

Tobias Wolbring

# Kausalanalytische Anforderungen an die Theoriebildung\*

**Zusammenfassung:** In den Sozialwissenschaften ergeben sich beim empirischen Nachweis kausaler Effekte erhebliche praktische Schwierigkeiten. Diese hängen u.a. damit zusammen, dass Aspekte der Prüfbarkeit bei der Theoriebildung nicht hinreichend beachtet werden. Im vorliegenden Beitrag werden daher kausalanalytische Anforderungen an sozialwissenschaftliche Theorien dargelegt. Neben der Vorstellung weithin geteilter Gütekriterien (z.B. Sparsamkeit, Realitätsnähe, Prognoseerfolg) und neuerer Überlegungen zu Mechanismus-basierten Erklärungen findet eine Auseinandersetzung mit dem Instrumentarium gerichteter azyklischer Graphen statt. Nach einer allgemeinen Darstellung dieses Ansatzes werden Vor- und Nachteile einer Graphen-basierten Theoriedarstellung anhand der Erklärung des Zusammenhangs zwischen physischer Attraktivität und Lehrveranstaltungsevaluation aufgezeigt. Obwohl nicht alle Probleme mit gerichteten azyklischen Graphen automatisch gelöst werden, wird aufgrund der erhöhten Transparenz und Präzision sowie der besseren Verzahnung von Theorie und Empirie für deren Verwendung bei der Theoriebildung plädiert.

Schlagwörter: Erklärung, Kausalität, gerichtete azyklische Graphen, Mechanismus, Methodologie

## Causal-Analytical Requirements for the Construction of Theories

**Abstract:** In the Social Sciences the empirical detection of causal effects bears considerable practical problems. Amongst others these are related to the fact that aspects of testability are not (sufficiently) accounted for in the construction of theories. Because of this, in this paper we propose causal-analytical requirements for social science theories. Besides the presentation of commonly accepted criteria of quality (e.g., parsimony, resemblance to reality, forecast success) and recent considerations on mechanism-based explanations we deal with the instrument of directed acyclic graphs. After a general presentation of this framework we use the explanation of the relationship between physical attractiveness and students' evaluations of teaching to illustrate the advantages and drawbacks of a graph-based theory presentation. Although directed acyclic graphs do not automatically solve all problems we pledge for their use in the construction of theories due to gains in transparency and precision and a better cogging between theory and empirical results.

Keywords: explanation, causality, directed acyclic graphs, mechanism, methodology

\* Diskussionen mit Norman Braun über wissenschaftstheoretische Fragen und mit Josef Brüderl über das Instrumentarium gerichteter azyklischer Graphen haben zum Verfassen dieses Beitrags ange-regt. Hilfreiche Anregungen, Hinweise und Verbesserungsvorschläge zu früheren Fassungen des Beitrags habe ich von Martina Kroher, den Herausgebern, zwei anonymen GutachterInnen sowie insbesondere von Norman Braun erhalten.

*»To offer a precise and defensible causal effect estimate, a well-specified theory is needed to justify assumptions about underlying causal relationships.«  
Stephan Morgan und Christopher Winship (2007: 30)*

## 1 Einleitung

Der empirische Nachweis von Ursache-Wirkungs-Zusammenhängen ist nicht nur bei Evaluationsstudien, sondern auch bei der Testung wissenschaftlicher Theorien von vordergründigem Interesse. Da Theorien definitionsgemäß empirisch prüfbare Hypothesen beinhalten, die in der Regel auf kausale Zusammenhänge abstellen und teilweise auch Wirkungskanäle einzelner Einflussgrößen explizieren, sollten kausalanalytische Methoden den zentralen Werkzeugkasten darstellen, um theoretische Erklärungen und Prognosen mit empirischem Datenmaterial abzugleichen. Dies gilt selbstverständlich in verstärktem Maße für diejenigen Ansätze, die einem explanatorischen Anspruch verpflichtet sind und damit im Zuge von Erklärungen auf allgemeine Gesetzmäßigkeiten oder empirische Regelmäßigkeiten verweisen (z.B. Hedström 2005; Hedström/Bearman 2009; Maurer/Schmid 2010).

Insbesondere in den Sozialwissenschaften ergeben sich jedoch bei dem Versuch, mittels quantitativ-empirischer Verfahren kausale Effekte aufzuzeigen, erhebliche praktische Schwierigkeiten. Diese reichen von der Unschärfe des theoretischen Begriffsapparats und daraus resultierenden Operationalisierungsproblemen über die mangelnde Explikation von Wirkungsmechanismen und einer damit einhergehenden Offenheit bei der Wahl der Teststrategie bis hin zur fehlenden Spezifikation der Kausalrichtung oder der Ursachen des Zusammenhangs. Entsprechende Schwierigkeiten hängen vermutlich damit zusammen, dass viele Theoretiker nicht um die praktischen Probleme bei der empirischen Prüfung theoretischer Überlegungen wissen<sup>1</sup> und daher bei der Theoriebildung Fragen der Prüfbarkeit kausaler Aussagen nicht hinreichend beachten. Dies erscheint besonders problematisch, wenn man die eingangs zitierte und unter Sozialwissenschaftlern weithin verbreitete Meinung akzeptiert, dass ein rein statistischer Nachweis kausaler Effekte nicht möglich ist und somit theoretischen Überlegungen die wesentliche Rolle bei diesem Vorhaben zukommt.

Ziel des Beitrags ist es daher, aus kausalanalytischer Perspektive Anforderungen an die Theoriebildung und damit Normen akzeptabler sozialwissenschaftlicher Erklärungen zu formulieren. Neben weithin bekannten Standardkriterien »guter« Wissenschaft wie der Falsifizierbarkeit, dem Informationsgehalt, der Bewährung und der Kohärenz von Theorien (z.B. Braun 2008; Esser 1999; Opp 2005a), die nicht nur aus kausalanalytischer Perspektive wünschenswert sind, wird im Folgenden besonderes Augenmerk auf das Instrumentarium gerichteter azyklischer Graphen gelegt. Dabei wird aufgezeigt, dass

1 Eine zentrale Ursache hierfür ist vermutlich die »recht strikte [...] Trennung zwischen Theorie und Empirie im Fach« (Braun 2008: 376), die auch in der universitären Ausbildung und vielen Lehrbüchern üblich ist.

graphische Verfahren einen Beitrag zur Theoriebildung und zur Verbesserung der oft beklagten, unzureichenden Verknüpfung zwischen theoretischen Überlegungen und empirischen Analysen leisten können. Der Beitrag schließt mit einer Illustration anhand des Effekts der physischen Attraktivität der Dozierenden auf deren Bewertung in studentischen Lehrveranstaltungsevaluationen.

## 2 Erklärungen: sparsam und realitätsnah?

Erklärungen sind Antworten auf Warum-Fragen unter Verweis auf allgemeine Regelmäßigkeiten und damit auf die Folgen bestimmter Ursachen und deren Konstellationen. Ein wesentliches Gütekriterium von Erklärungen und damit auch von Mengen von Erklärungen, nämlich Theorien, ist das Prinzip der Sparsamkeit. So ist nach Wilhelm von Ockham auf all diejenigen Annahmen zu verzichten, welche für die Erklärung eines Tatbestands nicht erforderlich sind. Eine Theorie sollte demnach so sparsam wie möglich, aber so komplex wie nötig sein. Diese Forderung ist in der Kausalanalyse besonders wichtig, da – wie später gezeigt wird – jedes zusätzliche erklärende Element unprüfbar Annahmen bei der Testung mit sich bringt. Für den Theoretiker impliziert diese Einsicht die Aufgabe, abstrakte essenzielle Elemente einer komplexen Realität zu isolieren (Prinzip der Isomorphie) und damit vereinfachende und somit unrealistische Prämissen zu treffen. Falsche Annahmen und Verzerrungen der Realität sind folglich ein notwendiges und unvermeidbares Übel bei der Theoriebildung (z.B. Braun 2008). Theorien sind deshalb immer nur Stückwerk und in diesem Sinne, gerade wenn ein ambitionierter realistischer Erklärungsanspruch verfolgt wird, stets falsch.

Milton Friedman (1953) hat diese Einsicht besonders pointiert präsentiert. In seinem berühmten Aufsatz »The Methodology of Positive Economics« macht er darauf aufmerksam, dass bei der Theoriebildung stets ein Zielkonflikt zwischen der Komplexität und Realitätsnähe der Annahmen auf der einen Seite, sowie der Generierung empirisch gehaltvoller Hypothesen und der analytischen Handhabbarkeit/Sparsamkeit der Modelle (Ockhams Razor) auf der anderen Seite besteht. Jeder Theoretiker ist in diesem Sinne zu einem gewissen Grade Instrumentalist, wenn ihm an der Formulierung eines empirisch prüfbar Ansatzes gelegen ist. Friedman folgert hieraus, dass die Güte einer Theorie nicht anhand der Realitätsnähe ihrer Annahmen, sondern nur am Prognoseerfolg bewertet werden sollte. Die These der Irrelevanz des deskriptiven Gehalts der Annahmen hat bekanntlich erheblichen Widerspruch ausgelöst, jedoch konnte Friedmans Position logisch nicht widerlegt werden (vgl. Blaug 1992; Boland 1979; 2008; Mäki 2009a; 2009b).

In der an Friedmans Beitrag anschließenden Diskussion betont jedoch Nagel (1963), dass hinsichtlich der getroffenen Prämissen die Unterscheidung wichtig ist, ob eine Theorie unrealistisch ist, da sie auf unvollständigen oder da sie auf falschen Annahmen beruht. Während unvollständige Annahmen unvermeidbar und sogar fruchtbar sind (siehe auch Kanazawa 1998), ist es zu unterlassen, von der Wahrheit drastisch abweichende, d.h. in diesem Sinne falsche Prämissen und Wirkmechanismen zu spezifizieren. Insbe-

sondere sollte man nicht dauerhaft an bekanntermaßen widerlegten Prämissen festhalten (siehe u.a. Brüderl 2004; Hausman 1992; Hedström 2005; Opp 2005a), da auch aus falschen Annahmen wahre Konklusionen gefolgert werden können. Selbst ein solide durchgeführter und den Ansprüchen der Kausalanalyse genügender empirischer Test theoretischer Implikationen gibt daher streng genommen keinen Aufschluss über die Gültigkeit der einzelnen Bestandteile einer Theorie. Es ist daher (insbesondere bei Vorliegen konkurrierender Erklärungen für dieselben Phänomene) sinnvoll, Theorien auf so viele verschiedene Arten wie möglich zu testen, was insbesondere auch die zentralen Annahmen einschließt.<sup>2</sup>

Bei dadurch motivierten Theoriemodifikationen muss jedoch ein Mittelweg zwischen den konkurrierenden Ansprüchen nach einer Annäherung an die Realität und nach einem handhabbaren Prognoseinstrument gefunden werden. Zwingende analytische Argumente dafür, wie viel Gewicht man dem pragmatischen Gesichtspunkt einer Theorie als Werkzeug und dem regulativen Element der Wahrheit bei der Theoriebildung beimessen sollte, gibt es jedoch nicht. Es erscheint daher sinnvoll, Lindenberg's (1991) pragmatischer Empfehlung zu folgen, als Ausgangspunkt für eine Analyse zunächst eine einfache, empirisch gehaltvolle Theorie heranzuziehen und diese schrittweise um komplexere, realitätsnähere Annahmen anzureichern, solange die Verhältnismäßigkeit des Zusatzaufwands gegenüber dem Zugewinn an Realitätsnähe gewährleistet ist. Zudem sind bei einer solchen Kosten-Nutzen-Bewertung die wissenschaftstheoretischen Einsichten zur Theoriegeladenheit empirischer Resultate und zur mangelnden Testbarkeit isolierter Hypothesen zu berücksichtigen.

