



Process Tracing in der Policy-Forschung

Peter Starke

Inhalt

1 Einleitung: Was ist Process Tracing?	1
2 Ursprünge, Annahmen, Forschungsziele und Fallauswahl	3
3 Datengrundlage: Systematik jenseits der Standardisierung	10
4 Datenauswertung: Spurenlesen als Methode	13
5 Anwendungen in der Policy-Forschung	16
6 Stärken, Schwächen und Gütekriterien von Process Tracing	19
7 Fazit	21
Literatur	22

Zusammenfassung

Process Tracing ist eine Untersuchungsmethode zur kausalen Erklärung, bei der vielfältige empirische Beobachtungen innerhalb eines oder mehrerer Fälle als Implikationen – oder „Spuren“ – theoretischer Kausalmechanismen verstanden werden. Die möglichst vollständige empirische Rekonstruktion kausaler Prozesse durch Fallstudien erlaubt Schlussfolgerungen über (alternative) theoretische Erklärungen. Der Beitrag stellt die wichtigsten Merkmale der Methode, Beispiele aus der Policy-Forschung, Stärken, Schwächen sowie methodische Gütekriterien vor.

Schlüsselwörter

Qualitative Methoden · Forschungsdesign · Fallstudien ·
Untersuchungsmethode · Fallauswahl

P. Starke (✉)
University of Southern Denmark, Odense, Dänemark
E-Mail: starke@sam.sdu.dk

1 Einleitung: Was ist Process Tracing?¹

Process Tracing – eine deutsche Übersetzung wie etwa „Prozessanalyse“ hat sich in der Forschung nicht durchgesetzt – ist eine der wichtigsten Untersuchungsmethoden für qualitative Fallstudien in Politikwissenschaft und Policy-Forschung. Einerseits unterscheidet sie sich als Methode für Kausalerklärungen von qualitativen diskursanalytischen und hermeneutischen Verfahren – trotz jüngerer Annäherungen (Nullmeier 2019; Norman 2015). Andererseits stehen im Unterschied zu vielen in erster Linie fallvergleichenden und/oder quantitativen Methoden jedoch nicht die Korrelationen zwischen unabhängigen Variablen und abhängiger Variable, sondern die Kausalmechanismen im Zentrum, die unabhängige und abhängige Variable verbinden und die sich innerhalb eines oder mehrerer Fälle nachweisen lassen. Dies kann entweder dem Theorietest, der Theoriegenerierung oder der Erklärung eines spezifischen Falles dienen. Die Fallzahl – „N“ im Sinne von Untersuchungseinheiten – ist beim Process Tracing in der Regel klein, die Datengrundlage – „N“ im Sinne von empirischen Beobachtungen – jedoch sehr breit und vielfältig. Verschiedenste Erhebungsmethoden und Datenformate können für Process Tracing genutzt werden, ebenso wie Sekundärdaten. Darin ähnelt die Methode historischen Untersuchungen, hat jedoch einen deutlich theoretischeren, abstrahierenden Zugang. Wie im Folgenden ausgeführt wird, gelingt die Validierung theoretischer Hypothesen beim Process Tracing über fall- und kontextgebundene Schlussfolgerungen. Empirische Beobachtungen werden als potenzielle empirische Implikationen – oder „Spuren“ – einer Theorie verstanden. Dieses Spurenlesen folgt spezifischen Regeln, die helfen, Fallbeobachtungen zu gewichten und miteinander in Beziehung zu setzen. Zusammenfassend lautet die diesem Beitrag zu Grunde liegende *Definition* demnach: Process Tracing ist eine Untersuchungsmethode zur kausalen Erklärung, bei der vielfältige empirische Beobachtungen innerhalb eines oder mehrerer Fälle als potenzielle Implikationen theoretischer Kausalmechanismen verstanden werden. Die möglichst vollständige empirische Rekonstruktion kausaler Prozesse durch Fallstudien erlaubt Schlussfolgerungen über (alternative) theoretische Erklärungen.

Eine Reihe von Ansätzen ist relativ eng mit dem Process Tracing verwandt, u. a. „analytic narratives“ (Bates et al. 1998; Levi 2002), „kausale Rekonstruktion“ (Mayntz 2002), „systematic process analysis“ (Hall 2006) und „causal narrative“ (Mahoney 2003). Bei anderen ist die Verbindung eher lose: So beruht etwa „pattern matching“ (Campbell 1975) zwar auf einigen ähnlichen Elementen – wie dem Versuch, Implikationen von Theorien innerhalb eines oder weniger Fälle nachzuweisen –, allerdings sind Kausalmechanismen von geringerer Bedeutung als beim Process Tracing. Klarer noch muss Process Tracing von älteren „Small-N“-Methoden, wie den Mill’schen Methoden (1968 [1872]) abgegrenzt werden, da es bei letzteren die systematischen Unterschiede in den Ausprägungen auf der

¹Für wertvolle Hinweise möchte ich den Herausgebern sowie den Teilnehmern des Autorenworkshops danken, insbesondere Claudio Wagemann, Detlef Jahn, Diana Panke und Jale Tosun. Darüber hinaus bin ich Henning Deters, Ingo Rohlfsing und Frank Stengel für ausführliche Kommentare dankbar.

unabhängigen und abhängigen Variablen sind, die die Basis für kausale Schlussfolgerungen über mehrere Fälle hinweg (*cross-case analysis*) bilden (kritisch hierzu, Lieberson 1991).² Ähnliches gilt für die „*congruence method*“ (George und Bennett 2005, S. 181–204; Beach und Pedersen 2016). Process Tracing hingegen beruht auf Ausprägungsmustern *innerhalb* des oder der untersuchten Fälle (*within-case analysis*). Gleichzeitig ist die *Untersuchungsmethode* Process Tracing nicht gleichbedeutend mit dem *Forschungsdesign* (Einzel-) Fallstudie bzw. Fallvergleich.³ Die Assoziation von Process Tracing und Fallstudiedesigns beruht eher auf forschungs-pragmatischen Entscheidungen als auf methodologischen Notwendigkeiten.

Dieser Beitrag zielt darauf ab, den Begriff und die Praxis des Process Tracing genauer zu fassen und ihre Möglichkeiten (und Grenzen) über eine Reihe von Beispielen zu verdeutlichen. Weiterführend sei besonders auf die programmatische Monografie von George und Bennett (2005), auf einen Sammelband von Bennett und Checkel (2015a) sowie auf die Lehrbücher von Rohlfing (2012) und von Beach und Pedersen (2012) verwiesen. Zudem gibt es verschiedene Kurzeinführungen in Artikelform (Collier 2011; Ricks und Liu 2018; Checkel 2008; Trampusch und Palier 2016). Auf deutsch liegt seit Kurzem auch eine Monographie von Nullmeier vor (Nullmeier 2021).

Der Beitrag ist wie folgt gegliedert: An einen kurzen historischen Abriss schließt eine Erläuterung der wissenschaftstheoretischen Annahmen an, die dem Process Tracing zu Grunde liegen sowie der Forschungsziele, die mit ihm üblicherweise verfolgt werden. Eng damit zusammen hängt die Fallauswahl. Danach widme ich mich den Daten, die typischerweise für das Process Tracing herangezogen werden. Die Datenauswertung, insbesondere die Gewichtung von Beobachtungen und die zentralen Arten von Schlussfolgerungen sind Thema des darauffolgenden Abschnitts. Es folgt eine kleine Auswahl an Beispielen aus der Policy-Forschung. Die Stärken, Schwächen und Gütekriterien von Process Tracing werden im vorletzten Abschnitt behandelt. Das Fazit fasst den Beitrag zusammen.

2 Ursprünge, Annahmen, Forschungsziele und Fallauswahl

2.1 Die Entwicklung der Methode

Der Begriff Process Tracing entstand vor mehr als 40 Jahren in der Psychologie und bezeichnete ursprünglich eine Methode, kognitive Entscheidungsprozesse auf Grundlage von direkten Äußerungen der Probanden, beobachtbarem Verhalten

²Die Differenzmethode und die Konkordanzmethode sind die bekanntesten Methoden, die John Stuart Mill beschrieben hat. Ursprünglich ausschließlich für experimentelle Forschungen gedacht, hatten sie einen bedeutenden Einfluss auf so genannte „Small-N-Vergleiche“ in der Sozialwissenschaft (Skocpol 1979). In der Politikwissenschaft zeichnet sich das „*most similar systems design*“ von Przeworski und Teune (1970) durch eine große Ähnlichkeit zu Mills Differenzmethode aus.

³Gleichzeitig muss sie von *Erhebungsmethoden* – etwa semistrukturierten Interviews – abgegrenzt werden, die auch für andere Untersuchungsmethoden genutzt werden können.

sowie physiologischen und neuropsychologischen Daten zu rekonstruieren (Ford et al. 1989; Einhorn et al. 1979; Hogarth 1974; Kühberger et al. 2011, zur Geschichte s. ausführlich Nullmeier 2021). In den späten 1970ern übernahm der Politikwissenschaftler Alexander L. George die Methode und entwickelte sie weiter, zunächst alleine (1979), später dann zusammen mit Timothy McKeown (George und McKeown 1985). In Abgrenzung von der psychologischen Literatur sollten allerdings nicht nur kognitive, sondern auch soziale Prozesse in die Analyse einbezogen werden (s. Bennett 2008a; Bennett und Checkel 2015b). Es dauerte allerdings weitere 20 Jahre bis Process Tracing den politikwissenschaftlichen Mainstream erreichte. Das Buch *Case Studies and Theory Development in the Social Sciences* von Alexander George und Andrew Bennett (2005) war hierfür der entscheidende Impuls. Mittlerweile ist Process Tracing Teil des Methodenkanons politikwissenschaftlicher Forschung. Die Methode ist heute weitaus weniger von der Marginalisierung bedroht als vielmehr von einem „buzzword problem [...], where process tracing is mentioned, but often with little thought or explication of how it works in practice“ (Bennett und Checkel 2015b, S. 4). Ziel dieses Beitrags ist auch zu zeigen, dass es sich bei Process Tracing um mehr als ein buzzword handelt.

2.2 Zentrale Annahmen

Eine Besonderheit des Process Tracing ist die Bedeutung, die Kausalmechanismen bei der Methode einnehmen. Dies steht in deutlichem Gegensatz einerseits zu einer klassisch positivistisch geprägten Perspektive in der Wissenschaftstheorie, andererseits zu vielen konstruktivistischen Ansätzen, die sich beide – wenn auch aus unterschiedlichen Gründen – durch eine Skepsis gegenüber kausalen Aussagen auszeichnen. Der Positivismus, ausgehend von David Hume, akzeptiert lediglich Aussagen über empirische Regelmäßigkeiten, möchte daraus aber keine Kausalitätsaussagen ableiten.⁴ Und konstruktivistische Ansätze blicken mit Skepsis auf Kausalaussagen, insbesondere wenn sie als soziale Gesetzmäßigkeiten formuliert sind (Yanow 2006; Schwartz-Shea 2006). Soziale Ursachen werden darüber hinaus nicht einfach als Teil der äußeren Welt, sondern oft als Ausdruck sozial geprägter Sinngebung gesehen. Kausalitätszuschreibungen können so selbst zum Untersuchungsobjekt werden.

