich aus „Experimenten“ mit Teilnehmenden ergeben, sowie methodische Probleme vermieden und Ressourcen freigesetzt werden. Weil sie in realen Settings situiert sind, können quasi-experimentelle Designs darüber hinaus eine höhere Kontextsensibilität und Generalisierbarkeit erreichen. Quasi-experimentelle Designs bedienen sich verschiedener Methoden, darunter *Propensity Score Matching*, *Difference-in-Differences*, *Regression Discontinuity Designs* oder *Instrumental Variables*. Diese verfügen über unterschiedliche Stärken und Limitationen und eignen sich jeweils für spezifische Anwendungsbereiche und Evaluationssettings.

Der Kausalitätsbegriff erweckt oft Assoziationen an Statistik und quantitative Methoden. Eine Herangehensweise, die bei Kausaluntersuchungen häufig übersehen wird, sind dedizierte qualitative Ansätze. Entgegen der weit verbreiteten Wahrnehmung eignen sich diese ebenfalls, um robuste Ergebnisse in Kausalanalysen zu generieren. In diesem Kontext haben sich stark systematisierte Ansätze zur Überprüfung von Evidenz hervorgetan (z.B. *Process Tracing* und *Qualitative Comparative Analysis* [QCA]). Diese Methoden liefern detaillierte und kontextsensible Einsichten und erlauben die Untersuchung komplexer, nicht-linearer Zusammenhänge.



Abb. 1: Quasi-experimentelle Methoden im Vergleich

Methode	Beschreibung	Stärken	Limitationen	Anwendungen
Propensity Score Matching (PSM)	konstruiert durch Matching künstliche Behandlungs- und Kontrollgruppen mithilfe eines Propensity Scores (Wahrscheinlichkeit, ‚Behandlung‘ zu erhalten)	ermöglicht die Berücksichtigung einer Vielzahl von Covariaten, leicht interpretier- und kommunizierbar	Annahme, dass alle relevanten Variablen in das PSM aufgenommen wurden	z.B. Beobachtungsstudien und Evaluationen mit reichhaltigen Surveydaten
Difference in Differences (DID)	vergleicht Veränderungen eines Outcomes vor und nach einer Intervention zwischen Behandlungs- und Kontrollgruppe und kontrolliert zeitlich variierende Effekte	Kontrolle aller gemeinsamen zeitlich variierenden und zeit-invarianten Faktoren leicht interpretier- und kommunizierbar	Verzerrungen, z.B. wenn Zeittrends zwischen Gruppen nicht parallel verlaufen Einfluss externer Ereignisse	z.B. Evaluation von Maßnahmen über Zeitspannen (Einführung von Gesetzen oder Interventionen) setzt Verfügbarkeit von Daten vor Interventionsbeginn voraus
Regression Discontinuity Design (RDD)	nutzt natürliche Schwellenwerte, um Einheiten knapp über und unter dieser Schwelle zu vergleichen	hohe Interne Validität transparente Darstellung kausaler Zusammenhänge	erhöhte externe Validität nur auf Kosten von interner Validität schätzt spezifische, ggf. nicht übertragbare Effekte	z.B. Analyse großer Projekte / Programme mit Teilnahme-kriterien (z.B. Einkommens- oder Altersgrenzen) Natürliches Experiment durch Schwellenwert notwendig
Instrumental Variables (IV)	statistische Methode zur Untersuchung kausaler Zusammenhänge zwischen Variablen	Flexibilität bei Modellierung komplexer Zusammenhänge effektive Adressierung von Endogenität	geeignete Instrumente sind schwer zu finden statistisch anspruchsvoll; tiefes Verständnis von Kontext und Bias erforderlich	z.B. bei Selektionseffekten für Projektteilnahme z.B. zur Verbesserung multivariater Regressionen / unvollständiger RCT-Randomisierung

Mixed-Methods-Designs kombinieren die Vorteile quantitativer und qualitativer Methoden. Dieser Ansatz ermöglicht umfassende und vielschichtige Erkenntnisse, indem quantitative Generalisierbarkeit und qualitativer Detailreichtum und -tiefe verbunden werden. Darüber hinaus können qualitative Daten auch gezielt Informationen für die Verbesserung von (quasi-)experimentellen Designs bereitstellen. Mixed-Methods-Evaluationen gelten in einer wachsenden Zahl internationaler Institutionen als ‚best-practice‘.