### 3 Theoriegetränktheit empirischer Resultate und die Grenzen der Testbarkeit<sup>3</sup>

Daten und damit auch empirische Zusammenhänge sind erst im Lichte einer (expliziten oder impliziten) Theorie interpretierbar. Es gibt deshalb keine voraussetzungslose empirische Forschung geschweige denn Nachweise kausaler Effekte. Vielmehr ist es notwendig, stets messtechnische und soziokulturelle Hintergrundannahmen zu treffen. Diese Prämissen werden in fast keiner Theorie expliziert, sind aber in der Praxis für die Bewertung empirischer Resultate unabdingbar. Weil es kein theorieloses empirisches Material gibt, kann es für Abweichungen zwischen Theorie und Daten immer auch Gründe geben,

- 2 Zu beachten ist dabei zum einen, dass nicht allen Annahmen der gleiche Stellenwert im theoretischen Gefüge zukommt (vgl. Blaug 1992; Braun 2008). Mit Musgrave (1981) ist dabei zu fordern, dass insbesondere diejenigen Annahmen geprüft werden, welche für die Ableitung der theoretischen Folgerungen zentral sind. Zum anderen beziehen sich Prämissen jeweils auf einen bestimmten Kontext und greifen teilweise auch ineinander. Sie sollten daher erstens im Hinblick auf den intendierten Anwendungsfall und zweitens nicht losgelöst voneinander getestet werden (z.B. Lehtinen/Kuorikoski 2007).
- 3 Der erste Teil dieses Abschnitts ist stark an Abschnitt 4.2.3 in Braun et al. 2012 angelehnt.

die unabhängig von dem Wahrheitsgehalt der geprüften Hypothesen sind. Bei systematischen Unterschieden zwischen Theorie und Empirie ist somit zunächst zu prüfen, ob diese durch eine mangelnde Datenqualität oder inadäquate Operationalisierungen erklärbar sind.<sup>4</sup> Falls dies nicht zutrifft, sind Modifikationen und Erweiterungen der Theorie in Betracht zu ziehen.

Insbesondere ist hierbei die von Duhem und Quine<sup>5</sup> beschriebene Problematik bei der empirischen Prüfung von Theorien zu beachten. Da für den empirischen Test einer Hypothese *H* immer Zusatzannahmen und Hilfhypothesen *A* benötigt werden, um die Kluft zwischen Theorie und Testsituation zu überbrücken, ist ein isolierter Test einzelner Hypothesen einer Theorie nicht möglich. Vielmehr werden immer *H* und *A* gemeinsam bzw. deren Implikationen geprüft. Widerspricht ein empirischer Befund *B* der Prognose *B*, kann daraus zwar abgeleitet werden, dass mindestens *H* oder *A* falsch ist. Da jedoch Zusatzannahmen und Hilfhypothesen falsch sein könnten, lässt eine Falsifikation nicht zwingend auf die mangelnde Gültigkeit der Hypothese schließen. Bereits Popper (1983: 187) hat diese frühe Einsicht von Duhem akzeptiert und daher betont, dass nur theoretische Systeme, aber nicht einzelne Hypothesen widerlegbar sind.

Aufgrund der Tatsache dieser empirischen Unterbestimmtheit von Theorien ist es für Kritische Rationalisten wünschenswert, möglichst einfache Theorien zu generieren, die gleichzeitig möglichst viel erklären und auch bisher unbeobachtete Phänomene implizieren (z.B. Homans 1969; Jasso 1988; Opp 2005a). Dadurch werden zum einen die Zahl unprüfbarer, aber notwendiger Zusatzannahmen und damit die Zahl nicht-theorieinhärenter Ursachen für eine Falsifikation gering gehalten. Zum anderen wird durch Befolgung dieses Prinzips die Prüfbarkeit des Ansatzes gewahrt. Überdies erscheint es aus kausalanalytischer Sicht vielversprechend, dem Problem der Unterdeterminanten durch Benennung spezifischer Wirkungspfade der jeweiligen Variablen zu begegnen, da dadurch der Informationsgehalt steigt und vielfältige Prüfungen der Hypothesen ermöglicht werden. Insbesondere sollte das Vertrauen in einen kausalen Effekt von *X* auf *Y* steigen, wenn sich nicht nur zeigen lässt, dass beide Faktoren miteinander assoziiert sind, sondern auch dass die erwarteten Mechanismen den Zusammenhang hervorbringen bzw. in erwarteter Weise moderieren. Es ist daher sinnvoll, sich etwas näher mit entsprechenden Überlegungen auseinanderzusetzen.

4 Hilfreich ist bei einer solchen Evaluation natürlich, wenn die Theorie bereits klare Messanweisungen der relevanten Konstrukte enthält.

5 Es sei an dieser Stelle darauf hingewiesen, dass zwischen den Überlegungen beider Wissenschaftstheoretiker entscheidende Unterschiede bestehen, obwohl üblicherweise von der Duhem-Quine-These die Rede ist. Nach Gillies (1993) liegt die Differenz der beiden Positionen hauptsächlich in der Reichweite der beschriebenen Problematik. Während Duhem (1954) seine These im Hinblick auf eine begrenzte Menge wissenschaftlicher Aussagen formulierte, zweifelte Quine (1951) an der grundsätzlichen Sinnhaftigkeit der Unterscheidung zwischen analytischen und synthetischen Sätzen.

## 4 Mechanismus-basierte Erklärungen

Der Vorschlag, Erklärungen spezifische Mechanismen zugrunde zu legen, ist keineswegs neu (siehe z.B. Hedström/Swedberg 1998) und mit Gerring (2010) und Opp (2005b) kann man sogar fragen, ob nicht in jeder theoretischen Erklärung auf bestimmte Wirkungsmechanismen verwiesen wird. Nichtsdestotrotz erfreut sich – angesichts der Kritik an deduktiv-nomologischen und induktiv-statistischen Erklärungen (z.B. Salmon 2006; Woodward 2003) – das Mechanismus-Konzept in jüngster Zeit, v.a. seit den kritischen Anstößen durch Elster (1989), in den Sozialwissenschaften zunehmender Beliebtheit (Bunge 2004; Goldthorpe 2001; Hedström 2005; Maurer/Schmid 2010; Mayntz 2004; Schmid 2006). Mit Hedström (2005: 25) lassen sich mechanismische Erklärungen dabei folgendermaßen charakterisieren:<sup>6</sup>

»Mechanisms can be said to consist of entities (with their properties) and the *activities* that these entities engage in, either by themselves or in concert with other entities. These activities bring about change, and the type of change brought about depends upon the properties of the entities and the way in which they are linked to one another. A social mechanism, as here defined, describes a constellation of entities and activities that are organized such that they regularly bring about a particular type of outcome.«

Vom Theoretiker wird also gefordert, generative Prozesse des interessierenden Tatbestandes offenzulegen, indem er diejenigen Aktivitäten von zwei oder mehr Einheiten angibt, welche Veränderungen ursächlich hervorbringen. In den Sozialwissenschaften sind etwa rationale Imitation, Vakanzketten und sich selbst-erfüllende Prophezeiungen verbreitete Mechanismen (für eine Zusammenschau siehe Hedström 2009). Wie Opp (2005b: 166) herausarbeitet, ist somit die Minimalvoraussetzung für das Vorliegen einer Mechanismus-basierten Erklärung, dass eine Regelmäßigkeit, meist der Zusammenhang zwischen zwei Variablen, durch mindestens einen dritten, intervenierenden Faktor aufgeklärt wird. In der Regel, aber nicht notwendigerweise wird die Analyse hierzu auf eine niedrigere Ebene verlagert und auf tieferliegende Mechanismen verwiesen (vgl. Stinchcombe 1991).<sup>7</sup>

Das Öffnen dieser Black Box führt dabei zu einer tieferen Erklärung eines Phänomens und ist daher wünschenswert. Mechanismus-basierte Erklärungen liefern nämlich in Ergänzung zu Sätzen der Art »X beeinflusst Y« ein Narrativ, über welche Wege X auf Y wirkt. Damit steigt auch der Informationsgehalt einer Theorie. Jedoch kann es zu einem infiniten Regress kommen, wenn versucht wird, für jede Regularität auf einer bestimm-

6 Wie in den Sozialwissenschaften üblich, finden sich auch zu diesem Begriff eine Vielzahl verschiedener Definitionen. Eine gute, kompakte Übersicht zu sieben verschiedenen Definitionsversuchen geben Hedström (2005: 25) sowie Hedström und Bearman (2009: 6), während Gerring (2008) sogar zehn verschiedene Bedeutungen des Begriffs herausarbeitet. Ausführlich werden unterschiedliche philosophische und soziologische Ansätze von Schmid (2006) behandelt.

7 Neben dem Verweis auf tieferliegende Zusammenhänge ist zudem insbesondere an Interaktionseffekte zu denken.

ten Ebene weitere Mechanismen aufzudecken, welche auf einer tieferliegenden Ebene zu verorten sind. So schreibt Blalock (1976: 88f.):

»Die Entdeckung, dass X Y verursacht, ist nicht einfach falsch. Sie wird nur plausibler, wenn man die Zwischenglieder kennt. Theoretisch gibt es offenbar kein Ende des Interpretationsprozesses, denn man wird immer zwischen je zwei kausal verbundenen Variablen eine intervenierende Variable einfügen können, wodurch eine ursprünglich direkte Beziehung in eine indirekte verwandelt wird.«

Jede Mechanismus-basierte Erklärung enthält daher notwendigerweise eine Black Box, die unerklärt bleiben muss. Die Entscheidung, wann eine Erklärung als hinreichend betrachtet wird, hängt dabei von disziplinspezifischen bzw. paradigmatischen Übereinkünften ab. Diese bezeichnet Hedström (2005) in Anlehnung an Miller (1987) als »stopping rules«, da sie angeben, wann die Suche nach Mechanismen zu einem Ende kommen soll. Je nach theoretischer Perspektive, von der aus ein Problem betrachtet wird, bedarf es demnach einer unterschiedlichen Tiefe der jeweiligen Erklärung.