Nun ist seit einiger Zeit die Kausalität in den politikwissenschaftlichen Mainstream zurückgekehrt, einerseits über den Umweg der Ökonomie mit ihrer „causal identification revolution“ und designbasierter quantitativer Forschung (Morgan 2016) und andererseits über Process Tracing. Die beiden Ansätze haben sich zwar unabhängig voneinander entwickelt, sie verbindet allerdings eine Kritik an herkömmlichen quantitativen Methoden, besonders an ihrer Unfähigkeit, zwischen kausalen und un-

⁴Die immer noch für Teile der Sozialwissenschaften maßgebliche Version des Positivismus, das deduktiv-nomologische Erklärungsmodell oder „covering law“-Modell von Hempel und Oppenheim (1948), folgt Hume und weicht Kausalitätsaussagen systematisch aus.

echten (spurious) Regelmäßigkeiten zu unterscheiden. Während sich die „causal identification“-Methoden jedoch an experimentelle Forschung anlehnen (dazu Janikowski und Tepe i. d. B.), basiert Process Tracing eher auf Erklärungsformen in der Geschichtswissenschaft und Kausalitätsvorstellungen des wissenschaftstheoretischen Realismus (Psillos 1999). Hier steht – auch im Kontrast zu experimenteller Forschung – nicht der *Effekt* einer unabhängigen Variable auf eine abhängige Variable im Vordergrund, sondern der kausale *Mechanismus*, der diesen Effekt produziert.

Aber was bedeutet Mechanismus eigentlich? Mechanismen verbinden Ursache und Wirkung, soviel ist allgemein anerkannt. Was genau „verbinden“ heißt und was der ontologische Status des Mechanismus selbst ist, bleibt jedoch umstritten (s. Gerring 2007b; Machamer et al. 2000; Glennan und Illari 2018). Mechanismen sind jedenfalls nicht mit „intervenierenden Variablen“ gleichzusetzen; sie generieren vielmehr selbst die Wirkung innerhalb eines spezifischen Settings oder Kontextes (vgl. Falletti und Lynch 2009). Umstritten ist unter anderem, ob Mechanismen letztlich unbeobachtbar sind oder ob vielmehr die Black Box der Kausalität „geöffnet“ werden kann (Autoren wie Beach und Pedersen 2012 sind hier besonders optimistisch). Die meisten Vertreter des Process Tracing gehen jedoch auch davon aus, dass der Mechanismus nicht als Ganzes beobachtbar ist, sondern weitgehend indirekt erschlossen werden muss. Bestimmte Beobachtungen von Zuständen und Ereignissen können demnach als empirische Implikationen eines Mechanismus (traces) verstanden werden, insbesondere wenn sie in dem Kontext und der Reihenfolge auftreten, die theoretisch erwartbar ist (Nullmeier 2021, S. 194–196). Die Verbindung zwischen theoretischer Spezifizierung eines Mechanismus und der Beobachtung von Implikationen kann theoretestend oder -generierend verstanden werden, wie im Folgenden erläutert wird.

2.3 Forschungsziel: Drei Arten des Process Tracing

Was das Forschungsziel angeht, so werden meist drei Arten von Process Tracing unterschieden (in Anlehnung an Beach und Pedersen 2012; Bennett und Checkel 2015b): (1) deduktiv-theoriestestende Analysen, (2) theorie- oder hypothesesgenerierende Analysen und (3) Erklärungen spezifischer Outcomes. Das jeweilige Forschungsziel bestimmt wiederum die Logik der Fallauswahl (s.u.) und den zeitlichen Ablauf des Forschungsprozesses.

Bei der ersten Art von Process Tracing werden, im Unterschied zum klassischen Hypothesentest in der quantitativen Forschung, Hypothesen nicht primär im Hinblick auf erwartete Effekte überprüft, sondern vielmehr durch Deduktion eine Vielzahl von zu testenden Implikationen einer Hypothese – den „Gliedern“ der Kausalkette – formuliert, die die Grundlage der empirischen Analyse bildet (s. Panke 2012). Zusätzlich zum Nachweis der erwarteten Ausprägungen von unabhängigen und abhängigen Variablen müssen möglichst viele dieser Implikationen des Mechanismus gefunden werden – manche Autoren verlangen gar den empirischen Nachweis der vollständigen Kausalkette (Beach und Pedersen 2012, S. 5). Zuweilen sind es tatsächlich weniger die Erklärungsvariablen, die in der Literatur umstritten sind, als vielmehr, welcher Mechanismus der Erklärung zu Grunde liegt. Ein klassisches Beispiel ist der

„demokratische Frieden“, bei dem nicht die empirische Korrelation – zwischen demokratischen Dyaden und Frieden –, sondern der Mechanismus umstritten ist (Russett 1994; s. George und Bennett 2005, Kap. 2; Müller 2002). In dieser Situation ist Process Tracing besonders hilfreich, weil damit alternative Mechanismen gegeneinander innerhalb eines oder mehrerer Fälle getestet werden können.

Theoriegenerierendes Process Tracing beginnt ohne eine plausible Erklärung für ein Outcome. Ziel ist, den Prozess zwischen Ursache(n) und Outcome zu rekonstruieren und einen möglichen abstrakten Mechanismus und seinen Auslöser induktiv zu erfassen und in Beziehung zum bestehenden Theorierepertoire zu stellen. Selbstverständlich geschieht das nicht „aus dem Nichts“, sondern setzt ein analytisches Instrumentarium von Begriffen, Typologien, Modellen und Mechanismen voraus, mit dem die Beobachtungen theoretisch einzuordnen sind. Durch den starken Einfluss des Falsifikationismus nach Popper in den Sozialwissenschaften ist die Literatur zur systematischen Theoriegenerierung im Vergleich zum Theorietest klar unterentwickelt, systematische Einführungen fehlen bislang (s. aber Swedberg 2011). Sowohl bei theorietestendem als auch bei theoriegenerierendem Process Tracing gehen die Schlussfolgerungen eindeutig über die untersuchten Fälle hinaus, werden also – zumindest in einem gewissen Maß – verallgemeinert.⁵ Fälle werden also als unter bestimmten Bedingungen vergleichbar angesehen; sie sind keine Phänomene *sui generis*; sie werden als „Fälle von etwas“ verstanden (s.u.).

Bei der Erklärung spezifischer Outcomes, dem dritten möglichen Forschungsziel, geht es zwar nicht um Generalisierung über die untersuchten Fälle hinaus, allerdings werden auch hier – im Gegensatz zu vielen geschichtswissenschaftlichen Analysen (s. Elman und Elman 2001) – die theoretischen Mechanismen der Erklärung explizit gemacht. Sie ist also nicht „atheoretisch“ im Sinne Lijpharts (1971, S. 691–693),⁶ hat aber im Sinne der Unterscheidung zwischen „idiografischer“ und „nomothetischer“ Forschung (Windelband 1894) einen „idiografischen“ Einschlag, da das einzelne Ereignis im Vordergrund steht (Gerring 2006b). Ziel ist nicht die möglichst vollständige Beschreibung des Gegenstandes, sondern eine „minimal hinreichende Erklärung“ eines Ereignisses oder Zustandes, die systematische und idiosynkratische Erklärungsfaktoren und -mechanismen kombiniert (Beach und Pedersen 2012, Kap. 4), wobei durchaus zwischen alternativen Erklärungen diskriminiert wird.⁷ Oft wird ein Fall zunächst als „Puzzle“ aus Sicht der Forschung oder des Alltagswissens identifiziert, nicht selten als „Ausreißer“ (deviant case). Ein Beispiel aus der Policy-Forschung ist die fallorientierte Erklärung von Krisen oder „Politikdesastern“ (Sagan 1995), wie etwa Diane Vaughans detaillierte Studie der Challenger-Katastrophe

⁵ Aussagen über die Stärke oder die konkrete Form eines Effekts können üblicherweise nicht im Rahmen von Prozessanalysen getroffen werden.

⁶ Es ist zudem zweifelhaft, ob es so etwas wie „atheoretische“ Forschung angesichts „theoriegeladener“ Daten überhaupt geben kann.

⁷ Minimal hinreichend ist eine Erklärung „that accounts for all of the important aspects of an outcome with no redundant parts being present“ (Beach und Pedersen 2012, S. 18). Es ist aber nicht ausgeschlossen, dass in anderen Fällen eine andere Erklärung für das Outcome gilt; sie ist also nicht notwendig.

(2009). Ziel ist zu verstehen, wie es zu solchen Desastern kommen konnte. Untersuchungsberichte im Gefolge von Skandalen und ähnlichen Ereignissen folgen häufig auch diesem Muster.

Die Grenzen zwischen den drei Arten des Process Tracing sind, wie so oft, fließend. Je stärker nicht nur rein fallspezifische, idiosynkratische, sondern allgemein anwendbare, abstrakte Erklärungsfaktoren auch für die Erklärung spezifischer Outcomes herangezogen werden, desto eher besteht die Möglichkeit, dass die Ergebnisse theoriegenerierenden oder sogar – testenden Charakter bekommen. In der Policy-Forschung ist dies, im Gegensatz zur Geschichtswissenschaft, der Normalfall.

2.4 Fälle und Fallauswahl

Eine bewusste Fallauswahl ist eine Voraussetzung dafür, dass auch Untersuchungen weniger Fälle oder sogar Einzelfallstudien in dieser Form über sich hinauswachsen können. Eine Vorbedingung der Fallauswahl ist das so genannte „casing“ (Ragin 1992; s. a. Rohlffing 2012, S. 24–28), d. h. die Antwort auf die Frage: Was ist ein Fall und von was ist es ein Fall? Beim „casing“ laufen Überlegungen zu Forschungsinteresse, Population, Daten und theoretischem Bezugsrahmen zusammen. In vielen Situationen ist das „casing“ relativ unproblematisch, da bewährte Forschungskonventionen und standardisierte Daten existieren. Zuweilen aber sind die Konventionen für neue Fragestellungen unbrauchbar oder es besteht keine Übereinkunft über die passende Untersuchungseinheit und die angemessene analytische Ebene. Hier ist explizites „casing“ unumgänglich. Für eine Untersuchung sozialpolitischer Reformen können beispielsweise entweder ganze Länder mit ihren sozialstaatlichen Konfigurationen, Einzelprogramme (z. B. Rente, Arbeitslosenversicherung), einzelne Gesetzesvorhaben oder Policy-Instrumente die Fälle der Untersuchung sein.

Besteht schließlich Klarheit darüber, wie die Fälle konstituiert werden, steht die Entscheidung an, welche konkreten Fälle in die Untersuchung einbezogen werden. Hier gibt es eine Reihe von bewährten Verfahren.⁸ Die Notwendigkeit (bewusster) Fallauswahl ergibt sich aus dem Umstand, dass Process Tracing auf zeitaufwändigen qualitativen Erhebungs- und Auswertungsmethoden beruht, die es nur selten erlauben, mehr als eine Handvoll Fälle einzubeziehen. Statt auf Zufallsauswahl beruht Process Tracing, wie im nächsten Abschnitt ausgeführt wird, auf logischen oder bayesianischen⁹ Schlussfolgerungen. Es sind nur sehr wenige Situationen denkbar,

⁸In diesem Beitrag werden nur jene Methoden der Fallauswahl vorgestellt, die konkret für die Prozessanalyse relevant sind. Methoden, deren Schlüsse auf „cross-case analysis“ beruhen, wie die Mill'schen Methoden, bleiben außen vor.