Beispiele für Qualitative Ansätze in der Kausalanalyse

Mit der **Qualitative Comparative Analysis (QCA)** werden Muster und Konfigurationen von Bedingungen identifiziert, die zu bestimmten Ergebnissen führen. Durch den Vergleich verschiedener Fälle und die Untersuchung von Bedingungskombinationen, die mit einem Ergebnis verbunden sind, untersucht QCA kausale Pfade. Die Method ermöglicht ein nuanciertes Verständnis von Kausalität bei komplexen sozialen Phänomenen und ermöglicht multikausale Erklärungen.

Mithilfe von **Process Tracing** werden kausale Mechanismen untersucht, indem schrittweise die Entfaltung von Ereignissen nachverfolgt wird. Die Methode beinhaltet die Rekonstruktion des kausalen Prozesses in Verbindung mit einer detaillierten Untersuchung kontextueller Faktoren und ermöglicht auf diese Weise ein differenziertes Verständnis der Mechanismen hinter einer Ursache-Wirkungs-Beziehung.

2. Methoden gegenstandsangemessen und pragmatisch auswählen

Pragmatismus in der Methodenwahl bedeutet, verfügbare Ressourcen, Zeitrahmen und Datenquellen realistisch einzuschätzen. Ein pragmatisches Vorgehen erkennt die spezifischen Limitationen und Stärken der methodischen Werkzeuge an, die in der Toolbox für Kausalanalysen zur Verfügung stehen. Eine gegenstandsangemessene Methodenwahl erfordert hingegen eine enge Abstimmung mit den spezifischen Merkmalen und Kontexten der zu evaluierenden Intervention. Dies schließt die Berücksichtigung von Besonderheiten in Zielgruppen, kulturellen Kontexten und sozialen Dynamiken ein. Zum Beispiel können sich für Evaluationen mit wenigen Projektteilnehmenden und in hochsensiblen Kontexten (z.B. in der Arbeit mit vulnerablen Zielgruppen) systematisierte qualitative Ansätze wie QCA für die Kausalanalyse gut eignen. Beide Aspekte sollte immer in Kontext des spezifischen Erkenntnisinteresses der Evaluation gedacht werden.

Kein methodischer Ansatz ist dem anderen per se ‚überlegen‘. Vielmehr unterstreichen Erfahrungen aus der Evaluationspraxis, dass in komplexen sozialen Interventionskontexten keine ‚one-size-fits-all‘-Lösungen existieren. Diesbezüglich lenken Debatten um Methodenhierarchisierung von einer pragmatischen und gegenstandsangemessenen Methodenwahl ab.

3. Kapazitäten für Kausal- und Wirkungsanalysen weiter stärken

In der Evaluationspraxis zeigt sich oft ein unvollständiges Verständnis von Kausalität. Häufig werden Ergebnisse quantitativer Untersuchungen als eindeutig präsentiert, obwohl grundlegende Herausforderungen von Kausalana-

PrEval

INFO BOX

Mixed-Methods: Qualitative Analysen verbessern auch (quasi-) experimentelle Ansätze und liefern Einsichten zu:

- Kausalen Mechanismen (Wie führt eine Ursache zu einer bestimmten Wirkung?)
- Effektheterogenität (Unterscheidet sich der Effekt für verschiedene Gruppen?)
- Wie wurde die Maßnahme durch Begünstigte angenommen?
- Was bedeutet die Maßnahme für unterschiedliche Personen? Kommt die Maßnahme bei allen Personen gleich an?
- Wie werden Variablen gemessen und wie werden sie verstanden?

lysen im Untersuchungsdesign nicht ausreichend Berücksichtigung finden. Entsprechend werden Kausalzusammenhänge suggeriert, die in dieser Einfachheit nicht haltbar sind.