Für die Soziologie empfiehlt Hedström (2005: 28f.), wie viele andere vor ihm (z.B. Boudon 1980; Coleman 1990; Esser 1999), einen methodologischen Individualismus, so dass handlungstheoretische und intentionale Aspekte den Kern mechanistischer Erklärungen in der Soziologie ausmachen. Während also das analytische Primat in der Soziologie auf der Ebene sozialer Systeme liegt, erscheint es fruchtbar, das theoretische Primat auf die Individualebene zu legen (vgl. Lindenberg 1991) und Erklärungen als Mehrebenenerklärungen anzulegen (Coleman 1990). Dies liegt schon alleine deshalb nahe, weil es keine direkten kausalen Verbindungen, sondern nur Korrelationen zwischen zwei Makro-Variablen geben kann (siehe z.B. Opp 2012). So hängen beispielsweise Krieg und Mortalität eng miteinander zusammen, die hierbei relevanten Ereignisse, wie das Auslösen einer Waffe, eine mangelnde Nahrungsaufnahme und die Infektion mit Krankheiten, sind jedoch auf tieferliegenden Ebenen angesiedelt.

Die methodologische Regel der Mikrofundierung impliziert jedoch keineswegs, strukturelle Gegebenheiten gänzlich auszublenden. Im Gegenteil sind alle sozialen Tatsachen, sowohl ihre Struktur als auch ihr Wandel, mit dem Begriffsinstrumentarium von Individuen, ihren Eigenschaften, Handlungen und ihren Beziehungen zueinander zu explizieren (vgl. Hedström/Bearman 2009: 8). Dieser Strukturindividualismus erfordert somit auf Grundlage eines handlungstheoretischen Kerns (Logik der Selektion), einerseits strukturelle Einflüsse auf Individualvariablen herauszuarbeiten (Logik der Situation) und andererseits Aggregationsprozesse individueller Handlungen aufzuzeigen (Logik der Aggregation) (Esser 1999). Verschiedene Mikrotheorien (z.B. Frame-Selektion-Modell, mikrointeraktionistische Ansätze, Theorien begrenzter und vollständiger Rationalität) sind hierbei als handlungstheoretischer Kern denkbar. Im Gefolge Poppers (2000) erscheint es bei der Wahl der Handlungstheorie jedoch besonders fruchtbar, von Akteuren auszugehen, die bemüht sind, situationsangemessen und in diesem Sinne vernünftig zu agieren (Rationalitätsprinzip).<sup>8</sup>

8 Poppers Gedanken zur Situationsanalyse werden von manchen als sein wichtigster soziologischer

Unabhängig davon, ob man dem Postulat nach einem methodologischen Individualismus und Poppers integrativem Spezifikationsvorschlag folgen möchte oder nicht, bringt die theoretische Herausarbeitung von Wirkungskanälen und intervenierenden Faktoren grundsätzlich verschiedene Vorzüge für die Analyse und Interpretation empirischer Daten mit sich. So betont Hedström (2009), dass Mechanismus-basierte Erklärungen – neben einer höheren Präzision und Verständlichkeit sowie einer durch die Explikation weniger zentraler sozialer Mechanismen bewirkten Reduktion der theoretischen Fragmentierung – das Vertrauen in einen Kausalschluss erhöhen, die Identifikation empirisch sonst nicht nachweisbarer Effekte ermöglichen und ein besseres Verständnis dafür liefern, warum Zusammenhänge eine bestimmte Form annehmen. Insbesondere motivieren Mechanismus-basierte Erklärungen theoriegeleitete Analysen und hemmen die sowohl aus explanatorischer, als auch aus kausalanalytischer Perspektive problematischen Tendenzen in Richtung einer »Variablen-Soziologie« (Esser 1996; Goldthorpe 2001).

Der Rekurs auf das Mechanismus-Konzept stellt jedoch nicht die einzige Möglichkeit dar, kausale Effekte analytisch zu fassen. Mit dem von Pearl (2009) entwickelten, allgemeineren Instrumentarium gerichteter azyklischer Graphen lassen sich sowohl diese, als auch andere Herangehensweisen an das kausalanalytische Identifikationsproblem repräsentieren. Die Struktur der kausalen Graphen ist dabei idealerweise theoretisch begründet, sodass die Wahl geeigneter Brückenannahmen, die Isolation essenzieller Elemente aus der Realität und die hinreichende Tiefe der Erklärung nicht automatisch aus einem graphischen Vorgehen folgt. Vielmehr hängt die Haltbarkeit der aus dem Graphen abgeleiteten Identifikationskriterien wesentlich von der Gültigkeit des zugrunde gelegten Kausalmodells ab. Insofern sind Mechanismus-basierte Ansätze und die Graphenmethodologie nicht als konkurrierende, sondern als komplementäre Ansätze zu begreifen. Neben der präziseren und transparenteren Abbildung der zugrunde gelegten Annahmen liegt der Vorteil einer visuellen Darstellung des theoretischen Arguments dabei vor allem darin, dass sie dem Empiriker recht klar vorgibt, welche Forschungsdesigns und Identifikationsstrategien vielversprechend sind. Um eine bessere Verknüpfung von Theorie und Empirie zu gewährleisten, ist es deshalb auch als Theoretiker wichtig, sich näher mit graphentheoretischen Konzepten und den damit einhergehenden Einsichten zu Möglichkeiten und Grenzen von Kausalnachweisen zu befassen.

## 5 Collider, Confounder und gerichtete azyklische Graphen

In seinem Buch »Causality: Models, Reasoning, and Inference« präsentiert Judea Pearl (2009)<sup>9</sup> sein kausalanalytisches Notationssystem, das auf dem Konzept gerichteter azyk-

und sozialwissenschaftlicher Beitrag bezeichnet – so zumindest die Hauptthese eines Beitrags von Hedström et al. (1998), in dem sich die Autoren mit der Bedeutung des Rationalitätsprinzips für die zeitgenössische Soziologie beschäftigen (siehe aber Schmid 1996 für eine Kritik an Poppers Vorschlag).

9 Eine kompaktere Darstellung der wesentlichen Elemente gibt Pearl (2010). Technisch weniger an-



lischer Graphen (directed acyclic graphs, im Weiteren: DAGs) basiert und insbesondere in der Causal-Inference-Community großen Anklang gefunden hat (siehe exemplarisch Elwert 2013; Morgan/Winship 2007; Woodward 2003). Pearls Arbeiten auf diesem Gebiet stellen dabei eine natürliche, aber fundamentale Erweiterung der pfadanalytischen Tradition in den Sozialwissenschaften dar (z.B. Blalock 1964; Duncan 1966; 1975) und sind gut kompatibel mit Lazarsfelds klassischen Arbeiten zur Interpretation statistischer Beziehungen (z.B. Kendall/Lazarsfeld 1950; Lazarsfeld 1955).<sup>10</sup> Das Konzept der DAGs ermöglicht dabei eine intuitive graphentheoretische Abbildung theoretischer Überlegungen, die aufgrund der Präzision des Notationssystems die Ableitung und rigorose Testung von Implikationen erleichtert.

Bevor die Vorzüge einer solchen Graphen-basierten Theoriebildung eingehender illustriert werden, ist es jedoch unabdingbar, einige zentrale Konzepte zu definieren: Graphen bestehen aus Knoten und Kanten, wobei Knoten im Folgenden theoretische Variablen<sup>11</sup> und Kanten deren theoretisch postulierte Zusammenhänge repräsentieren. Ein Graph ist gerichtet, wenn nicht nur angegeben ist, zwischen welchen theoretischen Variablen Zusammenhänge bestehen, sondern auch welche Kausalrichtung vorliegt. Beeinflusst  $X$   $Y$ , so ist  $X$  ein Vorfahre von  $Y$  und  $Y$  ein Nachfahre von  $X$ . Eine Variable kann nicht zugleich Vorfahre und Nachfahre einer anderen Variable (zu einem gegebenen Zeitpunkt) sein. Simultane Kausalitäten sind also ausgeschlossen, sodass kausale Effekte graphisch mittels unidirektionaler Pfeilen (gerichteten Kanten:  $\rightarrow$ ) zwischen den Knoten dargestellt werden.<sup>12</sup> Die zweite Bedingung der Azyklizität setzt voraus, dass kein direkter

spruchsvolle Einführungen sind Elwert (2013), Glymour (2006), Greenland und Pearl (2007) sowie dem Lehrbuch von Morgan und Winship (2007) zu entnehmen.

- 10 In der deutschsprachigen Soziologie haben sich u.a. Hummell und Ziegler (1976) um die Popularisierung dieser Arbeiten verdient gemacht.
- 11 Pearls Ansatz basiert auf einem probabilistischen Kausalitätsverständnis, weshalb streng genommen von Zufallsvariablen die Rede sein müsste. Hierbei ist zwischen beobachteten Zufallsvariablen (dargestellt durch Knoten ●) und unbeobachteten Zufallsvariablen (dargestellt durch offene Knoten ○) zu unterscheiden (vgl. Pearl 2009: 66f.). Diese Unterscheidung ist jedoch vor allem von empirischer Bedeutung und kann für die Theoriebildung vernachlässigt werden. Die in den DAGs dargestellten Kausalzusammenhänge beziehen sich zunächst einmal auf theoretische Konstrukte, denen, wie bereits Opp (1976; siehe auch 2005a) ausgeführt hat, in der Forschungspraxis Beobachtungsbegriffe bzw. Indikatoren zugeordnet werden müssen. Bei der Bewertung von DAGs im Hinblick auf die zu wählende Analysestrategie ist dann zu berücksichtigen, für welche theoretischen Begriffe geeignete Indikatoren zur Verfügung stehen und welche Möglichkeiten zur Identifikation des interessierenden Kausaleffekts praktisch umsetzbar sind.
- 12 Das Konzept gerichteter Graphen schließt somit aus, dass zu einem gegebenen Zeitpunkt (aber nicht im Zeitverlauf) sowohl  $X \rightarrow Y$  als auch  $Y \rightarrow X$  gilt. Jedoch räumt Pearl (2009: 144) die Möglichkeit ein, Zusammenhänge als bidirektionale (gestrichelte) Kanten darzustellen, wenn latente Größen für die Assoziation verantwortlich sind und die Darstellung dieser weiteren Zusammenhänge den Graphen überfrachten würde, ohne nennenswerten Mehrwert für die vorliegende Fragestellung zu erbringen (s.a. Morgan/Winship 2007: 63). Obwohl es sich bei Nutzung solcher Vereinfachungen im strengen Sinne nicht mehr um einen DAG handelt, kann man den Graphen weiterhin so behandeln, als ob dies der Fall wäre (vgl. Pearl 2009: 144).

Pfad besteht, der von einer Variable ausgeht und bei derselben Variable endet. Ausgeschlossen sind damit zirkuläre Zusammenhänge der Form:  $X \rightarrow Y \rightarrow Z \rightarrow X$ .<sup>13</sup>

Stellt man die theoretischen Überlegungen in einem DAG dar, wird anhand der Knoten und Kanten schnell deutlich, welche Variablen als theoretisch relevant erachtet werden und welche kausalen Zusammenhänge zwischen diesen unterstellt werden. Das graphentheoretische Vorgehen korrespondiert dabei mit einer einfachen Form der formalen Theoriebildung. Durch die Darstellung gerichteter Kanten zwischen Variablen wird nämlich eine nicht-parametrische Spezifikation funktionaler Zusammenhänge (z.B.  $Y = f(X)$ ) vorgenommen, wobei offen bleibt, ob die Assoziation positiv oder negativ ist, welche exakte funktionale Form vorliegt und ob der Effekt mit anderen Einflüssen variiert. DAGs sind daher ein erster Schritt in Richtung einer axiomatisch-deduktiven und damit präzisen und formal rigorosen Theoriebildung. Überdies besteht, wie Pearl (2009: Kap. 7) zeigt, eine Äquivalenz der DAG-Notation zum kontrafaktischen Modell der Kausalität (Imbens/Rubin 2008; Holland 1986; Rubin 1974), was eine Überführung von kausalen Modellen in Form von DAGs in etablierte statistische Modelle erheblich erleichtert.