⁹Bayesianische Schlüsse beruhen auf dem Satz von Bayes zur Berechnung bedingter Wahrscheinlichkeiten. Der Satz von Bayes erlaubt es, auf Grund bekannter, zuweilen subjektiver Wahrscheinlichkeiten (z. B. von empirischen Beobachtungen) auf andere Wahrscheinlichkeiten (z. B. von Erklärungen für diese Beobachtungen) zu schließen bzw. ein „update“ letzterer vorzunehmen. Bayesianische Methoden in der Statistik werden üblicherweise „frequentistischen“ Methoden gegenübergestellt, zu denen die meisten derzeit in der Politikwissenschaft verbreiteten quantitativen Verfahren gehören.

in denen überhaupt keine Fallauswahl vorgenommen werden muss, etwa wenn es sich um eine Art „Vollerhebung“ eines extrem seltenen Phänomens handelt. Allerdings beruhen nicht alle Methoden der Fallauswahl ausschließlich auf positiven Fällen (d. h. auf dem Vorliegen eines interessierenden Outcomes auf der abhängigen Variable); häufig müssen vielmehr negative Fälle (bei denen das Outcome nicht vorliegt) zusätzlich untersucht werden (Mahoney und Goertz 2004) oder es ist zunächst unklar, ob es sich um positive oder negative Fälle handelt, da auf der unabhängigen Variable ausgewählt wird.

2.5 Methoden der Fallauswahl

Methoden der Fallauswahl lassen sich grob in drei Gruppen einteilen (s. a. Seawright und Gerring 2008, in Anlehnung an Rohlfing 2012), S. 1) auf empirischen Merkmalsverteilungen, 2) auf empirischen Zusammenhängen und 3) primär auf theoretischen Erwartungen basierende Methoden (Tab. 1).

In die erste Kategorie fallen Methoden, bei denen Fälle gewählt werden, die sich durch „besondere“ Werte auf der abhängigen oder einer oder mehreren unabhängigen Variablen auszeichnen. Hierzu zählt insbesondere die *Extremfallmethode*. Die Extremfallmethode ist theoriegenerierend. Hierzu wählt man einen Fall aus, der extrem hohe oder niedrige Ausprägungen auf einer interessierenden – unabhängigen *oder* abhängigen – Variable aufweist. Die Merkmalsverteilung innerhalb der Population muss, zumindest annäherungsweise, bekannt sein (beispielweise durch Kenntnis der Sekundärliteratur zum Thema oder quantitativer Indikatoren). Man wählt z. B. ein Land, in dem radikale Reformen stattfanden, in der Annahme, dass auch ebenso klare Veränderungen in der dafür entscheidenden unabhängigen Variable stattfanden und – dies ist für Process Tracing zentral – ein Mechanismus deutlich nachweisbar ist. Alternativ kann ein extremer Wert auf einer unabhängigen Variable ausgewählt und seine Wirkungen untersucht werden (Seawright 2016) oder aber, wenn der Zusammenhang aus anderen Studien bereits bekannt ist, eine extreme Kombination aus Ursache und Wirkung, mit dem Ziel, den Mechanismus zu

Tab. 1 Systematik der Fallauswahl

Basis der Fallauswahl	Methode	Auswahl auf X und/oder Y	Forschungsorientierung
Verteilung	Extremfall	X oder Y	Theoriegenerierend
	Diverse cases	X	Theoriegenerierend oder -testend
	Diverse cases	Y	Theoriegenerierend
Zusammenhang	Typischer Fall	X/Y-Zusammenhang	Theoriegenerierend oder – testend
	Ausreißer (deviant case)	X/Y-Zusammenhang	Theoriegenerierend
Theorie	Most-likely	X	Theorietestend
	Least-likely	X	Theorietestend

beobachten. Ein Nachteil der Extremfallmethode ist die Möglichkeit u-förmiger oder anderer nicht-linearer Zusammenhänge, die die Logik der Extremfallmethode nichtig machen können. Zwar sollten Mechanismen innerhalb des Falles weiterhin beobachtbar sein, allerdings ist ein Schluss auf andere Fälle nicht mehr so einfach: Während unter Annahme eines linearen Zusammenhangs die Untersuchung eines extrem niedrigen und extrem hohen Falls Schlüsse auf die dazwischenliegenden Fälle zumindest nahelegt, trifft dies bei Nicht-Linearität nicht mehr zu. Die Möglichkeit sollte daher stets in Betracht gezogen werden. Des Weiteren sollte überlegt werden, ob es sich bei den gewählten Fällen eventuell um abweichende Fälle handelt oder gar um nicht-vergleichbare Fälle, die aus der relevanten Population fallen.

Die *Diverse-Cases-Methode* (Seawright und Gerring 2008) beruht auf der Idee, dass mehrere Fälle, die zusammen möglichst die gesamte Breite der Verteilung der Variablen abbilden, tendenziell eher repräsentativ sind (selbst wenn sie kein Zufalls-sample bilden). Hier wird also dem Zielkonflikt zwischen Fallzahl und Repräsentativität direkt Rechnung getragen. Vorstellbar ist die Auswahl von drei Fällen, einem mit extrem niedriger Ausprägung auf einer unabhängigen (oder der abhängigen) Variable, einem mit hoher und einem durchschnittlichen Fall. Für theoriegenerierendes Process Tracing kann die Auswahl auch auf der abhängigen Variable getroffen werden (King et al. 1994, S. 141–142). Eine etwas andere Begründung der *Diverse-Cases*-Auswahl ist die Forderung, Kausalmechanismen in möglichst unterschiedlichen Kontexten nachzuweisen, um so die Validität der Schlussfolgerungen zu erhöhen (Levi-Faur 2006). Hier dienen also nicht unabhängige und abhängige, sondern Kontextvariablen (z. B. Sektor, Region, Kultur) als Grundlage für die Fallauswahl.

Bei *typischen Fällen* oder ihrem Gegenteil, den *abweichenden Fällen* (Ausreißern), wird die Auswahl auf der Grundlage eines bekannten empirischen Zusammenhangs getroffen. Dies kann, muss aber nicht, im Rahmen eines Mixed-Methods-Designs geschehen (Rohlfing und Starke 2013; Lieberman 2005). Ein typischer Fall ist ein durch bestehende Theorien besonders gut erklärter Fall. Allerdings ist häufig der Mechanismus bzw. sind die Mechanismen ungeklärt oder es existieren alternative Mechanismen in der Literatur. Der typische Fall wird also herangezogen, um prozessanalytisch generalisierbare Mechanismen herauszuarbeiten (Theoriegenerierung) – die dann an weiteren Fällen oder neuem Datenmaterial getestet werden können – oder um zwischen Mechanismen zu diskriminieren (Theorietest). Process Tracing abweichender Fälle (s. Lijphart 1971) ist eindeutig theoriegenerierend, da hier sowohl der Mechanismus als auch die unabhängige Variable unbekannt sind, die dafür sorgen, dass der Fall nicht (gut) erklärt wird. Allerdings ist zu Beginn der Untersuchung nicht klar, ob die Ursache für die Abweichung rein kontingent und fallspezifisch oder auf andere Fälle übertragbar und potenziell generalisierbar ist. Mit anderen Worten, es ist häufig zu Beginn nicht klar, ob die Analyse letztendlich „nur“ eine Erklärung eines spezifischen Falls oder theoriegenerierend ist.

Zuletzt seien Arten der Fallauswahl erwähnt, die in erster Linie theoriebasiert sind. Bei *most-likely cases* liegen starke theoretische Gründe vor, die einen Mechanismus oder ein bestimmtes Outcome erwarten lassen. So wählte Paul Pierson (1994) für seine Untersuchung von sozialpolitischer Kürzungspolitik die Regierungen Reagan und Thatcher aus, da vor dem Hintergrund der Parteiendifferenztheorie

(sowie den Absichtserklärungen der Akteure) Kürzungen zu erwarten waren. Umgekehrt sind diese Erwartungen bei einem *least-likely case* besonders schwach.¹⁰ Der Witz dieser Methode ist nun, dass ihre Ergebnisse besonders dann von Interesse sind, wenn die Erwartungen *nicht* erfüllt werden, also z. B. bestehende Theorien eine bestimmte Reform erwarten lassen, aber überraschenderweise nichts passiert ist. Es bleibt umstritten, ob Theorien tatsächlich mit Hilfe von einzelnen *most-likely* oder *least-likely*-Fällen effektiv getestet werden können (Gerring 2007a; Eckstein 1975). Process Tracing sollte daher nicht bei der Konstatierung des „falschen“ Outcomes stehenbleiben, sondern die eventuell dahinterstehenden alternativen Erklärungen und Mechanismen beleuchten. So nimmt etwa Pierson das Ausbleiben von radikalen Kürzungen als Ausgangspunkt für seinen Ansatz der „new politics of the welfare state“.

3 Datengrundlage: Systematik jenseits der Standardisierung

3.1 Prozessbeobachtungen als Datengrundlage

Process Tracing eines einzelnen Falls sollte nicht als nicht-determinierte N = 1-Studie missverstanden werden. Als within-case analysis beruht sie stets auf einer Vielzahl von Datenpunkten und profitiert damit, in den Worten des Methodologen Donald Campbell von der „discipline coming from a richness of relevant details“ (1975, S. 175). In der Praxis ist daher nicht ein Mangel an Daten, sondern eher ihr Überfluss und ihre ungeordnete Form problematisch.

Während in quantitativen Analysen und teilweise auch in einigen fallvergleichenden Designs die Standardisierung der Daten am Anfang stehen muss, erlaubt Process Tracing hier größere Freiheiten. An dieser Stelle wird oft die Unterscheidung zwischen „data-set observations“ (DSOs) und „causal-process observations“ (CPOs) verwendet (Collier et al. 2010, S. 182–196). CPOs sind dabei etwas anderes als ein Wert bzw. eine Zeile mit Werten in einem tabellarischen Datensatz (DSOs); sie bezeichnen vielmehr „an insight or piece of data that provides information about context or mechanism“ (Collier et al. 2010, S. 184). Auch wenn also Kausalität selbst nicht beobachtbar ist, so bleibt die Analyse von Prozessen und somit die Temporalität der Daten zentral (Grzymala-Busse 2011). Bereits George and McKeown (1985) betonten, dass bei Process Tracing nicht der einzelne Datenpunkt, sondern „a stream of behavior through time“ (1985, S. 36) theoretischen Wert besitzt. Es geht also um möglichst gesicherte Informationen über den Ablauf und die Reihenfolge von Ereignissen, ihre Dauer und ihr Tempo, beteiligte Akteure und deren relative Positionen und Ressourcen, Erwartungen vor und Einschätzungen nach zentralen Ereignissen, (informelle oder formelle) Entscheidungsprozesse,

¹⁰Bisweilen wird in diesem Zusammenhang auch der „crucial case“ erwähnt, der aber lediglich eine besonders starke Variante des *most-likely/least-likely*-Designs darstellt (Rohlfing 2012; Eckstein 1975; Gerring 2007a).

relevante und verworfene Optionen und Ähnliches. Weder müssen diese Beobachtungen unabhängig voneinander sein – was bei Prozessen ohnehin widersinnig ist – noch sind sie gleichgewichtig. Im Gegenteil, beim Process Tracing erhalten häufig einige wenige Beobachtungen ein besonderes Gewicht für die Argumentation (s. nächster Abschnitt).

Auch das Datenformat ist vielfältig. Eine besondere Stärke von Process Tracing liegt in der grundsätzlichen Offenheit gegenüber völlig unterschiedlich gearteten Daten im Hinblick auf den Grad ihrer Standardisierung, Skalenniveau, Quelle, öffentliche Verfügbarkeit etc. Sekundär- und Primärdaten und deren Erhebung durch Feldforschung sind grundsätzlich zulässig und wünschenswert. Die verwendeten Daten beschränken sich zudem nicht auf qualitative Formate; es ist vielmehr möglich, ja sogar von Vorteil, wenn quantitative Daten, sofern vorhanden, auch in primär qualitativen Fallstudien verwendet werden (Maxwell 2010).¹¹ Teilweise ähnelt die Vorgehensweise der Triangulation, bei der auch verschiedene Datenformate und -quellen kombiniert werden, um die Reliabilität zu erhöhen (s. Siewert i. d. B.), sie ist aber weniger formalisiert. Es ist sogar möglich, dass innerhalb einer Studie derselbe (Teil eines) Mechanismus in verschiedenen Fällen unterschiedlich gemessen wird, je nachdem, was für den jeweiligen Fall angemessen – d. h. valide – ist.