Konsequente Untersuchungen von Kausalität erfordern akribische Arbeit und stellen Evaluierende vor anspruchsvolle Aufgaben. Daher ist es wichtig, dass Evaluierende ein möglichst umfassendes Verständnis für die verfügbaren Methoden entwickeln. Speziell in der Extremismusprävention, Demokratieförderung und politischen Bildung wird das Potenzial von Wirkungsanalysen oft nicht ausgeschöpft, was für die Qualitätssicherung von Projekten allerdings von großer Bedeutung wäre. Auch für Interventionsplanende und politische Entscheidungsträger*innen ist ein grundlegendes Verständnis kausaler Analysen wichtig, um eine ineffektive Ressourcennutzung oder Fehlschlüsse aus Ergebnissen zu vermeiden.

Zum Ausbau dieser Fähigkeiten sollten Räume für Wissensaustausch und methodologische Weiterentwicklung geschaffen werden. Durch Kooperationen und den Aufbau von Netzwerken aus Forschenden, Evaluierenden und Praktizierenden können vorhandenes Wissen und Kapazitäten gesichert und weitergetragen werden. Dafür müssen entsprechende Austauschformate und Plattformen bereitgestellt werden. Auch Schulungsmaterialien und Online-Ressourcen spielen eine wichtige Rolle für den Wissensaustausch und die Sicherung und Verbreitung von Expertise. So könnte beispielsweise ein fachspezifischer Helpdesk Praktizierende und Interventionsplanende angemessen beraten und Unterstützungsformate zur Verfügung stellen.

4. Wirkungsfragen von Beginn an in die Projekt- und Programmplanung integrieren

Eine frühzeitige Integration von Wirkungsfragen in die Planung von Projekten und Programmen ist entscheidend, um eine erfolgreiche Wirkungsevaluation zu ermöglichen. Damit einher geht die Notwendigkeit einer frühzeitigen und präzisen Zieldefinition und die Definition geeigneter Indikatoren. Darüber hinaus sollte bereits zu diesem Zeitpunkt eine präzise Formulierung der Interventionslogik erfolgen und geeignete Wirkungsmodelle entwickelt werden. Für Kausaluntersuchungen sollte im Rahmen der Interventionsplanung auch bereits eine methodische Stoßrichtung vorgegeben werden. Ein Beispiel hierfür ist die Notwendigkeit einer sorgfältigen Planung von *Baselines* (Ausgangspunkt als Referenz für Vergleichszwecke), da bestimmte Designs ohne diese nicht realisiert werden können. Diese Entscheidungen in der Planungsphase haben einen maßgeblichen Einfluss auf die Qualität der Evaluation und sind wichtig für eine passgenaue Anwendung von Evaluationsmethoden.

Weiterführende Informationen

Die zugehörige *PrEval Expertise*, Lektüreempfehlungen und Informationen zu den Autor*innen finden Sie unter: <https://preval.hsfk.de/handreichung0324>



V.i.S.d.P.: Franziska Heil (PRIF)

Baseler Straße 27–31, Frankfurt am Main, Deutschland
Telefon: (069)959104-0, E-Mail: preval@prif.org
Internet: www.prif.org

Gefördert durch:



aufgrund eines Beschlusses
des Deutschen Bundestages



Peace Research Institute Frankfurt
Leibniz-Institut für
Friedens- und Konfliktforschung



PrEval – Zukunftswerkstätten

ist ein Forschungs- und Transfervorhaben zu Evaluation in der Extremismusprävention, Demokratieförderung und politischen Bildung in Deutschland. Die vorliegende Handreichung entstand im Rahmen der PrEval Zukunftswerkstatt „Evaluationsforschung“, die vom Deutschen Zentrum für Integrations- und Migrationsforschung (DeZIM) koordiniert wird.

Weitere Informationen zum Projekt und den beteiligten Partnern unter <https://preval.hsfk.de>