Wie Elwert (2013) hervorhebt, rücken zudem bei einer Repräsentation des theoretischen Arguments in graphentheoretischer Sprache insbesondere Annahmen über fehlende Kausalzusammenhänge und fehlende Assoziationen klarer in den Vordergrund, als dies in verbalen sozialtheoretischen Ausführungen üblich ist. Dies ist besonders wichtig, da falsche Annahmen über das Fehlen von Zusammenhängen in der Regel schwerwiegendere Konsequenzen für die empirische Testung haben als falsche Annahmen über das Vorliegen von Zusammenhängen. Erst die Absenz bestimmter Kausaleffekte und Assoziationen im Graphen ermöglicht nämlich die Identifikation des interessierenden kausalen Effekts (vgl. Elwert 2013). Für den Fall, dass der Kausaleffekt bei einem gegebenen DAG nicht isolierbar ist, kann es daher vielversprechend sein, weitere Variablen (u.a. Variablen, durch die Treatment und Outcome beeinflusst werden) in Betracht zu ziehen.

Überdies hilft die Darstellung in Form von DAGs dabei, unterschiedliche strukturelle Positionen der Drittvariablen  $Z$  mit Blick auf den interessierenden Kausaleffekt  $X \rightarrow Y$  zu identifizieren. Generell können Variablen dabei verschiedene Rollen in DAGs zukommen:<sup>14</sup>

1. *Unabhängige Einflüsse/Interaktionseffekte*: Interessiert nur der kausale Effekte von  $X$ , so kann von der Kontrolle auf unabhängige Einflüsse (d.h. Variablen, die nur mit  $X$  oder nur mit  $Y$  zusammenhängen und den Effekt anderer Variablen nicht moderieren) abgesehen werden. *Abbildung 1(a)* gibt ein Beispiel, in dem Multikausalität herrscht, also  $X$  und  $Z$  unabhängig voneinander wirken. Die Variable  $Z$  mag zwar ein wichtiger Erklärungsfaktor für die Variation von  $Y$  sein, kann jedoch bei der Theorie-

13 Die Forderung nach Azyklichkeit ist dabei nicht zu verwechseln mit der Eigenschaft der Transitivität von Relationen. Azyklichkeit ist ein allgemeineres Prinzip als Transitivität und daher auch weniger restriktiv. Transitive Graphen sind nämlich stets azyklisch, während die umgekehrte Aussage nicht gilt (für weitere Details siehe Sen 1970: 15f., 47).

14 Bei der hier vorgeschlagenen Klassifikation handelt es sich um eine Typologie des Autors, die keinen Anspruch auf Vollständigkeit erhebt, jedoch die wichtigsten Konstellationen abdecken sollte.

bildung ausgespart bleiben, sofern das Ziel ist, den Effekt von  $X$  auf  $Y$  zu identifizieren. Eine Besonderheit von Pearls Notationssystem ist nun, dass auch Interaktionseffekte graphisch wie in *Abbildung 1(a)* dargestellt werden.<sup>15</sup> Entsprechende Wechselwirkungen sind häufig von kausalanalytischem Interesse, da sie ein besseres Verständnis für tieferliegende Prozesse liefern. Interaktionseffekte sind aber in DAGs strukturell nicht von unabhängigen Einflüssen unterscheidbar.<sup>16</sup>

2. *Instrumente*: Neben der Möglichkeit unabhängiger Einflüsse von  $X$  und  $Z$  auf  $Y$  ist auch vorstellbar, dass  $Z$  nur  $X$ , aber nicht  $Y$  beeinflusst. Für den Theoretiker erscheint dieser Fall irrelevant, da  $Z$  außer über  $X$  nicht mit  $Y$  assoziiert ist und somit nicht zur Erklärung der Ausprägung von  $Y$  beiträgt. Für den Empiriker ist aber gerade diese Konstellation interessant, wenn wie in *Abbildung 1(b)* eine direkte Schätzung des kausalen Effekts von  $X$  auf  $Y$  durch andere Einflüsse erschwert wird. In diesem Fall kann  $Z$  als Instrumentalvariable genutzt werden, sodass der Kausaleffekt<sup>17</sup> trotz der Komplikationen bestimmbar ist (siehe z.B. Angrist/Pischke 2009: Kap. 4; Morgan/Winship 2007: Kap. 7). Für den empirischen Nachweis von Ursache-Wirkungs-Zusammenhängen können also aus theoretischer Sicht irrelevant erscheinende Zusammenhänge von praktischer Bedeutung sein.
3. *Confounder*: Diese Einflussgrößen hat man üblicherweise bei der Drittvariablenkontrolle im Auge. Confounder sind Variablen, die, wie in *Abbildung 1(c)*, sowohl die unabhängige als auch die abhängige Variable beeinflussen bzw. in Pearls Terminologie einen Back-Door-Pfad zwischen  $X$  und  $Y$  öffnen (d-connected).<sup>18</sup> Die Nicht-Beachtung solcher Pfade in empirischen Analysen führt zu einer verzerrten Schätzung des interessierenden Kausaleffekts. Um eine solide Testung theoretischer Implikationen

15 VanderWeele und Robins (2007) arbeiten anhand von DAGs vier Arten von Effektmodifikatoren heraus. Mit der allgemeineren Problematik von Effektheterogenität bei Regressionsschätzungen aus einer kausalanalytischen Perspektive setzen sich Elwert und Winship (2010) auseinander. Zudem legt Pearl (z.B. 2010: 131) nahe, bei Vorliegen von Effektheterogenität einen (allgemeinen oder individualspezifischen) »natürlichen« Baselinelevel für  $Z$  festzulegen und den »natürlichen« Effekt von  $X$  auf  $Y$  für diese Randbedingung zu bestimmen.

16 Üblicherweise stellt man in pfadanalytischer Tradition Interaktionseffekte durch einen Pfeil dar, der auf einen anderen Pfeil gerichtet ist. Dies läuft jedoch der Logik von Pearls Notationssystem zuwider.

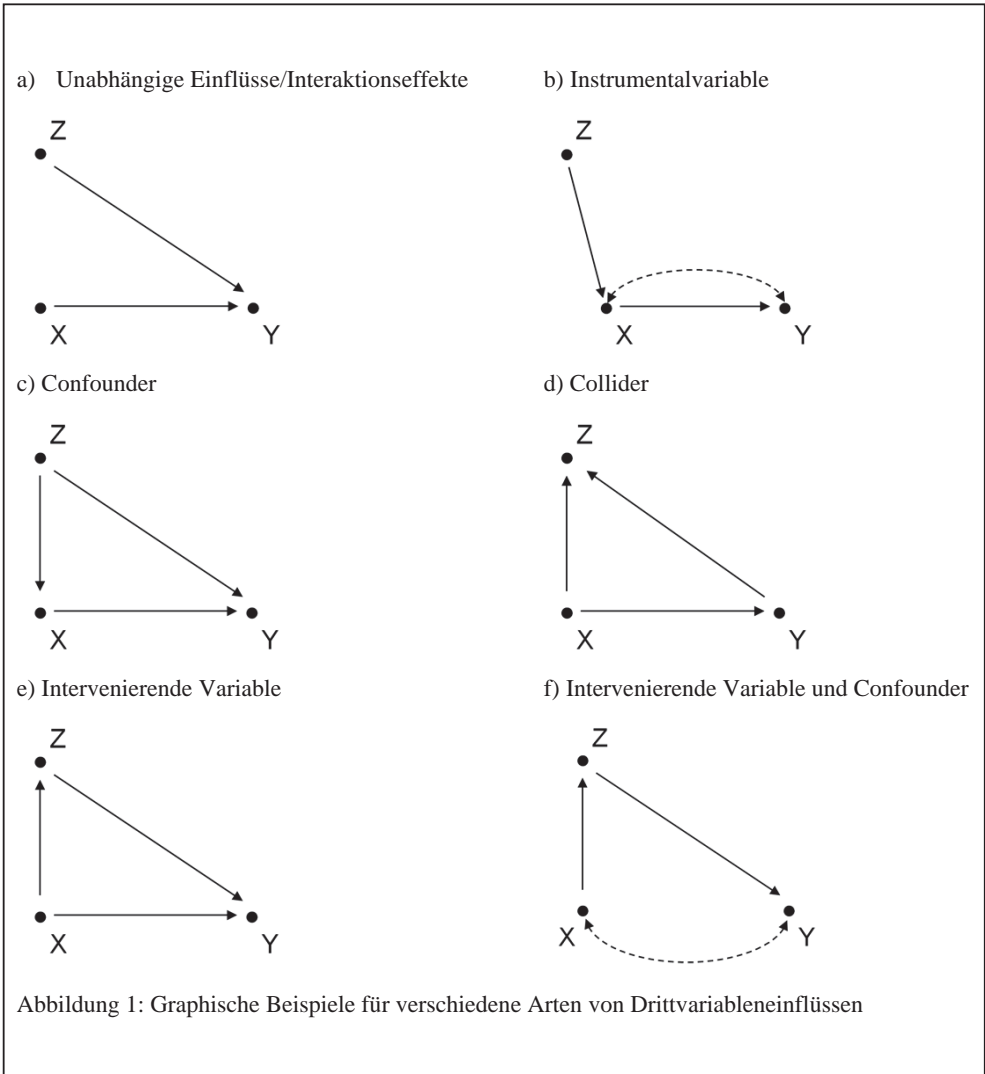
17 Genauer wird mittels Instrumentalvariablenregressionen der local average treatment effect (LATE) geschätzt, also der durchschnittliche kausale Effekt für eine Subgruppe aller Beobachtungen (siehe z.B. Winship/Morgan 2007: Abschnitt 7.4). Dabei handelt es sich um diejenigen Personen, welche sich an die Zuteilung zur Treatment-/Kontrollgruppe gehalten haben (compliers), also in diesem Fall entsprechend der Ausprägung der Instrumentalvariable reagieren.