Um eine ausreichende Datenbasis für Process Tracing zu erhalten, muss also das Netz sehr weit ausgeworfen werden. Diese Vorgehensweise wird üblicherweise mit der eines Detektivs verglichen (Collier 2011). Ein Detektiv versucht, möglichst viele verfügbaren Beobachtungen eines Falls mit den theoretischen Implikationen alternativer Erklärungen – vergleichbar mit den Verdächtigen eines Kriminalfalls – abzugleichen und so zu einer kohärenten Schlussfolgerung zu gelangen (s. a. nächster Abschnitt). Der Kontrast mit hochstandardisierten quantitativen (aber auch einigen qualitativen) Verfahren könnte kaum größer sein, dennoch kann dabei eine große Systematik erreicht werden, allerdings in erster Linie über den Dialog zwischen Theorie und Empirie, d. h. die enge Verknüpfung des Nachdenkens über die Implikationen von (alternativen) Kausalmechanismen und der empirischen Beobachtung: „All pieces of evidence are relevant to the central argument (they are not „random“), but they do not comprise observations in a larger sample“, so Gerring (2006a, S. 178). Die Kehrseite der Detektivmethode ist der große Aufwand, der zur Datenerhebung für überzeugendes Process Tracing notwendig ist (George und McKeown 1985, S. 41) (Pierson 1994).

3.2 Sekundäranalyse und Primärerhebung

Prozessanalytische Forschung bewegt sich zwischen Sekundäranalyse und primärer Forschung, zwischen der Auswertung von Dokumenten und Feldforschung. Die Praxis, die bestehende historische Literatur für Process Tracing sekundäranalytisch

¹¹ Erst wenn auch quantitativ-statistische Auswertungsverfahren verwendet werden, spricht man von einem „Mixed-Methods“-Design.

„auszubeuten“, ist jedoch weit verbreitet (Mahoney und Villegas 2007). Der Vorteil besteht darin, dass mehrere Fälle z. T. über größere Zeiträume verglichen werden können. Im Sinne der disziplinären Arbeitsteilung verlässt man sich hier auf das Urteil von Länder- oder Epochenexpertinnen, um stärker theoretisch orientiert zu arbeiten und die hohen zeitlichen und finanziellen Kosten von eigenen Recherchen zu vermeiden (Skocpol 1984). Das kann gut gehen, muss es aber nicht. Der schlechteste Fall ist das unkritische „Rosinenpicken“ („cherry picking“) von historischen Einzelbeobachtungen, die außerhalb ihres ursprünglichen Analysekontextes keine Gültigkeit aufweisen (Lustick 1996). Auch kann es passieren, dass (aus welchen Gründen auch immer) verzerrte Darstellungen von Historikerinnen kritiklos übernommen werden und so ihr *bias* reproduziert wird. Zur Steigerung von Validität und Reliabilität historischer Sekundäranalysen schlägt Cameron Thies einige Grundregeln vor, u. a.: wenn möglich, immer mehrere Autoren zu Rate zu ziehen, sich dabei chronologisch von neueren zu älteren Publikationen vorzuarbeiten und den ursprünglichen Forschungskontext der vermeintlich neutral berichteten Fakten mitzudenken (Thies 2002).

Aufgrund der genannten Probleme empfiehlt es sich, stets die sekundäranalytische Datengenerierung primäranalytisch zumindest zu ergänzen. Die wichtigsten Methoden sind Archivrecherche und Interviews, oft im Rahmen von ausgedehnten Feldforschungen (zu praktischen Aspekten, s. Lieberman et al. 2004). Die Archivrecherche fristet ein Schattendasein in der politikwissenschaftlichen Methodenliteratur (eine Ausnahmepublikation bildet Frisch et al. 2012). Das ist insofern überraschend, als sich viele Politikwissenschaftlerinnen mit Vorgängen befassen, die in Akten nachweisbar sind (sofern man weiß, wie man sie findet). In der politikwissenschaftlichen Ausbildung sind Grundlagen in der Archivkunde eine Ausnahmeerscheinung und so müssen sich interessierte Studierende und Forscherinnen diese Methoden zwangsläufig selbst aneignen.¹² Die zunehmende Verfügbarkeit von Dokumenten online ist eine Chance, doch sie birgt die Gefahr der Verzerrung, wenn nicht zusätzlich nach ausschließlich „offline“ vorhandenem Material gesucht wird. Es ist leicht vorstellbar, dass bestimmte politische Akteure bei der (Online-)Publikation von Dokumenten höchst selektiv und mit Blick auf ihre eigene Öffentlichkeitswirkung vorgehen. Primärquellen müssen daher aktiv gesucht und aufbereitet werden. Dies darf aber nicht darüber hinwegtäuschen, dass die Quellenlage selbst auch historisch und politisch geprägt ist:

Historical evidence is drastically and irremediably incomplete; it is often of dubious validity; and it tends to be biased – favoring the victors rather than the losers, the lasting developments rather than the historical dead ends and detours, the rich and the educated rather than the poor and illiterate, and so forth (Rueschemeyer und Stephens 1997, S. 56; s. a. Thies 2002).

¹²Dobson und Ziemann bieten z. B. einen Überblick über die Herausforderungen verschiedener Textarten, wie Reden, Briefe, autobiografische Quellen (Dobson und Ziemann 2009). Auf Deutsch existieren eine Reihe von Einführungen in die Archivarbeit für Historiker, die teilweise auch für Politikwissenschaftler relevant sind (Beck und Henning 2012; Brenner-Wilczek et al. 2006).

3.3 Interviews

Qualitative Interviews sind als Form der Primärerhebung deutlich besser etabliert als die Archivrecherche. Für Politikwissenschaftlerinnen ist neben der allgemeinen Literatur zu qualitativen Interviews aus den Nachbardisziplinen Soziologie, Ethnologie und Psychologie insbesondere die Spezialliteratur zu Experten- und Eliteninterviews von Belang (Mosley 2013; Dexter 2006 [1970]; Bogner et al. 2002; Leech 2002; Goldberg und Hildebrandt 2020).¹³ Insbesondere bei Interviews mit Entscheidungsträgerinnen und Interessengruppen können eine Reihe von generellen Schwierigkeiten von Interviews – Machtasymmetrien, Manipulation, *hindsight bias* – besonders ausgeprägt sein. Die Auswahl der Teilnehmerinnen wird bewusst – also nicht über Zufallsauswahl – vorgenommen und orientiert sich am Kontextwissen über Entscheidungsprozesse (Lynch 2013; Tansey 2007). Für Process Tracing eignen sich in erster Linie längere, semistrukturierte Interviews. Bei der Auswertung gibt es unterschiedliche Herangehensweisen. Oftmals geht es um die Rekonstruktion von manifesten Ereignissen und Prozessen mit Hilfe von Beteiligten bzw. um Kontextwissen (etwa zu einem Politikfeld). Hier kann die Auswertung relativ *ad hoc* geschehen. Ähnlich verhält es sich mit illustrativen Zitaten. Wenn allerdings Merkmale wie Präferenzen, Argumentationsweisen und Positionen von mehreren Akteuren erhoben werden sollen, so liegt eine systematische Codierung der Interviews mit Hilfe von qualitativer Analysesoftware (z. B. MAXQDA, Atlas.ti, NVivo) nahe. An diesem Punkt wäre allerdings eine bessere Rezeption der interpretativen Sozialwissenschaft – aus der solche Auswertungsverfahren stammen – durch kausal argumentierendes Process Tracing wünschenswert (für eine hilfreiche Einordnung, s. Gläser und Laudel 2013). Noch ist diese Form der systematischen Auswertung von Interviews und anderen Primärdaten bei Process Tracing unüblich.

4 Datenauswertung: Spurenlesen als Methode

4.1 Van Everas vier Arten empirischer Evidenz

Wie genau wird nun aus den vielfältigen Spuren, die detektivisch zusammengetragen wurden, überzeugendes Process Tracing? Gibt es konkrete Auswertungsregeln? Klar ist, dass bei Process Tracing nicht die reine Anzahl von Beobachtungspunkten etwas über die Qualität der Schlussfolgerungen aussagt. Vielmehr können einzelne Beweisstücke ein besonders großes Gewicht annehmen. Die Nähe zu kriminalistischen Ermittlungen und juristischen Beweisen, aber auch zur klinischen Diagnostik ist auch hier offenbar. Die aktuelle Debatte dreht sich nun darum, was genau die Gewichtung von

¹³Die Literatur zu qualitativen Interviews allgemein ist sehr umfangreich und reicht von theoretischen bis zu eher praktisch orientierten Abhandlungen (Gläser und Laudel 2010; Rubin und Rubin 2011; Rathbun 2008; Harvey 2011).

Tab. 2 4 Arten von Tests

		Gewissheit (certainty)	
		Hoch	Niedrig
Trennschärfe (uniqueness)	Hoch	„Doubly-decisive test“	„Smoking gun test“
	Niedrig	„Hoop test“	„Straw-in-the-wind test“

Quelle: eigene Darstellung auf Basis von Van Evera (1997)

Beobachtungen bestimmt.¹⁴ Einen ersten wichtigen Beitrag leistete Van Evera (1997) mit seiner idealtypischen Unterscheidung von vier Arten qualitativer Tests (s. Tab. 2): „straw-in-the-wind test“, „hoop test“, „smoking gun test“ und „doubly-decisive test“. Die vier Tests unterscheiden sich im Grad der „Gewissheit“ (certainty), d. h. der Wahrscheinlichkeit, mit der eine Hypothese stimmt, wenn gewisse Evidenz vorliegt, sowie im Grad ihrer „Trennschärfe“ (uniqueness)¹⁵, d. h. die Wahrscheinlichkeit, mit der bei Nichtvorliegen automatisch die alternative Hypothese korrekt ist. Letzteres hängt davon ab, ob die empirischen Vorhersagen von mehreren Theorien gleichzeitig gemacht werden (Zaks 2017). Diese Typologie klingt zunächst einmal verwirrend, was insbesondere damit zu tun hat, dass in einigen Fällen das Vorliegen und das Nicht-Vorliegen von Evidenz unterschiedlich starke Schlussfolgerungen begründen kann. Daher gehe ich im Folgenden etwas genauer auf die vier unterschiedlichen „Tests“ ein.¹⁶ Wichtig ist zu beachten, dass sich Gewissheit und Trennschärfe jeweils auf mehr als eine alternative Erklärung/Hypothese eines Outcomes beziehen. Man kann die beiden Dimensionen zudem auch in der Terminologie von notwendiger und hinreichender Bedingung verstehen (Rohlfing 2012, S. 182–183; Mahoney 2012).