18 Pearl (2009: Abschnitt 3.3.1) hat das Back-Door-Pfad-Konzept zur Präzisierung der Confounder-Problematik eingeführt. Ein Pfad ist dabei jede Verbindung, die sich zwischen zwei Knoten mittels der Kanten eines Graphs – ungeachtet der Kantenrichtung – herstellen lässt. Ein Back-Door-Pfad zwischen  $X$  und  $Y$  liegt dann vor, wenn ein Pfad zwischen beiden Variablen besteht, bei welchem  $X$  Nachfahre einer dritten Variable  $Z$  ist. Formal lässt sich zeigen, dass es für die Identifikation des kausalen Effekts von  $X$  hinreicht, für eine Teilmenge an Kovariaten  $Z$  zu adjustieren.  $Z$  ist dabei so zu wählen, dass es keine Nachfahren von  $X$  enthält und alle Back-Door-Pfade blockiert (Back-Door-Kriterium).

zu ermöglichen, ist daher die Spezifikation aller potenziellen Back-Door-Pfade notwendig. Geschieht dies nicht, ist eine empirische Überprüfung mittels der üblichen Analyseverfahren (meist Regression oder Matching) streng genommen nicht möglich – schließlich könnte immer auch die Auswahl der Kovariablen unvollständig sein. Da Theorien häufig wenig informativ über mögliche Confounder sind, obliegt die fehleranfällige Entscheidung, welche Variablen bei der Analyse berücksichtigt werden, in der Praxis oft dem empirisch arbeitenden Forscher.

4. *Collider*: Ein Collider liegt immer dann vor, wenn auf einem Pfad zwei Pfeile in Richtung einer Variable deuten. Ein Spezialfall davon ist, dass eine Variable direkt oder indirekt von den beiden Variablen, deren Zusammenhang interessiert, beeinflusst wird (siehe *Abbildung 1(d)*). Ist die Variable  $Z$  ein Collider, so blockiert sie den Back-Door-Pfad  $X \rightarrow Z \leftarrow Y$ . Ein Pfad ist also blockiert (d-separated), wenn entweder eine nicht-kollidierende Variable auf dem Pfad kontrolliert wird oder ein unbeachteter Collider auf dem Pfad liegt. Nimmt man den Collider oder einen seiner Nachfahren jedoch als Kontrollvariable in die Analyse auf, so wird der zuvor blockierte Pfad geöffnet. Eine zentrale Einsicht, die sich bei der Auseinandersetzung mit DAGs ergibt, ist somit, dass die Kontrolle auf bestimmte Faktoren zusätzliche Verzerrungen der Schätzung bewirken kann.<sup>19</sup> Dies impliziert zum einen eine Verabschiedung von der unter quantitativ arbeitenden Sozialwissenschaftlern weit verbreiteten »Kontrolliere alles!«-Strategie. Zum anderen ist es aufgrund dieser Einsicht unzureichend, in der Theorie nur zu benennen, welche Drittvariablen von Bedeutung sein könnten. Vielmehr ist zusätzlich zu spezifizieren, welche Variablen Collider sind.
5. *Intervenierende Variable*: Häufig werden Zusammenhänge zwischen zwei Variablen durch eine dritte Größe vermittelt. Diese Situation ist nach Pearl (2009: 81ff.) durch das Vorliegen eines Front-Door-Pfads gekennzeichnet. In *Abbildung 1(e)* hat  $X$  beispielsweise nicht nur einen direkten kausalen Effekt auf  $Y$ , sondern wirkt sich auch indirekt über  $Z$  auf  $Y$  aus. Bei Vorliegen eines Front-Door-Pfads der Form  $X \rightarrow Z \rightarrow Y$  existiert somit ein Mechanismus, der zur Erklärung der Assoziation zwischen  $X$  und  $Y$  beiträgt. Das Instrumentarium gerichteter azyklischer Graphen ist folglich gut kompatibel mit Mechanismus-basierten Erklärungen. DAGs, die intervenierende Variablen beinhalten, bergen dabei teilweise die Möglichkeit, kausale Zusammenhangsvermutungen selbst dann zu überprüfen, wenn ungemessene konfundierende Einflüsse vorliegen und ein empirischer Nachweis eigentlich aussichtslos erscheint (Morgan/Winship 2007: Kap. 8). In *Abbildung 1(f)* ist ein solcher Fall abgebildet, in dem

19 Zudem lassen sich verschiedene Arten von selektionsbedingten Verzerrungen als Collider-Problem auffassen (siehe Morgan/Winship 2007; Elwert/Winship im Erscheinen). Wird beispielsweise die Teilnahme an einer Befragung sowohl von der Outcome- als auch der Treatment-Variable beeinflusst (z.B. Einkommen und Bildung), so stellt die Variable »Teilnahme an der Befragung« einen Collider dar. Dieser wird (unabsichtlich) kontrolliert, wenn in der Analyse nur diejenigen Personen berücksichtigt werden, die sich an der Befragung beteiligt haben und in der Folge wird der Effekt von Bildung auf Einkommen verzerrt geschätzt.



der kausale Zusammenhang zwischen  $X$  und  $Y$  nur bei Kenntnis der moderierenden Rolle von  $Z$  unverzerrt geschätzt werden kann, indem man die Effekte von  $X$  auf  $Z$  und von  $Z$  auf  $Y$  separat bestimmt. Auch aus kausalanalytischer Sicht sind daher Mechanismus-basierte Erklärungen zu begrüßen.

Die jeweilige strukturelle Position der Einflussgrößen bestimmt also das weitere empirische Vorgehen (für eine Zusammenschau analyseleitender Kriterien siehe Morgan/Winschup 2007 ; Pearl 2009; Elwert 2013). Welche der fünf referierten strukturellen Positionen eine Variable einnimmt, ist dabei immer im Hinblick auf den betrachteten Pfad zu bewerten. Je nach Wirkungspfad kann die Rolle einer Einflussgröße daher variieren. So ist es beispielweise möglich, dass eine Variable auf einem Pfad ein Confounder und auf ei-

nem anderen Pfad ein Collider ist.<sup>20</sup> Weiterhin ist zu beachten, dass sich strukturelle Positionen durch Kontrolle auf andere Einflussgrößen verändern können. Es gibt daher häufig mehr als eine Möglichkeit, alle geöffneten Back-Door-Pfade zu schließen, sodass auch meist mehr als eine Identifikationsstrategie vielversprechend ist. Mittlerweile existiert auch Computer-Software, die diese Schritte erleichtert und (auch für komplexe Graphen mit mehr als hundert Variablen) alle minimalen Teilmengen der zu kontrollierenden Variablen eines DAGs aufzeigt, die zur Identifikation des Kausaleffekts hinreichend sind (minimal sufficient adjustment sets) (s. Knüppel/Stang 2010; Textor et al. 2011).

Will man diese Vorzüge von DAGs nutzen, sind zumindest fünf Aspekte bei der Theoriebildung zu beachten. Erstens ist mit Braun (2008: 385f.) – neben der präzisen Herausarbeitung des im Vordergrund stehenden Kausaleffekts – »die Explikation aller anderen Variablen, die als potenziell einflussreich betrachtet werden und deren Konstanz für die Gültigkeit der Hypothesen der Theorie vorausgesetzt wird«, zu fordern. Zweitens sind diese Angaben um Informationen zu den kausalen Zusammenhängen zwischen diesen weiteren Einflussgrößen und damit zu deren strukturellen Positionen zu ergänzen. Hierbei sind drittens Simultanitäten und Azyklizitäten der unterstellten Kausalstrukturen zu vermeiden. Dies bedeutet nicht, dass komplexe Wechselwirkungen zwischen zwei oder mehr Faktoren ausgeschlossen werden. Um das Argument einer kausalanalytischen Prüfung mit Hilfe von DAGs zugänglich zu machen, muss der Theorie aber die zeitliche Abfolge der Wechselwirkungen klar zu entnehmen sein. Viertens steigt die Wahrscheinlichkeit, den interessierenden kausalen Effekt identifizieren zu können, durch die Benennung von Wirkungspfaden mit intervenierenden Variablen. Durch die Herausarbeitung derartiger Mechanismen bieten sich nämlich zusätzliche Prüfmöglichkeiten. Schließlich erscheint es fünftens aus kausalanalytischer Sicht wünschenswert, sich auf einen oder wenige ausgewählte Kausalzusammenhänge zu konzentrieren und Theorien mittlerer Reichweite im Sinne Mertons (1967) für das jeweilige »problem at hand« anzustreben. Eine solche Komplexitätsreduktion erleichtert die fokussierte Herausarbeitung relevanter Einflüsse und fördert somit eine saubere empirische Testung der Theorie. Dies wird abschließend anhand eines einfachen Beispiels veranschaulicht.

## 6 Illustration: Attraktivitätseffekte bei Lehrveranstaltungsevaluationen

Ausgangspunkt ist der empirische Befund, dass Lehrveranstaltungen attraktiver Dozierender im Durchschnitt bei der Lehrveranstaltungsevaluation (im Folgenden LVE) besser bewertet werden (z.B. Hamermesh/Parker 2005; Klein/Rosar 2006; Wolbring 2010a). Verschiedene Erklärungen wurden für diesen Zusammenhang auf Veranstaltungsebene

20 In einem solchen Fall ergeben sich weitere Erschwernisse bei der Identifikation des Kausaleffekts, da beide strukturellen Positionen unterschiedliche Analysestrategien nahelegen. Trotzdem gibt es zum Teil auch für diese Schwierigkeiten eine Lösung (z.B. Kontrolle auf die Confounder/Collider-Variable und auf eine weitere Variable, die auf dem »Collider-Pfad« liegt).

in der Literatur vorgelegt, wobei zu berücksichtigen ist, dass studentische Lehrveranstaltungsbewertungen das Resultat der Aggregation individueller Handlungen sind. Auf Grundlage evolutionärer Überlegungen, die davon ausgehen, dass physische Attraktivität die Fitness eines potenziellen Partners signalisiert, und unter Berücksichtigung kultureller Einflüsse lassen sich u.a. folgende Mechanismen unterscheiden:<sup>21</sup>