Der „straw-in-the-wind test“ ist der schwächste der vier Tests, da er sich weder durch Gewissheit noch durch Trennschärfe auszeichnen. Van Evera (1997, S. 32) nennt hier als Beispiel das Fehlen eines schriftlichen Führerbefehls zum Holocaust. Weder ist ein solches Dokument notwendig, um zu argumentieren, dass Hitler persönlich den Holocaust befahl; noch erlaubt seine Abwesenheit automatisch starke Schlüsse darüber, dass Hitler ihn nicht unterstützte oder über alternative historische Ursachen für die nationalsozialistische Vernichtungspolitik. Würde man ein solches Dokument finden, so hätte es sicher seinen Platz in der Argumentation, würde aber nicht ausreichen, um alternative Erklärungen zu entkräften.

Beim „hoop test“ sind Beobachtungen gewiss, aber nicht trennscharf. Negative Evidenz kann eine Hypothese zwar entkräften, aber positive Evidenz hat nur sehr geringe Aussagekraft. Das Beispiel hier ist etwa die Anwesenheit des Verdächtigen in der Stadt, in der ein Verbrechen stattfand. Kann ein Beschuldigter beweisen, dass er

¹⁴ Es ist umstritten, ob und wie genau auch kontrafaktische Schlüsse in Prozessanalysen verwendet werden sollten (Lebow 2010). Während George und Bennett solchen Schlüssen kritisch gegenüberstehen (2005, S. 167–170), plädieren Autoren wie Rohlfing dafür, sie systematisch mit einzubeziehen (2012, S. 175–178).

¹⁵ Ich danke Henning Deters für den Vorschlag dieser Übersetzung.

¹⁶ Um die Verwirrung noch zu steigern, können die vier „Tests“ nicht nur bei theoretestenden Analysen zum Einsatz kommen. Test ist hier eher als diagnostischer Test oder Prüfverfahren zu verstehen.

außer Landes war, so ist er wohl unschuldig, seine Anwesenheit selbst beweist aber nur, dass es ihm möglich war, die Tat zu begehen, nicht, dass er sie begangen hat.

Bei einem „*smoking gun test*“ verhält es sich genau entgegengesetzt. Geringe Gewissheit und hohe Trennschärfe eines solchen Tests bewirken, dass zwar ein positiver Nachweis entscheidend sein kann, eine Hypothese ihren Alternativen vorzuziehen, ein negativer Befund jedoch nicht. Der Verdächtige mit dem rauchenden Colt war wohl der Täter. Nur, weil ein Verdächtiger keinen rauchenden Colt in der Hand hält, kann er allerdings noch nicht als entlastet gelten.

Der „*doubly-decisive*“ Test ist vergleichbar mit Aufnahmen eines Verbrechens (z. B. eines Banküberfalls) durch eine Überwachungskamera. Einerseits ist positive Evidenz (z. B. ein deutlich erkennbares Gesicht) in hohem Maße inkriminierend (= hohe Gewissheit), andererseits kann sie als entlastend gelten, wenn der Beschuldigte nicht zu sehen ist (= Trennschärfe).

Die vier Testformen von Van Evera sind in der Folgezeit ausführlich diskutiert worden.¹⁷ Insbesondere Bennetts bayesianisch inspirierte Interpretation ist erhellt (Bennett 2008b, 2015). Inferenz bei Process Tracing wird analog zum bayesianischen Schließen mit bedingten Wahrscheinlichkeiten als ein „Updaten“ von a priori Wahrscheinlichkeiten gesehen (s. a. Beach und Pedersen 2012). Vereinfacht gesagt: Je nachdem, wie überraschend oder einzigartig Beobachtungen aus dem Blickwinkel bestimmter Hypothesen, des allgemeinen Fallwissens und den theoretischen Erwartungen der Literatur sind, desto stärker ihr Gewicht in der Argumentation. Im Gegensatz zu falsifikationistischen Vorstellungen von Tests gehen bayesianische Tests davon aus, dass wir durch Beobachtung bestimmter Evidenz die (alternativen) Hypothesen, mit denen wir an einen Fall herangehen, „updaten“ können. Häufig hängt dieses Updaten davon ab, wie etabliert eine Theorie und wie „überraschend“ (vereinfacht gesprochen) die gefundene empirische Evidenz vor dem Hintergrund der bestehenden Literatur sind. Eine bislang sehr gut belegte Theorie (hohe Priorwahrscheinlichkeit) wird von neuer widersprechender Evidenz nur marginal entkräftet werden können, jedoch spielt es eine große Rolle, wie überraschend und neuartig diese Evidenz ist.¹⁸

Das Ziel von Process Tracing ist nun, geeignete Evidenz für möglichst starke Tests zu finden, also z. B. gescheiterte „*hoop tests*“, positive „*smoking gun tests*“ und vielleicht sogar einen „*doubly decisive test*“.¹⁹ Dafür wiederum ist es unerlässlich, dass die empirischen Implikationen alternativer Erklärungen (einschließlich ihrer Mechanismen!) systematisch herausgearbeitet und miteinander verglichen werden, um ihre theoretische Einzigartigkeit zu erfassen.²⁰ Was sind die empirischen

¹⁷ Die Verwirrung, wie genau die vier Tests zu verstehen sind, besteht jedoch auch in der Literatur weiter und die Interpretationen von Bennett, Rohlfing, Mahoney, Beach und Pedersen und Collier unterscheiden sich im Detail.

¹⁸ Etwas komplizierter wird es, wenn im direkten Verhältnis zu alternativen Erklärungen upgedated werden soll.

¹⁹ Dies bedeutet nicht, dass „*straw-in-the-wind tests*“ wertlos seien. Allerdings muss hier die fehlende Stärke des Tests durch die Anzahl der empirischen Beobachtungen ausgeglichen werden.

²⁰ Bereits Campell empfiehlt Autoren von Fallstudien, neben Feldnotizen auch die theoretisch zu erwartenden Implikationen detailliert festzuhalten (1975, S. 186).

Erwartungen, wenn Theorie A stimmt? Aber auch: Was dürfte *nicht* zu beobachten sein? Wie überraschend sind bestimmte Beobachtungen? Gibt es einzelne Beobachtungen, mit denen klar zwischen Theorie A und B diskriminiert werden kann? Je systematischer und transparenter solche Fragen behandelt werden, desto besser. In ihrem Lehrbuch beschreiben Beach und Pedersen (2012) die verschiedenen Formen bayesianischer Schlussfolgerungen ausführlich und erläutern mit vielen Beispielen, woran man besonders starke Schlüsse erkennt und wie man sie verwendet. In den letzten Jahren hat sich innerhalb der bayesianisch orientierten Process Tracing Literatur eine überaus technische Debatte entwickelt, in der u. a. versucht wurde, qualitativen Inferenzen wieder quantifizierte Wahrscheinlichkeiten zuzuweisen (Fairfield und Charman 2017). Ob dies überhaupt möglich ist und ob der Aufwand dabei im Verhältnis zum Ertrag steht, ist zweifelhaft (Zaks 2021). In jedem Fall kann „Spurenlesen“ im Rahmen von Process Tracing heute differenziert und transparent durchgeführt werden, wenn man ausreichend spezifische Theorien und detailliertes empirisches Wissen besitzt.

4.2 Die Darstellung prozessanalytischer Befunde

Bislang wenig beachtet wurde die Aufbereitung der Ergebnisse von Process Tracing. Fallstudien werden auch in der Policy-Forschung üblicherweise in narrativer Form präsentiert. In Abgrenzung von historischen Ansätzen und dem Stereotyp des „Geschichtenerzählens“ findet sich bei manchen Methodikern allerdings ein großer Widerstand gegen narrative Forschung (Beach und Pedersen 2012; Tilly 1999). Versteht man ein Narrativ jedoch, wie in der Literaturwissenschaft üblich, als eine Form der *Darstellung* einer Geschichte (eines „Plots“), so ist dieser Widerstand nicht nachvollziehbar. Denn der Plot kann durchaus stark abstrahiert und theoriegeleitet sein (ein gutes Beispiel sind die analytic narratives von Bates et al. 1998; Levi 2002). Aufgrund der Bedeutung von Kausalmechanismen verträgt sich die Narration sogar sehr gut mit Process Tracing, denn bei beiden geht es darum, selektiv Abläufe und Ereignisse zu repräsentieren, die Zeitlichkeit und Kausalität aufweisen und sich vor dem Hintergrund eines Kontexts (dem „setting“ der Geschichte) abspielen (s. a. Abbott 2008). Daher führt kaum ein Weg um narrative Darstellungen herum. Hinweise, was eine „gut erzählte“ Fallstudie ausmacht, gibt es jedoch wenige (s. aber Büthe 2002; Ruback 2010).

5 Anwendungen in der Policy-Forschung

5.1 Theoretestendes Process Tracing: Steinmo (1989)

Das erste Beispiel für (implizites) Process Tracing in der Policy-Forschung ist Sven Steinmos Studie zur Steuerpolitik in Großbritannien, Schweden und den USA (Steinmo 1989). Zu Beginn unterscheidet Steinmo drei Ansätze zur Erklärung von Länderunterschieden in Höhe und Struktur der Besteuerung: „„interests, „values“

and „the state“ (1989, S. 500). Der erste Ansatz bezieht sich auf Theorien, die Parteien und Interessenorganisationen ins Zentrum stellen, der zweite auf kulturalistische Erklärungen und politische Kultur und der dritte auf (historisch-) institutionalistische Theorien. In der Folge soll getestet werden, welcher der drei Ansätze die Unterschiede zwischen den Ländern am besten erklären kann. Ausgangspunkt ist die kontraintuitive empirische Beobachtung, dass die USA traditionell stärker auf umverteilende Steuern und auf Unternehmenssteuern setzen als Schweden (Großbritannien befindet sich zwischen diesen beiden Polen). Hinzu kommt der außerordentliche Grad der Fragmentierung im US-amerikanischen Steuersystem. Steinmo stützt sich in seiner Untersuchung auf Sekundärquellen, Einnahmestatistiken und eine Reihe von Interviews mit steuerpolitischen Akteuren. Seine empirische Analyse beschreibt die allgemeinen Muster der Steuerpolitik in den drei Ländern, teilweise in einer eher stilisierten Form: pluralistische Politik in den USA, korporatistische in Schweden und Parteipolitik in Großbritannien. Diese Muster werden hauptsächlich auf historische Unterschiede in den politischen Institutionen (insbesondere Wahlsystemen und Strukturen der Interessenorganisationen) zurückgeführt. Einzelne Steuerreformen werden hingegen eher illustrativ herangezogen. Dabei vergleicht Steinmo die Präferenzen der Akteure, die Prozesse der Steuergesetzgebung und die strukturellen Outcomes mit den Erwartungen der theoretischen Ansätze. Mechanismen werden nicht abstrakt-deduktiv, sondern eher *ad hoc* formuliert. Der Autor ist sich jedoch der methodischen Grenzen seiner Studie bewusst und schreibt:

A more complete explanation would require a more fully elaborated analysis of the development of American political institutions over time and the dynamic interaction of these institutions and the large, fragmented, and expanding economy. Unfortunately, space does not allow us to delve into this discussion here (Steinmo 1989, S. 511).

Steinmos Studie beruht allerdings nicht ausschließlich auf drei unverbundenen Fallstudien, sie ist vielmehr eine Art Mixed-Methods-Design sowohl mit Elementen des Process Tracing als auch eines Small-N-Vergleichs. Nicht nur die Entwicklungen *innerhalb* der Fälle, sondern auch die Kontraste *zwischen* den Fällen spielen daher eine Rolle in der komplexen Gesamtargumentation. Sein Theoretetest zeigt, dass der institutionalistische Ansatz die strukturellen Unterschiede und typischen Entscheidungsmuster in der Steuerpolitik besser erklären kann als interessens- oder wertorientierte Ansätze alleine.