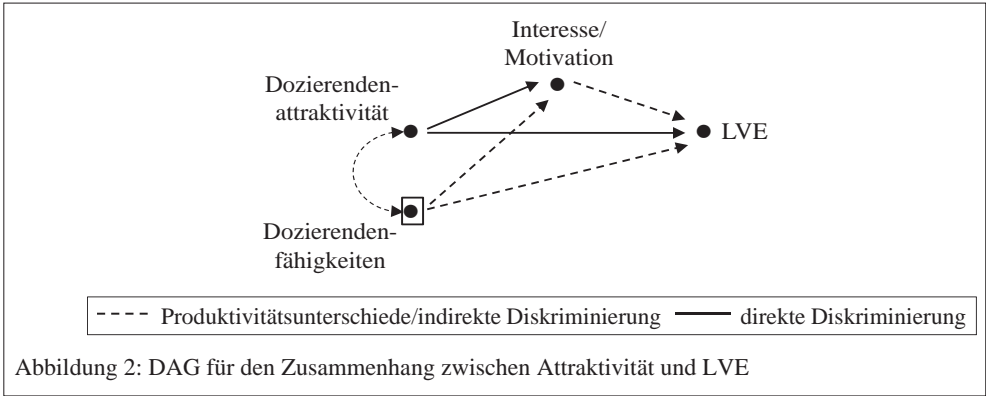
1. *Attraktivitätskonsens*: Bestimmte Merkmale des Partners sind für die Weitergabe und das Überleben der eigenen Gene förderlich. Evolutionär haben sich im Zuge von Variation, Selektion und Vererbung diejenigen Strategien durchgesetzt, die diese Charakteristika nutzen. Insbesondere die physische Attraktivität des Gegenübers ist ein bewährtes Signal für die Evaluation der Fitness. Es besteht daher eine große interpersonelle Übereinstimmung in den Urteilen über die physische Attraktivität einer Person. Zudem bewirken auch kulturelle Einflüsse eine Homogenisierung der Bewertungsstandards. Ein Attraktivitätskonsens ist Voraussetzung für einen Zusammenhang zwischen Attraktivität und LVE. Würde eine starke Heterogenität in den Einschätzungen bestehen, könnten sich Attraktivitätseffekte, obwohl auf Individual-ebene vorhanden, auf Veranstaltungsebene ausgleichen.
2. *Attention Boost*: Eine evolutionäre Perspektive legt überdies nahe, dass attraktive Menschen mehr Aufmerksamkeit erhalten. Für die Partnerwahl ist dies offensichtlich, wobei sowohl potenzielle Partner als auch potenzielle Konkurrenten mit steigender Attraktivität an Bedeutung gewinnen. Menschen nutzen diese tief verankerte und kulturell gestützte Heuristik auch in anderen Kontexten, in denen unvollständige Information über das Gegenüber herrscht. Asymmetrische Information kennzeichnet auch den Lehrkontext, sodass auch attraktive Dozierende eher wahrgenommen werden und mehr Aufmerksamkeit erhalten. Bessere LVE attraktiver Lehrender könnten daher Folge eines gesteigerten studentischen Interesses und Lernerfolges sein. Der Attention Boost sollte sich zudem in einem erhöhten Besuch der Lehrveranstaltung widerspiegeln.
3. *Attribution/Glamoureffekt*: An das äußere Erscheinungsbild werden Stereotype geknüpft, die unabhängig von dem ursprünglichen Kontext der Partnerwahl wirken können. Attraktiven Menschen werden (aufgrund evolutionärer Ursachen und/oder kultureller Einflüsse) zahlreiche positive Eigenschaften wie Freundlichkeit, Verlässlichkeit, Intelligenz, Sozialkompetenz und Gesundheit attribuiert. Es würde daher nicht verwundern, wenn Studierende attraktiven Lehrenden auch eine bessere Lehrbefähigung zuschreiben (Attribution). Da attraktive Menschen in einem positiveren Licht wahrgenommen werden, werden ihnen auch (kleinere) Verfehlungen eher nachgesehen (Glamoureffekt). Fehler können hierbei einerseits bagatellisiert und andererseits auf externe Ursachen zurückgeführt werden (vgl. Rosar 2009).

Diese verschiedenen Mechanismen resultieren in einem generellen Attractiveness Treatment Advantage. Neben diesen verschiedenen Wirkungspfaden sind weiterhin mögliche

21 Ausführlichere Darstellungen dieser stark verkürzten Präsentation der theoretischen Überlegungen finden sich in Rosar (2009) und Wolbring (2013).

Confounder zu beachten. Die Identifikation des kausalen Effekts könnte im vorliegenden Fall insbesondere durch systematische Unterschiede in den Fähigkeiten attraktiver und unattraktiver Dozierender beeinträchtigt werden (vgl. Wolbring 2010b). Zum einen könnte man aus evolutionärer Sicht argumentieren, dass das äußere Erscheinungsbild tatsächlich ein valides Signal ist, das genetisch beeinflusste Fähigkeiten reflektiert, die im Lehrkontext von Bedeutung sind. Zum anderen ist aus soziologischer Sicht an die Möglichkeit indirekter Diskriminierung zu denken: Da attraktive Menschen seit frühester Kindheit von Personen in ihrem Umfeld mehr Aufmerksamkeit und Förderung erfahren und besser behandelt werden, könnten LVE schlicht sozialisationsbedingte Fähigkeitsunterschiede (z.B. Intelligenz, Selbstbewusstsein, Eloquenz) reflektieren. Gleichviel welche Ursache zutrifft, würde eine Nicht-Beachtung dieses Back-Door-Pfades in den statistischen Analysen zu verzerrten Schätzungen des Kausaleffekts der physischen Attraktivität auf die LVE führen.

Der in *Abbildung 2* abgetragene Graph illustriert die beiden Wirkungspfade und das Problem bei der Schätzung des kausalen Attraktivitätseffekts. Die Korrelation zwischen Attraktivität und Fähigkeiten konfundiert den interessierenden Zusammenhang. Zudem lässt sich der indirekte, über das Interesse und die Motivation vermittelte Effekte der Attraktivität nicht bestimmen, da es sich bei der intervenierenden Variable ›Interesse/Motivation‹ auf dem Pfad ›Attraktivität‹ → ›Interesse/Motivation‹ ← ›Fähigkeiten‹ → ›LVE‹ um einen Collider handelt.



Die nahe liegendste Identifikationsstrategie zur Isolation der kausalen Pfade wäre nun, wie in *Abbildung 2* anhand der Box um den Knoten Dozierendenfähigkeiten dargestellt, den Pfad über den Confounder zu blockieren. Dies wird üblicherweise *ex post facto* durch Drittvariablenkontrolle in Regressionen, Matching oder Stratifizierung gelöst. Diese Adjustierungsstrategien sind im vorliegenden Fall jedoch nicht umsetzbar, da keine geeigneten Messungen der Lehrleistung verfügbar sind – ein nahe liegender Indikator stellt schließlich die Gesamtbewertung bei der LVE dar, die jedoch die interessierende Outcome-Variable ist. Alternativ kann man versuchen, ein bezüglich der Fähigkeiten homogenes Subsample zu analysieren. Auch dies ist mit regulären Umfragedaten schwerlich

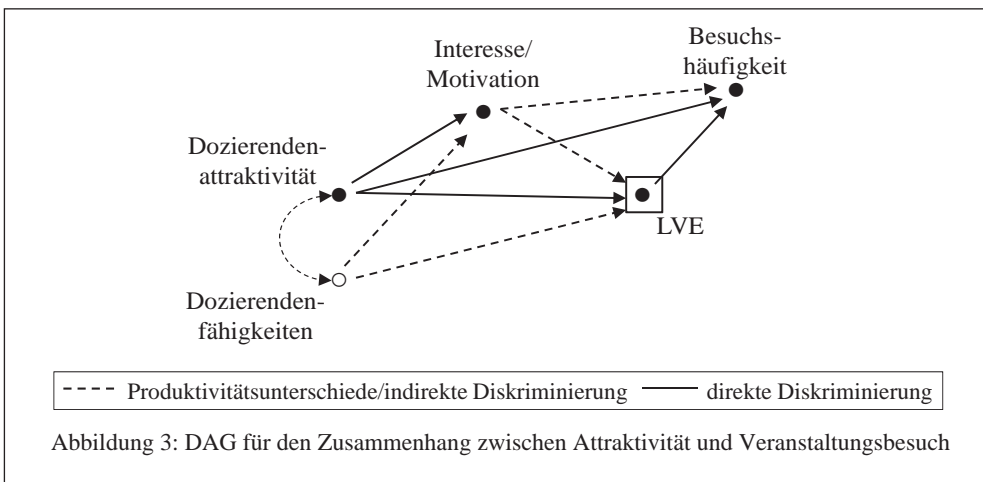


möglich, da dafür wiederum ein geeigneter Fähigkeitsindikator benötigt würde. Vielversprechend ist dagegen eine experimentelle Versuchsanordnung, in der die physische Attraktivität der Lehrenden systematisch variiert wird, während die Lehrleistung unverändert bleibt (für ein Beispiel siehe Wolbring 2013: Kap. 6). Enthält der Datensatz außerdem Indikatoren zu der Motivation und dem Interesse der Studierenden, kann man überdies direkte und indirekte Kausaleffekte voneinander separieren.

Enthalten die Daten jedoch kein valides Maß für die Dozierendenfähigkeiten und bestehen Zweifel an der skizzierten experimentellen Versuchsanordnung (z.B. aus Sorge vor dessen mangelnder externer Validität), ist der kausale Effekt der Attraktivität für die in *Abbildung 2* skizzierte Kausalstruktur nicht identifizierbar. Es kann in solchen Situationen hilfreich sein, weitere Implikationen des theoretischen Arguments zu bedenken und den Graphen um zunächst irrelevant erscheinende Variablen zu ergänzen.

Aus den kursorisch dargestellten theoretischen Überlegungen ergeben sich im vorliegenden Fall weitere kausale Zusammenhänge, die einen Test auf die verzerrende Wirkung von Fähigkeitsunterschieden ermöglichen. Es ist nämlich zu erwarten, dass die physische Attraktivität auch den Veranstaltungsbesuch beeinflusst. Erstens könnte sich eine entsprechende Wirkung vermittelt über ein gesteigertes Interesse einstellen (Attention Boost). Zweitens könnte auch eine (in der studentischen Wahrnehmung) bessere Lehrbefähigung besuchssteigernd wirken (Attribution). Schließlich ist drittens möglich, dass Studierende einen genuinen Nutzen daraus ziehen, mit attraktiven Mitmenschen zu interagieren.

*Abbildung 3* spiegelt diese Überlegungen wider, wobei angenommen wird, dass die Dozierendenfähigkeiten nicht gemessen wurden. Dem Graphen ist zu entnehmen, dass Besuchshäufigkeit ein Collider ist, wenn man den ursächlichen Einfluss der Dozierendenattraktivität auf die LVE isolieren möchte, und dass auch bei der Identifikation des Kausaleffekts der Attraktivität auf die studentische Anwesenheit der konfundierende Einfluss der Dozierendenfähigkeiten zu beseitigen ist. Da aber nun nicht mehr der kausale Zusammenhang der Attraktivität mit der LVE, sondern mit der Besuchshäufigkeit im Vorder-



grund steht, lässt sich auf einen Nachfahren der Dozierendenfähigkeiten, nämlich die LVE, statistisch kontrollieren und sich dadurch der unerwünschte Einfluss teilweise beheben.

Hierbei sind zwei Punkte zu beachten: Zum einen sind für eine solchen Adjustierung in Regressionen und auch beim Matching weitere parametrische Annahmen (z.B. funktionale Form des Zusammenhangs, Berechnung des Propensity Scores) notwendig, die fehleranfällig sind und nicht aus dem nicht-parametrischen DAG folgen. Zweitens wird anhand von *Abbildung 3* deutlich, dass auch mit der Strategie, eine zweite Outcome-Variable zu betrachten, in diesem Falle keine vollständige Isolierung des kausalen Effekts möglich ist. Die Variable ›Dozierendenfähigkeiten‹ kann – auch bei Kontrolle auf die Lehrbewertung – vermittelt über das Interesse und die Motivation der Studierenden die Schätzung konfundieren, wenngleich ein zweiter (und wahrscheinlich der wichtigere) Konfundierungspfad über die subjektive Qualitätswahrnehmung der Studierenden blockiert wird. Trotz dieser Grenzen der Identifizierbarkeit sollte das Vertrauen in die kausale Natur des Attraktivitätseffekts zunehmen, wenn das äußere Erscheinungsbild der Dozierenden mit dem Besuch assoziiert ist, gleichwohl Unterschiede in der studentischen Bewertung der Lehre berücksichtigt werden.