5.2 Theoriegenerierendes Process Tracing: Trampusch (2006)

Ein weiteres Beispiel für Process Tracing ist Christine Trampuschs Studie der sozialpolitischen Reformpolitik der Regierung Schröder (Trampusch 2006). Einer traditionell inkrementalistischen Sichtweise, die die vielfältigen Blockademöglichkeiten im politischen System der Bundesrepublik in den Mittelpunkt rückt und (parteipolitische) Präferenzen als weitgehend statisch und exogen gegeben sieht, stellt sie einen Ansatz gegenüber, der Präferenzen (auch) als endogen durch

Rückkoppelungseffekte der Policies beeinflusst ansieht. Der erste Ansatz – von ihr „Interaktionsanalyse“ genannt – hat Schwierigkeiten, die rot-grünen Sozialreformen zu erklären, sie erscheinen als abweichender Fall sowohl vom üblichen inkrementellen Politikmuster der Bundesrepublik als auch generell von Sozialreformen in Systemen mit vielen Vetopunkten. Die Autorin zeigt nun empirisch, dass der Reformprozess empirisch in zentralen Punkten nicht den Annahmen des Interaktionsmodells entsprach, sondern eher einem historisch-institutionalistischen Modell des endogenen Präferenzwandels. Exogene Schocks etwa, die Interaktionsanalysen üblicherweise zur Erklärung radikalen Politikwandels heranziehen, spielten zwar eine gewisse Rolle, reichen jedoch zur Erklärung des Wandels nicht aus. Vielmehr entwickelt Trampusch ein komplexes Narrativ, in dem die Interaktionsperspektive um eine sequenzorientierte Perspektive ergänzt wird, in der Rückkopplung und Präferenzänderungen eine zentrale Bedeutung haben.

Methodisch entspricht Trampuschs Untersuchung in vielerlei Hinsicht dem Process Tracing. Im Mittelpunkt stehen nicht nur Korrelationen zwischen unabhängigen und abhängigen Variablen über mehrere Fälle hinweg, sondern Prozesse innerhalb eines Falles und die Sequenz von zentralen, theoretisch bedeutsamen Ereignissen. Sie nutzt den Status der deutschen Sozialpolitik unter Rot-Grün als abweichender Fall, um zur Weiterentwicklung von existierenden Theorien des Policy-Wandels beizutragen. Ausdrücklich zielt sie dabei auf eine Modifikation, und nicht auf die Falsifikation der Interaktionstheorie. Neben Sekundäranalysen werden auch Primärquellen (z. B. Koalitionsverträge, Stellungnahmen) und Medienberichte herangezogen, um Präferenzen und Entscheidungsabläufe zu belegen. Auch diese Untersuchung weicht allerdings vom Idealmodell von Process Tracing ab, wenn es um die Spezifizierung von Kausalmechanismen geht. Statt alternative Ursachen und Outcomes mit klar ausformulierten theoretischen Mechanismen und durchgehenden Kausalketten zu verbinden, benutzt Trampusch die beiden theoretischen Perspektiven eher als Heuristiken (2006, S. 63) und untersucht, inwiefern empirische Politikprozesse eher dem einen oder anderen recht allgemeinen Policy-Modell ähneln, um schließlich zu einer Synthese der Ansätze zu gelangen.

5.3 Erklärung eines spezifischen Outcomes: Mos (2014)

Ein gutes Beispiel für historisch ausgerichtetes Process Tracing zur Erklärung eines spezifischen Outcomes ist Martijn Mos’ Analyse der Antidiskriminierungsklausel hinsichtlich sexueller Orientierung in Artikel 13 des Amsterdamer EU-Vertrag von 1997 (Mos 2014). Dieses Ereignis war insofern überraschend, als zum Zeitpunkt der Vertragsverhandlungen nur eine kleine Minderheit von Mitgliedstaaten vergleichbare Klauseln auf nationaler Ebene verwirklicht hatte und auf intergouvernementaler Ebene kein großes Interesse an einer Ausweitung des Nichtdiskriminierungsschutzes auf weitere Gruppen herrschte. Mos zeigt, dass das Europäische Parlament hier eine wichtige Rolle einnehmen konnte, was den dominanten liberal-intergouvernementalistischen und institutionalistischen Erklärungen des Amsterdamer Vertrags widerspricht. Detailliert wird nachgezeichnet, wie Mitglieder des Parlaments seit den

frühen 1980ern die Rechte Homosexueller trotz fehlender Kompetenzgrundlage auf die supranationale Agenda brachten (und dort hielten) und schließlich zusammen mit *Gay-Rights*-Aktivisten die Vertragsverhandlungen beeinflussen konnten. Er stützt sich dabei auf offizielle Dokumente, eigene Archivrecherche sowie auf 33 semistrukturierte Interviews mit Vertretern von EU-Institutionen, Mitgliedstaaten und Interessenverbänden.

Mos bezeichnet seine Analyse selbst als theoriegenerierendes Process Tracing (2014, S. 634). Im Unterschied zum theoretischen Anspruch bei Trampusch erscheint Mos' Erklärung jedoch weit stärker vom überraschenden empirischen Fall als vom theoretischen Puzzle getrieben. Er betont, seine Fallstudie

does not call for a wholesale rejection of leading theoretical perspectives, but [...] suggests that overall grand theory should coexist with an appreciation of the more contextual dynamics of specific episodes of policymaking (Mos 2014, S. 634).

Das Ergebnis der Analyse ist dann auch keine neue Theorie, sondern zwei Bedingungen, unter denen das Europäische Parlament zur Einflugschneise für Antidiskriminierungsrechte werden konnte: „agenda overload“ und spezifische „issue characteristics“. Ob diese Bedingungen auf andere Policies übertragbar sind, bleibt offen. Wie so oft ist auch hier der Unterschied zwischen den Formen von Process Tracing in der Praxis schwer festzumachen.

Allen drei Beispielen ist gemeinsam, dass die Erklärungsmechanismen nicht lehrbuchmäßig mitsamt allen Gliedern der Kausalkette ausformuliert werden. In dieser Hinsicht gleichen sie der Methode des „pattern matching“, wo auch untersucht wird, ob einzelne, teilweise relativ lose zusammenhängende empirische Aspekte zu verschiedenen Theorien passen oder nicht. Auch wenn sich letztendlich nicht alle Teile der Erklärung empirisch belegen lassen, so würde jedoch eine transparentere theoretische Spezifizierung weiterhelfen. Eine typische Schwäche von Process Tracing-Analysen ist die Ungleichbehandlung alternativer Erklärungen. Auch dies findet sich in den – ansonsten vorbildhaften – Beispielen wieder. So wird etwa die „kulturalistische“ Perspektive bei Steinmo eher stiefmütterlich behandelt. Damit wird allerdings die Chance vertan, eine der potenziellen Stärken des Process Tracing voll auszuspielen.

6 **Stärken, Schwächen und Gütekriterien von Process Tracing**

Eine Stärke von Process Tracing liegt in der Möglichkeit, Kausalschlüsse über alternative Erklärungen auch bei wenigen Fällen über den Weg der empirischen Analyse von Kausalmechanismen zu erzielen. Zweitens muss als Stärke die hohe Flexibilität hinsichtlich der Datenquellen und -formate, die einbezogen werden, genannt werden. Drittens sind die Methoden, Schlüsse auf Basis von teilweise wenigen, aber stark ins Gewicht fallenden Beobachtungen zu treffen, mittlerweile recht ausgefeilt.

Kritisch ist zum einen anzumerken, dass Theorien in der Policy-Forschung häufig nicht in dem Maß spezifiziert sind, wie das für eine starke Deduktion von empirischen Implikationen notwendig wäre. Insbesondere die theoretischen Mechanismen werden häufig nur vage skizziert, sind umstritten oder fehlen gar völlig. Argumente, dass bestimmte Faktoren – etwa bestimmte Akteure – „eine Rolle spielen“ sind nicht unüblich. Theorietests über detaillierte Mechanismen, wie sie in der Literatur zu Process Tracing meist im Vordergrund stehen, wären demnach häufig kaum möglich.

Dieser Punkt kann noch grundlegender formuliert werden. Obwohl Process Tracing in die Nähe zum wissenschaftstheoretischen Realismus gestellt wird, sind bestimmte Aspekte weiterhin positivistisch geprägt. Insbesondere die theoretestende Variante basiert auf dem hypothetisch-deduktivistischen Modell wissenschaftlicher Erkenntnis, nach dem Theorien eine Reihe relativ unproblematisch ableitbare Implikationen enthalten, die in basale empirische Beobachtungen übersetzt und so getestet werden können. Das Problem ist jedoch, dass diese Ableitung und Übersetzung selbst theoriegetrieben ist (Schindler 2013). Idealerweise müssten, bevor ein Theorietest durchgeführt werden kann, diese „Hilfstheorien“ also selbst getestet werden (und wiederum deren Hilfstheorien davor). Eher pragmatistisch geprägte Alternativen zum hypothetisch-deduktivistischen Modell werden etwa von Clarke und Primo (2012) für die Politikwissenschaft aufgezeigt (s. a. Friedrichs und Kratochwil 2009). Eine weitere Alternative bieten „abduktive“ Ansätze in der Sozialwissenschaft (Reichertz 2013). Deren Ideal vom Spurenlesen und iterativen Verfeinern von Erklärungen anhand von einzelnen oder wenigen Fällen kommt der Praxis von Process Tracing sehr nahe, ohne „testen“ zu wollen (s. Fairfield und Charman 2019). Immerhin betonen auch Bennett und Checkel, dass Process Tracing meist eine Mischung aus Deduktion und Induktion ist.

Schließlich ist ein ungelöstes Problem die Frage der Generalisierung von prozessanalytischen Erkenntnissen. Quantitative Forschung kann sich in ihrem Anspruch der Generalisierung von einer Stichprobe auf die Population auf die klassische Wahrscheinlichkeitstheorie berufen. Dieses klare Fundament fehlt dem Process Tracing (noch). Auf dem Kontinuum zwischen reiner Einzelbetrachtung und der Erforschung von universellen Gesetzmäßigkeiten beansprucht Process Tracing den Bereich der „kontingenten Generalisierung“ (George und Bennett 2005; s. a. Mitchell 2002). George und Bennett verstehen darunter eine Art von Generalisierung, die Kontextbedingungen mit einbezieht und nur innerhalb einer vergleichbaren Gruppe von Fällen (z. B. ausgebauten Wohlfahrtsstaaten, ethnisch gemischte Gesellschaften, ressourcenreiche Autokratien) verallgemeinert. Die Vergleichbarkeit des Kontexts kann aus ähnlichen historischen Entwicklungen, kulturellen oder geografischen Gegebenheiten abgeleitet werden. Es ist jedoch unklar, was dies genau bedeutet. Welches sind die Kriterien, die uns angeben, ob und inwiefern unsere Schlüsse auch externe Validität behaupten können? Ein alternatives Konzept ist die „generalization by mechanism“ (Bengtsson und Herting 2014; Ylikoski 2019), die davon ausgeht, dass unter bestimmten Bedingungen Erkenntnisse über Kausalmechanismen aus einem Kontext auf einen anderen generalisiert werden können.