Gleiches gilt für theoretisch abgeleitete Interaktionseffekte. Ihr Nachweis ist nicht hinreichend, um einen kausalen Effekt exakt zu identifizieren. Interaktionseffekten kommt deshalb in DAGs kein eigenständiger Platz zu. Dennoch sollte das Vertrauen in die Existenz des postulierten Ursache-Wirkungs-Zusammenhangs zunehmen, wenn sich Hypothesen über theoretisch erwartete Wechselwirkungen empirisch bewähren. Im Hinblick auf das Anwendungsbeispiel könnte man etwa eine Interaktion zwischen Verhaltenserwartungen und geäußertem Verhalten vermuten. Wenn das Verhalten attraktiver Dozierender nämlich nicht den an sie gerichteten, erhöhten Erwartungen (z.B. freundlich, kooperativ, vertrauenswürdig) entspricht, dann ist die Enttäuschung unter den Studierenden ebenfalls höher und somit auch die Neigung, die Lehrenden abzustrafen. Das Beauty Premium sollte sich demnach bei starken Verfehlungen (z.B. Stellen eines deutlich zu schweren und als unfair wahrgenommenen Tests) in eine Beauty Penalty umkehren.<sup>22</sup> In Pearls Notationssystem werden derartige Interaktionseffekte – analog zu *Abbildung 1(b)* – als zwei separate Pfade ohne graphische Explikation der Wechselwirkung abgetragen (›Attraktivität‹ → ›Erwartungen‹ → ›LVE‹; ›Verhalten‹ → ›LVE‹).

## 7 Fazit

Das hier präsentierte Beispiel ist sicherlich vereinfacht und daher leicht kritisierbar. Zum einen wurden bedeutsame Einflüsse ausgespart. Beispielsweise ist aus evolutionärer Sicht das Geschlecht der Dozierenden und Studierenden für die Wirkung der Attraktivität von

22 Wolbring (2013) hat damit konforme empirische Evidenz vorgelegt: Studierende verhalten sich bei der LVE generell reziprok im Hinblick auf die Testgestaltung der Lehrenden und sanktionieren physisch attraktive Dozierende stärker für Verfehlungen (in Form des Stellens eines zu schwierigen Tests vor der LVE) als weniger attraktive Lehrende.

Bedeutung. Ebenso könnte man darüber diskutieren, ob nicht der studentische Lernerfolg in dem Graphen als eigenständiger Knoten zu berücksichtigen ist. Zum anderen erscheinen die in den DAGs als anwesend und abwesend angenommenen Kausalzusammenhänge diskussionswürdig. So könnte ein Zusammenhang zwischen Dozierendenfähigkeiten und Besuchshäufigkeit bestehen, der nicht über Interesse/Motivation und auch nicht über die wahrgenommene Lehrqualität (gemessen mittels LVE) vermittelt ist (vgl. Wolbring 2013).

Beide Punkte zeigen, dass sich die Diskussion bei Nutzung von DAGs auf die Wahl der Kanten und Knoten zuspitzt. Dies ist der große Vorteil einer graphischen Darstellungsweise: Die theoretischen Annahmen über die Existenz und Abwesenheit von kausalen Einflüssen werden transparent dargestellt und sind daher leichter zu kritisieren. Dies fördert den wissenschaftlichen Diskurs über die Adäquanz der getroffenen Annahmen, weshalb sich die Nutzung von DAGs positiv auf die Erfüllung weithin geteilter Gütekriterien für die Theoriebildung auswirken sollte. Die Wahl geeigneter Brückenannahmen, die Isolation essenzieller Elemente aus der Realität und die hinreichende Tiefe der Erklärung folgen zwar nicht automatisch aus einem graphischen Vorgehen, alle diese wünschenswerten Eigenschaften können aber durch die Explikation der theoretischen Annahmen und der damit einhergehenden leichteren Kritisierbarkeit befördert werden. Zugleich ergibt sich aus der visuellen Darstellung eine starke Vorstrukturierung des empirischen Vorgehens, da präzise dargelegt wird, welche strukturellen Positionen den Drittvariablen zukommen, wie sie in den Analysen zu behandeln sind und ob überhaupt ein Kausaleffekt in dem postulierten Kausalmodell identifizierbar ist.

Jedoch sind mehrere Defizite zuzugestehen. Erstens werden graphische Darstellungen bei Zusammenhängen zwischen einer Vielzahl von Variablen schnell unübersichtlich. In Folge lassen sich auch die theoretischen Implikationen nicht mehr so leicht erfassen. Diese Kritik trifft jedoch für verbale Darstellungen komplexer Zusammenhänge in noch stärkerem Maße zu. Eine formalisierte Form der Theoriebildung erscheint daher eine gute Ergänzung zu dem hier vorgeschlagenen Vorgehen darzustellen. Graphische Relationen lassen sich nämlich kompakt in Matrixform darstellen (siehe z.B. Roberts 1976) und auf dieser Grundlage wurden bereits formale Identifikationskriterien (z.B. Back-Door-Kriterium) und dazugehörige Algorithmen entwickelt, die von Computer-Software zur Analyse von DAGs genutzt werden. Zweitens erscheint es unerfreulich, dass Interaktionseffekten in dem vorgestellten Instrumentarium kein eigenständiger Platz zukommt. Sie sind zwar für die Identifikation von Kausaleffekten unerheblich, die Testung auf theoretisch erwartete Interaktionseffekte gibt aber häufig Aufschluss über die Haltbarkeit der postulierten Mechanismen und erscheint daher wertvoll. Schließlich löst eine Darstellung der Theorie mittels DAGs nicht das altbekannte Aggregationsproblem. Im Graphen werden nämlich nur kausale Einflüsse, aber nicht unbedingt Gleichgewichtsprozesse abgetragen. Gerade wenn konkurrierende Mechanismen vorliegen, bleibt somit offen, welche Einflüsse dominieren und welches Gleichgewicht sich einstellt.

Obwohl also offensichtlich keineswegs alle Probleme bei der Theoriebildung in den Sozialwissenschaften mit gerichteten azyklischen Graphen automatisch gelöst werden, wird für eine graphische Darstellung theoretischer Überlegungen plädiert. Zum einen

schärfen DAGs nämlich das theoretische Argument und sorgen für mehr Transparenz der getroffenen Annahmen. Zum anderen ermöglichen sie eine bessere Verzahnung von Theorie und Empirie, was wechselseitige Lernprozesse fördert.

## Literatur

- Angrist, Joshua/Pischke, Jörn-Steffen (2009): *Mostly Harmless Econometrics. An Empiricist's Companion*. Princeton: Princeton University Press.
- Blalock, Hubert M. (1964): *Causal Inferences in Nonexperimental Research*. Chapel Hill: University of North Carolina Press.
- Blalock, Hubert M. (1976): »Die Beurteilung kausaler Modelle«. In: Hummell, Hans J./Ziegler, Rolf (Hg.): *Korrelation und Kausalität. Band 1*. Stuttgart: Ferdinand Enke Verlag, S. 68-90.
- Blaug, Mark (1992): *The Methodology of Economics. Or How Economists Explain* (2. Aufl.). Cambridge: Cambridge University Press.
- Boland, Lawrence A. (1979): »A Critique of Friedman's Critics«. In: *Journal of Economic Literature* 17(2), S. 503-522.
- Boland, Lawrence A. (2008): »Assumptions Controversy«. In: Durlauf, Steven N./Blume, Lawrence E. (Hg.): *The New Palgrave Dictionary of Economics* (2. Aufl.). Hampshire: Palgrave Macmillan. <http://www.dictionaryofeconomics.com/dictionary> (zuletzt abgerufen am 22.01.2013).
- Boudon, Raymond (1980): *Die Logik gesellschaftlichen Handelns. Eine Einführung in die soziologische Denk- und Arbeitsweise*. Darmstadt/Neuwied: Luchterhand.
- Braun, Norman (2008): »Theorie in der Soziologie«. In: *Soziale Welt* 59(4), S. 373-395.
- Braun, Norman/Keuschnigg, Marc/Wolbring, Tobias (Hg.) (2012): *Wirtschaftssoziologie I: Grundzüge*. München: Oldenbourg.
- Brüderl, Josef (2004): »Die Überprüfung von Rational-Choice-Modellen mit Umfragedaten«. In: Diekmann, Andreas/Voss, Thomas (Hg.): *Rational-Choice-Theorie in den Sozialwissenschaften. Anwendungen und Probleme*. München: Oldenbourg, S. 163-180.
- Bunge, Mario (2004): »How Does It Work? The Search for Explanatory Mechanisms«. In: *Philosophy of the Social Sciences* 34(2), S. 182-210.
- Coleman, James S. (1990): *Foundations of Social Theory*. Cambridge: Belknap Press.
- Duhem, Pierre M. (1954[1906]): *The Aim and Structure of Physical Theory*. Princeton: Princeton University Press.
- Duncan, Otis D. (1966): »Path Analysis: Sociological Examples«. In: *American Journal of Sociology* 72(1), S. 1-16.
- Duncan, Otis D. (1975): *Introduction to Structural Equation Models*. New York: Academic.
- Elster, Jon (1989): *Nuts and Bolts for the Social Sciences*. Cambridge: Cambridge University Press.
- Elwert, Felix (2013): »Graphical Causal Models«. In: Morgan, Stephen L. (Hg.): *Handbook of Causal Analysis for Social Research*. Dordrecht: Springer, S. 245-274.
- Elwert, Felix/Winship, Christopher (2010): »Effect Heterogeneity and Bias in Main-Effects-Only Regression Models«. In: Dechter, Rina/Geffner, Hector/Halpern, Joseph Y. (Hg.): *Heuristics, Probability and Causality: A Tribute to Judea Pearl*. London: College Publications, S. 327-336.
- Elwert, Felix/Winship, Christopher (im Erscheinen): »Endogenous Selection Bias«. In: *Annual Review of Sociology* 40.
- Esser, Hartmut (1996): »What is Wrong with ›Variable Sociology‹?«. In: *European Sociological Review* 12(2), S. 159-166.
- Esser, Hartmut (1999): *Soziologie. Allgemeine Grundlagen* (3. Aufl.). Frankfurt a.M.: Campus.
- Friedman, Milton (1953): »The Methodology of Positive Economics«. In: Friedman, Milton (Hg.): *Essays in Positive Economics*. Chicago: University of Chicago Press, S. 1-43.
- Gerring, John (2008): »The Mechanismic Worldview: Thinking Inside the Box«. In: *British Journal of Political Science* 38(1), S. 161-179.

- Gerring, John (2010): »Causal Mechanisms: Yes, But«. In: *Comparative Political Studies* 43(11), S. 1499-1526.
- Gillies, Donald (1993): *Philosophy of Science in the Twentieth Century. Four Central Themes*. Oxford: Blackwell.
- Glymour, M. Maria (2006): »Using Causal Diagrams to Understand Common Problems in Social Epidemiology«. In: Oakes, J. Michael/Kaufman, Jay S. (Hg.): *Methods in Social Epidemiology*. San Francisco: Jossey-Bass, S. 387-422.
- Goldthorpe, John H. (2001): *On Sociology. Numbers, Narratives, and the Integration of Research and Theory*. Oxford: Oxford University Press.
- Greenland, Sander/Pearl, Judea (2007): »Causal Diagrams«. In: Boslaugh, Sarah (Hg.): *Encyclopedia of Epidemiology*. Thousand Oaks: Sage, S. 149-156.
- Hamermesh, Daniel S. (2011): *Beauty Pays*. Princeton: Princeton University Press.
- Hamermesh, Daniel S./Parker, Amy M. (2005): »Beauty in the Classroom. Instructors' Pulchritude and Putative Pedagogical Productivity«. In: *Economics of Education Review* 24(4), S. 369-376.
- Hausman, Daniel M. (1992): *The Inexact and Separate Science of Economics*. Cambridge: Cambridge University Press.
- Hedström, Peter/Swedberg, Peter/Udéhn, Lars (1998): »Popper's Situational Analysis and Contemporary Sociology«. In: *Philosophy of the Social Science* 28(3), S. 339-364.
- Hedström, Peter (2005): *Dissecting the Social. On the Principles of Analytical Sociology*. Cambridge: Cambridge University Press.
- Hedström, Peter (2009): »Studying Mechanisms to Strengthen Causal Inferences in Quantitative Research«. In: Box-Steffensmeier, Janet M./Brady, Henry E./Collier, David (Hg.): *The Oxford Handbook of Political Methodology*. Oxford: Oxford University Press, S. 319-335.
- Hedström, Peter/Bearman, Peter (2009): »What is Analytical Sociology All About? An Introductory Essay«. In: Dies. (Hg.): *The Oxford Handbook of Analytical Sociology*. New York: Oxford University Press, S. 3-24.
- Hedström, Peter/Swedberg, Richard (Hg.) (1998): *Social Mechanisms. An Analytical Approach to Social Theory*. Cambridge: Cambridge University Press.
- Holland, Paul (1986): »Statistics and Causal Inference«. In: *Journal of the American Statistical Association* 81, S. 945-960.
- Homans, George C. (1969): *Was ist Sozialwissenschaft?* Opladen: Westdeutscher Verlag.
- Hummell, Hans J./Ziegler, Rolf (Hg.) (1976): *Korrelation und Kausalität. 3 Bände*. Stuttgart: Ferdinand Enke Verlag.
- Imbens, Guido W./Rubin, Donald (2008): »Rubin Causal Model«. In: Durlauf, Steven N./Blume, Lawrence E. (Hg.): *The New Palgrave Dictionary of Economics* (2.Aufl.). Hampshire: Palgrave Macmillan. DOI:10.1057/9780230226203.1466.
- Jasso, Guillermina (1988): »Principles of Theoretical Analysis«. In: *Sociological Theory* 6(1), S. 1-20.
- Kanazawa, Satoshi (1998): »In Defense of Unrealistic Assumptions«. In: *Sociological Theory* 16(2), S. 193-204.
- Kendall, Patricia L./Lazarsfeld, Paul F. (1950): »Problems of Survey Analysis«. In: Lazarsfeld, Paul F./Merton, Robert K. (Hg.): *Continuities in Social Research: Studies in the Scope and Method of »The American Soldier«*. New York: Free Press, S. 133-196.
- Klein, Markus/Rosar, Ulrich (2006): »Das Auge hört mit! Der Einfluss der physischen Attraktivität des Lehrpersonals auf die studentische Evaluation von Lehrveranstaltungen – eine empirische Analyse am Beispiel der Wirtschafts- und Sozialwissenschaftlichen Fakultät der Universität zu Köln«. In: *Zeitschrift für Soziologie* 35(4), S. 305-316.
- Knüppel Sven/Stang Andreas (2010): »DAG Program: Identifying Minimal Sufficient Adjustment Sets«. In: *Epidemiology* 21(1), S. 159.
- Lazarsfeld, Paul F. (1955): »Interpretation of Statistical Relations as a Research Operation«. In: Lazarsfeld, Paul F./Rosenberg, Morris (Hg.): *The Language of Social Research*. Glencoe, Ill.: Free Press, S. 115-125.

- Lehtinen, Aki/Kuorikoski, Jaakko (2007): »Unrealistic Assumptions in Rational Choice Theory«. In: *Philosophy of the Social Sciences* 37(2), S. 115-138.
- Lindenberg, Siegwart (1991): »Die Methode der abnehmenden Abstraktion: Theoriegesteuerte Analyse und empirischer Gehalt«. In: Esser, Hartmut/Troitzsch, Klaus (Hg.): *Modellierung sozialer Prozesse*. Bonn: Informationszentrum Sozialwissenschaften, S. 29-78
- Mäki, Uskali (2009a): »Unrealistic Assumptions and Unnecessary Confusions: Rereading and Rewriting F53 as a Realist Statement«. In: Mäki, Uskali (Hg.): *The Methodology of Positive Economics. Reflections on the Milton Friedman Legacy*. Cambridge: Cambridge University Press, S. 90-116.
- Mäki, Uskali (2009b): »Realistic Realism about Unrealistic Models«. In: Kincaid, Harold/Ross, Don (Hg.): *A Handbook of the Philosophy of Economics*. Oxford: Oxford University Press, S. 68-98.
- Maurer, Andrea/Schmid, Michael (2010): *Erklärende Soziologie. Grundlagen, Vertreter und Anwendungsfelder eines soziologischen Forschungsprogramms*. Wiesbaden: VS Verlag.
- Mayntz, Renate (2004): »Mechanisms in the Analysis of Social Macro-Phenomena«. In: *Philosophy of the Social Sciences* 34(2), S. 237-259.
- Merton, Robert K. (1967): *On Theoretical Sociology: Five Essays, Old and New*. New York: Free Press.
- Miller, Richard W. (1987): *Fact and Method: Explanation, Confirmation and Reality in the Natural and the Social Sciences*. Oxford: Princeton University Press.
- Morgan, Stephen L./Winship, Christopher (2007): *Counterfactual and Causal Inference. Methods and Principles for Social Research*. Cambridge: Cambridge University Press.
- Musgrave, Alan (1981): »Unreal Assumptions in Economic Theory: The F-Twist Untwisted«. In: *Kyklos* 34(3), S. 377-387.
- Nagel, Ernst (1963): »Assumptions in Economic Theory«. In: *American Economic Review: Papers and Proceedings* 53(2), S. 211-219.
- Opp, Karl-Dieter (1976): »Theoretische Begriffe, Beobachtungsbegriffe und Kausalanalyse«. In: *Soziale Welt* 27(2), S. 139-143.
- Opp, Karl-Dieter (2005a): *Methodologie der Sozialwissenschaften: Einführung in Probleme ihrer Theoriebildung und praktischen Anwendung* (6. Aufl.). Wiesbaden: VS Verlag.
- Opp, Karl-Dieter (2005b): »Explanations by Mechanisms in the Social Sciences. Problems, Advantages and Alternatives«. In: *Mind & Society* 4(2), S. 163-178.
- Opp, Karl-Dieter (2012): »Can there be Causal Effects on the Macro Level?«. In: *Sociologica* 1/2012, doi 10.2383/36897.
- Pearl, Judea (2009 [2000]): *Causality: Models, Reasoning, and Inference* (2. Aufl.). Cambridge: Cambridge University Press.
- Pearl, Judea (2010): »The Foundations of Causal Inference«. In: *Sociological Methodology* 40(1), S. 75-149.
- Popper, Karl R. (2000 [1967]): »Das Rationalitätsprinzip«. In: Miller, David (Hg.): *Karl R. Popper Lesebuch. Ausgewählte Texte zu Erkenntnistheorie, Philosophie der Naturwissenschaften, Metaphysik, Sozialphilosophie*. Tübingen: Mohr Siebeck, S. 350-359.
- Popper, Karl R. (1983): *Realism and the Aim of Science*. London: Hutchinson.
- Quine, Willard van Orman (1951): »Two Dogmas of Empiricism«. In: *Philosophical Review* 60, S. 20-53.
- Roberts, Fred S. (1976): *Discrete Mathematical Models with Applications to Social, Biological, and Environmental Problems*. Englewood Cliffs: Prentice Hall.
- Rosar, Ulrich (2009): *Physische Attraktivität und soziale Ungleichheit. Ein Forschungsprogramm*. Habilitation: Universität Köln.
- Rubin, Donald B. (1974): »Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies«. In: *Journal of Educational Psychology* 66(5), S. 688-701.
- Salmon, Wesley C. (2006): *Four Decades of Scientific Explanation*. Minneapolis: University of Minnesota Press.
- Schmid, Michael (1996): *Rationalität und Theoriebildung. Studien zu Karl R. Poppers Methodologie der Sozialwissenschaften*. Amsterdam: Rodopi.
- Schmid, Michael (2006): *Die Logik mechanistischer Erklärungen*. Wiesbaden: VS Verlag.

- Sen, Amartya K. (1970): *Collective Choice and Social Welfare*. San Francisco u.a.: Holden-Day.
- Stinchcombe, Arthur (1991): »The Conditions of Fruitfulness of Theorizing About Mechanisms in Social Science«. In: *Philosophy of the Social Sciences* 21(3), S. 367-388.
- Textor, Johannes/Hardt, Juliane/Knüppel, Sven (2011): »Letter to the Editor: DAGitty: A Graphical Tool for Analyzing Causal Diagrams.« In: *Epidemiology* 22(5): S. 745.
- VanderWeele, Tyler J./Robins, James M. (2007): »Four Types of Effect Modification – a Classification Based on Directed Acyclic Graphs«. In: *Epidemiology* 18(5), S. 561-568.
- Wolbring, Tobias (2010a): »Physische Attraktivität, Geschlecht und Lehrveranstaltungsevaluation. Eine Replikationsstudie zu den Befunden von Hamermesh und Parker (2005) und Klein und Rosar (2006) mit Hilfe von Individualdaten«. In: *Zeitschrift für Evaluation* 9(1), S. 29-48.
- Wolbring, Tobias (2010b): »Weshalb die Separierung von Produktivitätseffekten und Diskriminierung bei der studentischen Lehrveranstaltungsbewertung misslingt. Selektive Stichproben, fehlende Drittvariablenkontrolle und die Konfundierung von Effekten«. In: *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 62(2), S. 317-326.
- Wolbring, Tobias (2013): *Fallstricke der Lehrevaluation. Möglichkeiten und Grenzen der Messbarkeit von Lehrqualität*. Frankfurt am Main: Campus.
- Woodward, James (2003): *Making Things Happen. A Theory of Causal Explanation*. Oxford: Oxford University Press.

*Anschrift:*

Dr. Tobias Wolbring  
 Departement für Juristen-, Sozial- und Staatswissenschaften  
 ETH Zürich  
 Mühlegasse 21  
 CH-8001 Zürich  
 tobias.wolbring@gess.ethz.ch