Ungeachtet dieser offenen wissenschaftstheoretischen Fragen ist Process Tracing nicht gleich Process Tracing und die Debatte um angemessene methodische Gütekriterien ist in vollem Gange. Bennett und Checkel etwa stellen 10 Kriterien für gute Process Tracing auf (s. Kasten, s. a. die Checkliste im Appendix von Beach und Pedersen 2012), die u. a. den Aspekt der alternativen Erklärungen betonen. Dies ist sinnvoll, denn ein Risiko für die Validität von Process Tracing ergibt sich aus psychologischen Effekten, wie dem Bestätigungsbiass oder spezifischen Problemen bayesianischer Inferenz, wie der „base rate fallacy“ (Bar-Hillel 1980), die dazu führt, dass die „Gewissheit“ und „Trennschärfe“ von Evidenz falsch eingeschätzt wird. Hier kann es helfen, alternative Erklärungen bei der empirischen Analyse stets systematisch und gleichberechtigt mit einzubeziehen.

Kasten 1: Gütekriterien für Process Tracing

- 1) Cast the net widely for alternative explanations
- 2) Be equally tough on the alternative explanations
- 3) Consider the potential biases of evidentiary sources
- 4) Take into account whether the case is most or least likely for alternative explanations
- 5) Make a justifiable decision on when to start
- 6) Be relentless in gathering diverse and relevant evidence, but make a justifiable decision on when to stop
- 7) Combine process tracing with case comparisons when useful for the research goal and feasible
- 8) Be open to inductive insights
- 9) Use deduction to ask „if my explanation is true, what will be the specific process leading to the outcome?“
- 10) Remember that conclusive process tracing is good, but not all good process tracing is conclusive

Quelle: Bennett und Checkel (2015b)

7 Fazit

Process Tracing zeichnet sich durch eine bestimmte Herangehensweise an Theorien, Daten und Schlussfolgerungen aus, die in diesem Beitrag vorgestellt wurde. Die Methode eignet sich für viele Fragen der Policy-Forschung. Merkmale wie Fallorientierung, Kausalerklärung, Prozesscharakter, Akteurzentrierung, Theoriepluralismus und Datenpluralismus sind gute Gründe für die traditionell starke Affinität von Policy-Forschung und Process Tracing. Gleichzeitig ist die Policy-Forschung methodisch eine pluralistische Subdisziplin. Daher sollte die Ausdifferenzierung der Methodendiskussion nicht dazu führen, dass die Anknüpfungspunkte zwischen den methodischen Ansätzen aus dem Blick geraten (s. zu interpretativen Methoden etwa Nullmeier 2019). Ein stärkerer Dialog mit der quantitativen Forschung, QCA-Anwenderinnen und der interpretativ-qualitativen Forschung nicht nur über wissenschaftstheoretische Grundlagen, sondern über die tatsächliche Forschungspraxis ist

wichtiger denn je. Denn wenn es etwa um Validitätsprobleme von Interviewdaten geht oder Probleme der Fallauswahl, sind die Gräben zwischen den verschiedenen Camps oft weit weniger tief als vermutet. Zudem besteht die Erwartung, dass Mixed-Methods-Designs die Stärken der verschiedenen Methoden komplementär verknüpfen können. Auch hier kann Process Tracing mit der Betonung von Kausalmechanismen und detailliertem Fallwissen eine wichtige Rolle erfüllen.

Literatur

- Abbott, H. Porter. 2008. *The Cambridge introduction to narrative*. Cambridge: Cambridge University Press.
- Bar-Hillel, Maya. 1980. The base-rate fallacy in probability judgments. *Acta Psychologica* 44(3): 211–233.
- Bates, Robert, Avner Greif, Margaret Levi, Jean-Laurent Rosenthal, und Barry R. Weingast, Hrsg. 1998. *Analytic narratives*. Princeton: Princeton University Press.
- Beach, Derek, und Rasmus B. Pedersen. 2012. *Process-tracing methods: Foundations and guidelines*. Ann Arbor: University of Michigan Press.
- Beach, Derek, und Rasmus Brun Pedersen. 2016. *Causal case study methods: Foundations and guidelines for comparing, matching, and tracing*. Ann Arbor: University of Michigan Press.
- Beck, Friedrich, und Eckart Henning, Hrsg. 2012. *Die archivalischen Quellen. Mit einer Einführung in die Historischen Hilfswissenschaften*, 5. Aufl. Wien/Köln/Weimar: Böhlau.
- Bengtsson, Bo, und Nils Herting. 2014. Generalization by mechanism: Thin rationality and ideal-type analysis in case study research. *Philosophy of the social sciences* 44(6): 707–732.
- Bennett, Andrew. 2008a. Building communities, bridging gaps: Alexander George's contributions to research methods. *Political Psychology* 29(4): 489–507.
- Bennett, Andrew. 2008b. Process tracing: A Bayesian perspective. In *The Oxford handbook of political methodology*, Hrsg. Janet M. Box-Steffensmeier, Henry E. Brady und David Collier, 702–721. Oxford: Oxford University Press.
- Bennett, Andrew. 2015. Appendix: Disciplining our conjectures. Systematizing process tracing with Bayesian analysis. In *Process tracing: From metaphor to analytic tool*, Hrsg. Andrew Bennett und Jeffrey T. Checkel, 276–298. Cambridge: Cambridge University Press.
- Bennett, Andrew, und Jeffrey T. Checkel, Hrsg. 2015a. *Process tracing. From metaphor to analytic tool*. Cambridge: Cambridge University Press.
- Bennett, Andrew, und Jeffrey T. Checkel. 2015b. Process tracing: From philosophical roots to best practices. In *Process tracing. From metaphor to analytic tool*, Hrsg. Andrew Bennett und Jeffrey T. Checkel, 3–37. Cambridge: Cambridge University Press.
- Bogner, Alexander, Beate Littig, und Wolfgang Menz, Hrsg. 2002. *Das Experteninterview: Theorie, Methode, Anwendung*. Opladen/Wiesbaden: Leske + Budrich/Springer.
- Brenner-Wilczek, Sabine, Gertrude Capl-Kaufmann, und Max Plassmann. 2006. *Einführung in die moderne Archivarbeit*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Büthe, Tim. 2002. Taking temporality seriously: Modeling history and the use of narratives as evidence. *American Political Science Review* 96(3): 481–493.
- Campbell, Donald T. 1975. „Degrees of freedom“ and the case study. *Comparative Political Studies* 8(2): 178–193.
- Checkel, Jeffrey T. 2008. Process tracing. In *Qualitative methods in international relations. A pluralist guide*, Hrsg. Audie Klotz und Deepa Prakash, 114–127. Heidelberg/Berlin: Springer.
- Clarke, Kevin A., und David M. Primo. 2012. *A model discipline: Political science and the logic of representations*. Oxford: Oxford University Press.
- Collier, David. 2011. Understanding process tracing. *PS: Political Science & Politics* 44(4): 823–830.

- Collier, David, Henry E. Brady, und Jason Seawright. 2010. Sources of leverage in causal inference: Toward an alternative view of methodology. In *Rethinking social inquiry*, Hrsg. Henry E. Brady und David Collier, 2. Aufl., 161–199. Lanham: Rowman & Littlefield.
- Dexter, Lewis A. 2006 [1970]. *Elite and specialized interviewing, ECPR classics*. Colchester: ECPR Press.
- Dobson, Miriam, und Benjamin Ziemann, Hrsg. 2009. *Reading primary sources: The interpretation of texts from nineteenth- and twentieth-century history*. London: Routledge.
- Eckstein, Harry. 1975. Case study and theory in political science. In *Handbook of political science*, Hrsg. Fred I. Greenstein und Nelson W. Polsby, 79–137. Reading: Addison-Wesley.
- Einhorn, Hillel J., Don N. Kleinmuntz, und Benjamin Kleinmuntz. 1979. Linear regression and process-tracing models of judgment. *Psychological Review* 86(5): 465.
- Elman, C., und M. F. Elman, Hrsg. 2001. *Bridges and boundaries: Historians, political scientists, and the study of international relations*. Cambridge, MA: MIT Press.
- Fairfield, Tasha, und Andrew Charman. 2019. A dialogue with the data: The Bayesian foundations of iterative research in qualitative social science. *Perspectives on Politics* 17(1): 154–167.
- Fairfield, Tasha, und Andrew E. Charman. 2017. Explicit Bayesian analysis for process tracing: Guidelines, opportunities, and caveats. *Political Analysis* 25(3): 363–380.
- Falleti, Tulia G., und Julia F. Lynch. 2009. Context and causal mechanisms in political analysis. *Comparative Political Studies* 42(9): 1143–1166.
- Ford, J. Kevin, Neal Schmitt, Susan L. Schechtman, Brian M. Hults, und Mary L. Doherty. 1989. Process tracing methods – Contributions, problems, and neglected research questions. *Organizational Behavior and Human Decision Processes* 43(1): 75–117.
- Friedrichs, Jörg, und Friedrich Kratochwil. 2009. On acting and knowing: How pragmatism can advance international relations research and methodology. *International Organization* 63(4): 701–731.
- Frisch, Scott A., Douglas B. Harris, Sean Q. Kelly, und David C. W. Parker. 2012. *Doing archival research in political science*. New York: Cambria Press.
- George, Alexander L. 1979. Case studies and theory development: The method of structured, focused comparison. In *Diplomacy: New approaches in history, theory and policy*, Hrsg. G. Lauren, 43–68. New York: The Free Press.
- George, Alexander L., und Andrew Bennett. 2005. *Case studies and theory development in the social sciences*. Cambridge, MA: MIT Press.
- George, Alexander L., und Timothy J. McKeown. 1985. Case studies and theories of organizational decision-making. *Advances in Information Processing in Organizations* 2:21–58.
- Gerring, John. 2006a. *Case study research: Principles and practices*. Cambridge: Cambridge University Press.
- Gerring, John. 2006b. Single-outcome studies: A methodological primer. *International Sociology* 21(5): 707–734.
- Gerring, John. 2007a. Is there a (viable) crucial-case method? *Comparative Political Studies* 40(3): 231–253.
- Gerring, John. 2007b. The mechanistic worldview: Thinking inside the box. *British Journal of Political Science* 38(1): 161–179.
- Gläser, Jochen, und Grit Laudel. 2010. *Experteninterviews und qualitative Inhaltsanalyse*, 4. Aufl. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Gläser, Jochen, und Grit Laudel. 2013. Life with and without coding: Two methods for early-stage data analysis in qualitative research aiming at causal explanations. *Forum Qualitative Sozialforschung/Forum Qualitative Social Research* 14(2). <https://doi.org/10.17169/fqs-14.2.1886>.
- Glennan, Stuart, und Phyllis Illari, Hrsg. 2018. *The Routledge handbook of mechanisms and mechanical philosophy*. London: Routledge.
- Goldberg, Felix, und Achim Hildebrandt. 2020. Experteninterviews. In *Handbuch Methoden der Politikwissenschaft*, Hrsg. Claudius Wagemann, Achim Goerres und Markus B. Siewert, 267–284. Berlin/Heidelberg: Springer.

- Grzymala-Busse, Anna. 2011. Time will tell? Temporality and the analysis of causal mechanisms and processes. *Comparative Political Studies* 44(9): 1267–1297.
- Hall, Peter A. 2006. Systematic process analysis: When and how to use it. *European Management Review* 3(1): 24–31.
- Harvey, William S. 2011. Strategies for conducting elite interviews. *Qualitative Research* 11(4): 431–441.
- Hempel, Carl Gustav, und Paul Oppenheim. 1948. Studies in the logic of explanation. *Philosophy of Science* 15(2): 135–175.
- Hogarth, Robin M. 1974. Process tracing in clinical judgment. *Behavioral Science* 19(5): 298–313.
- King, Gary, Robert O. Keohane, und Sidney Verba. 1994. *Designing social inquiry: Scientific inference in qualitative research*. Princeton: Princeton University Press.
- Kühberger, Anton, Michael Schulte-Mecklenbeck, und Rob Ranyard. 2011. Introduction: Windows for understanding the mind. In *A handbook of process tracing methods for decision research: A critical review and user's guide*, Hrsg. Michael Schulte-Mecklenbeck, Anton Kühberger und Rob Ranyard, 1–18. New York: Psychology Press.
- Lebow, Richard N. 2010. *Forbidden fruit: Counterfactuals and international relations*. Princeton: Princeton University Press.
- Leech, B. L. 2002. Asking questions: Techniques for semistructured interviews. *PS: Political Science & Politics* 35(4): 665–668.
- Levi, Margaret. 2002. Modeling complex historical processes with analytic narratives. In *Akteure – Mechanismen – Modelle. Zur Theoriefähigkeit makro-sozialer Analysen*, Hrsg. Renate Mayntz, 108–127. Frankfurt a. M./New York: Campus.
- Levi-Faur, David. 2006. Varieties of regulatory capitalism: Getting the most out of the comparative method. *Governance* 19(3): 367–382.
- Lieberman, Evan S. 2005. Nested analysis as a mixed-method strategy for comparative research. *American Political Science Review* 99(3): 435–452.
- Lieberman, Evan S., Marc Morje Howard, und Julia Lynch. 2004. Symposium: Field research. *Qualitative Methods: Newsletter of the American Political Science Association Organized Section on Qualitative Methods* 2(1): 2–15.
- Lieberson, Stanley. 1991. Small N's and big conclusions: An examination of the reasoning in comparative studies based on a small number of cases. *Social Forces* 70(2): 307–320.
- Lijphart, Arend. 1971. Comparative politics and the comparative method. *American Political Science Review* 65(3): 682–693.
- Lustick, Ian S. 1996. History, historiography, and political science: Multiple historical records and the problem of selection bias. *American Political Science Review* 90(3): 605–618.
- Lynch, Julia F. 2013. Aligning sampling strategies with analytic goals. In *Interview research in political science*, Hrsg. Layna Mosley, 31–44. Ithaca: Cornell University Press.
- Machamer, Peter, Lindley Darden, und Carl F. Craver. 2000. Thinking about mechanisms. *Philosophy of Science* 67(1): 1–25.
- Mahoney, James. 2003. Strategies of causal assessment in comparative historical analysis. In *Comparative historical analysis in the social sciences*, Hrsg. James Mahoney und Dietrich Rueschemeyer, 337–372. Cambridge: Cambridge University Press.
- Mahoney, James. 2012. The logic of process tracing tests in the social sciences. *Sociological Methods & Research* 41(4): 570–597.
- Mahoney, James, und Gary Goertz. 2004. The possibility principle: Choosing negative cases in comparative research. *American Political Science Review* 98(4): 653–669.
- Mahoney, James, und Claudio M. Villegas. 2007. Historical enquiry and comparative politics. In *The Oxford handbook of comparative politics*, Hrsg. Carles Boix und Susan Stokes, 73–89. Oxford: Oxford University Press.
- Maxwell, Joseph A. 2010. Using numbers in qualitative research. *Qualitative Inquiry* 16(6): 475–482.

- Mayntz, Renate. 2002. Zur Theoriefähigkeit makro-sozialer Analysen. In *Akteure – Mechanismen – Modelle. Zur Theoriefähigkeit makro-sozialer Analysen*, Hrsg. Renate Mayntz, 7–43. Frankfurt/New York: Campus.
- Mill, John Stuart. 1968 [1872]. *System der deduktiven und induktiven Logik: Eine Darlegung der Grundsätze der Beweislehre und der Methoden wissenschaftlicher Forschung, Gesammelte Werke, übers. und hrsg. von Theodor Gomperz*, Bd. 1. 2., dt. Aufl., Neudr. der Ausg. Leipzig, 1884 ed. Aalen: Scientia.
- Mitchell, Sandra D. 2002. Contingent generalizations: Lessons from biology. In *Akteure – Mechanismen – Modelle: Zur Theoriefähigkeit makro-sozialer Analysen*, Hrsg. Mayntz, 179–195. Frankfurt a. M./New York: Campus.
- Morgan, Kimberly J. 2016. Process tracing and the causal identification revolution. *New Political Economy* 21(5): 489–492.
- Mos, Martijn. 2014. Of gay rights and Christmas ornaments: The political history of sexual orientation non-discrimination in the treaty of Amsterdam. *JCMS: Journal of Common Market Studies* 52(3): 632–649.
- Mosley, Layna, Hrsg. 2013. *Interview research in political science*. Ithaca: Cornell University Press.
- Müller, Harald. 2002. Antinomien des demokratischen Friedens. *Politische Vierteljahresschrift* 43(1): 46–81.
- Norman, Ludvig. 2015. Interpretive process tracing and causal explanations. *APSA Qualitative and Multi-method Analysis Newsletter* 13(2): 4–9.
- Nullmeier, Frank. 2019. Interpretative Politikforschung und kausale Mechanismen. *Zeitschrift für Politikwissenschaft* 29(2): 153–171.
- Nullmeier, Frank. 2021. *Kausale Mechanismen und Process Tracing: Perspektiven der qualitativen Politikforschung*. Frankfurt/New York: Campus.
- Panke, Diana. 2012. Process tracing: Testing multiple hypotheses with a small number of cases. In *Research design in European studies: Establishing causality in Europeanization*, Hrsg. Theofanis Exadaktylos und Claudio M. Radaelli, 125–140. Basingstoke: Palgrave Macmillan.
- Pierson, Paul. 1994. *Dismantling the Welfare State? Reagan, Thatcher, and the Politics of Retrenchment*. Cambridge, MA: Cambridge University Press.
- Przeworski, Adam, und Henry Teune. 1970. *The logic of comparative social inquiry*. New York: Wiley-Interscience.
- Psillos, Stathis. 1999. *Scientific realism: How science tracks truth*. London: Routledge.
- Ragin, Charles C. 1992. „Casing“ and the process of social inquiry. In *What is a case? Exploring the foundations of social inquiry*, Hrsg. Charles C. Ragin und Howard S. Becker, 217–226. Cambridge: Cambridge University Press.
- Rathbun, Brian C. 2008. Interviewing and qualitative field methods: Pragmatism and practicalities. In *The Oxford handbook of political methodology*, Hrsg. Janet M. Box-Steffensmeier, Henry E. Brady und David Collier, 685–701. Oxford: Oxford University Press.
- Reichertz, Jo. 2013. *Die Abduktion in der qualitativen Sozialforschung. Über die Entdeckung des Neuen*, 2. Aufl. Wiesbaden: Springer VS.
- Ricks, Jacob I., und Amy H. Liu. 2018. Process-tracing research designs: A practical guide. *PS: Political Science & Politics* 51(4): 842–846.
- Rohlfing, Ingo. 2012. *Case studies and causal inference: An integrative framework*. Hounds mills/Basingstoke: Palgrave Macmillan.
- Rohlfing, Ingo, und Peter Starke. 2013. Building on solid ground: Robust case selection in multi-method research. *Swiss Political Science Review* 19(4): 492–512.
- Ruback, Timothy J. 2010. ‘Let me tell the story straight on.’ *Middlemarch*, Process-tracing methods, and the politics of narrative. *British Journal of Politics and International Relations* 12(4): 477–497.
- Rubin, Herbert J., und Irene S. Rubin. 2011. *Qualitative interviewing: The art of hearing data*. London/New York: Sage.

- Rueschemeyer, Dietrich, und John D. Stephens. 1997. Comparing historical sequences – A powerful tool for causal analysis. *Comparative Social Research* 16(1): 55–72.
- Russett, Bruce. 1994. *Grasping the democratic peace: Principles for a post-Cold War world*. Princeton: Princeton University Press.
- Sagan, Scott. 1995. *The limits of safety: Organizations, accidents, and nuclear weapons*. Princeton: Princeton University Press.
- Schindler, Samuel. 2013. Observation and theory-ladenness. In *Encyclopedia of philosophy and the social sciences*, Hrsg. Byron Kaldus, 695–698. London: Sage.
- Schwartz-Shea, Peregrine. 2006. Judging quality: Evaluative criteria and epistemic communities. In *Interpretation and method: Empirical research methods and the interpretive turn*, Hrsg. Dwora Yanow und Peregrine Schwartz-Shea, 89–113. Armonk/London: M.E. Sharpe.
- Seawright, Jason. 2016. The case for selecting cases that are deviant or extreme on the independent variable. *Sociological Methods & Research* 45(3): 493–525.
- Seawright, Jason, und John Gerring. 2008. Case selection techniques in case study research: A menu of qualitative and quantitative options. *Political Research Quarterly* 61(2): 294–308.
- Skocpol, Theda. 1979. *States and social revolutions: A comparative analysis of France, Russia, and China*. Cambridge: Cambridge University Press.
- Skocpol, Theda. 1984. Emerging agendas and recurrent strategies in historical sociology. In *Vision and method in historical sociology*, Hrsg. Theda Skocpol, 356–391. Cambridge: Cambridge University Press.
- Steinmo, Sven. 1989. Political institutions and tax policy in the United States, Sweden, and Britain. *World Politics* 41(4): 500–535.
- Swedberg, Richard. 2011. Theorizing in sociology and social science: Turning to the context of discovery. *Theory and Society* 41(1): 1–40.
- Tansey, Oisin. 2007. Process tracing and elite interviewing: A case for non-probability sampling. *PS: Political Science & Politics* 40(4): 765–772.
- Thies, Cameron. 2002. A pragmatic guide to qualitative historical analysis in the study of international relations. *International Studies Perspectives* 3(3): 351–372.
- Tilly, Charles. 1999. The trouble with stories. In *The social worlds of higher education. Handbook for teaching in a new century*, Hrsg. Ronald Aminzade und Bernice Pescosolido, 256–270. Thousand Oaks: Pine Forge Press.
- Trampusch, Christine. 2006. Sequenzorientierte Policy-Analyse. *Berliner Journal für Soziologie* 16(1): 55–76.
- Trampusch, Christine, und Bruno Palier. 2016. Between X and Y: How process tracing contributes to opening the black box of causality. *New political economy* 21(5): 437–454.
- Van Evera, Stephen. 1997. *Guide to methods for students of political science*. Ithaca: Cornell University Press.
- Vaughan, Diane. 2009. *The Challenger launch decision: Risky technology, culture, and deviance at NASA*. Chicago: University of Chicago Press.
- Windelband, Wilhelm. 1894. Geschichte und Naturwissenschaft. Rede zum Antritt des Rektorats der Kaiser-Wilhelms-Universität Straßburg, gehalten am 01.05.1894. Sitzungsberichte der Heidelberger Akademie der Wissenschaften, Philosophisch-Historische Klasse; Jg. 1910, Abh. 14. <https://doi.org/10.11588/diglit.20767>. Zugegriffen am 06.04.2022.
- Yanow, Dwora. 2006. Introduction. In *Interpretation and method: Empirical research methods and the interpretive turn*, Hrsg. Dwora Yanow und Peregrine Schwartz-Shea, xi–xxvi. Armonk/London: M.E. Sharpe.
- Ylikoski, Petri. 2019. Mechanism-based theorizing and generalization from case studies. *Studies in History and Philosophy of Science Part A* 78:14–22.
- Zaks, Sherry. 2017. Relationships among rivals (RAR) a framework for analyzing contending hypotheses in process tracing. *Political Analysis* 25(3): 344–362.
- Zaks, Sherry. 2021. Updating Bayesian(s): A critical evaluation of Bayesian process tracing. *Political Analysis* 29(1): 58–74.