

Forschungsdesign in der Politikwissenschaft – Kausale Perspektiven versus kontrastive Theorietests

Steffen Ganghof^{1,*}

¹ Lehrstuhl für vergleichende Politikwissenschaft, Universität Potsdam, Germany

* E-Mail: ganghof@uni-potsdam.de

Zusammenfassung

Die politikwissenschaftliche Literatur unterscheidet zwei Grundtypen von Forschungsdesigns: x- und y-zentriert. Dieser Beitrag argumentiert, dass ein „kontrastives“ Forschungsdesign als dritter Grundtyp abgegrenzt werden sollte. Die drei Designs unterscheiden sich durch die Anzahl der betrachteten Theorien und dadurch, ob mehrere Theorien konkurrierend oder komplementär sind. Die typologische Abgrenzung des kontrastiven Designs verdeutlicht auch die Vor- und Nachteile x- und y-zentrierter Designs. Anhand verschiedener Beispielstudien (experimentell und nicht-experimentell, quantitativ und qualitativ) werden die Charakteristika der drei Designs sowie ihre Kombinationsmöglichkeiten herausgearbeitet. Darüber hinaus wird das kontrastive Design als verbindendes Element zwischen den quantitativen und qualitativen Forschungs-„Kulturen“ hervorgehoben.

Schlüsselwörter

Forschungsdesign, kausale Perspektiven, Theorietest, x-zentriert, y-zentriert, Experimente

Research Design in Political Science – Causal perspectives versus contrastive theory testing

Abstract

The political science literature distinguishes two basic types of research designs: x- and y-centered. The article argues for the distinction of a third basic type: the “contrastive” design. The three designs differ in the number of relevant theories and in whether they see theories as competing or complementary. The typological differentiation of the contrastive research design helps to clarify the pros and cons of x- und y-centered designs. The article uses exemplary studies (experimental and observational, quantitative and qualitative) to illustrate the characteristics of the three designs as well as the possibilities of combining them. The contrastive design also constitutes a common element of the quantitative and qualitative research „cultures”.

Keywords

research design, causal perspectives, theory test, effects of causes, causes of effects, experiments

Danksagung:

Ich danke Sebastian Eppner, Katja Heeß, Ulrich Kohler, Alexander Pörschke, Stefan Schukraft, Kai Schulze und den anonymen Gutachtern der OZP für hilfreiche Diskussionen und Kommentare. Für verbliebene Unzulänglichkeiten bin ich verantwortlich.

The author has declared that no competing interests exist.

1. Einleitung

Die politikwissenschaftliche Methodenliteratur arbeitet seit längerem mit einer grundlegenden Unterscheidung zwischen x- und y-zentrierten Forschungsdesigns oder „Kausalperspektiven“ (Ganghof 2005; Gschwend/Schimmelfennig 2007; Wenzelburger et al. 2014, 73; Schnapp 2015; Kubbe 2016, 44). *X-zentrierte Designs* sind auf eine bestimmte erklärende Variable fokussiert; sie wollen den kausalen Effekt schätzen, den diese Variable auf eine bestimmte abhängige Variable ausübt. *Y-zentrierte Designs* sind hingegen auf ein zu erklärendes Phänomen oder Ereignis fokussiert; auf der Suche nach befriedigenden Erklärungen kombinieren sie mehrere theoretische Ansätze und aus diesen abgeleitete Erklärungsvariablen.

Ich plädiere im Folgenden für die Abgrenzung eines dritten, ebenso grundlegenden Designs, in dem es um den vergleichenden Test konkurrierender Theorien geht. Dieser Typ wird im Folgenden als *kontrastives Forschungsdesign* bezeichnet. Diese Terminologie knüpft an die wissenschaftstheoretische Position des „kontrastiven Empirismus“ (Sober 1990) an, der zufolge „the goal of science is to bring observations to bear on the comparison of theories“ (Sober 2013, 167). Die drei Designs lassen sich als Idealtypen anhand von drei Kriterien unterscheiden: (1) der Anzahl der betrachteten Theorien, (2) dem logischen Verhältnis dieser Theorien zueinander und (3) der Art der verwendeten empirischen Evidenz.

Die präzise typologische Abgrenzung der drei Forschungsdesigns schärft meines Erachtens den Blick für ihre jeweiligen Stärken und Schwächen. X-zentrierte Designs sind problematisch, wenn es konkurrierende theoretische Spezifikationen des zu untersuchenden kausalen Effekts gibt. Kontrastive Designs sind umgekehrt problematisch, wenn sich die vermeintlich konkurrierenden Theorien bei genauerer Betrachtung gegenseitig ergänzen. Y-zentrierte Designs sind insofern problematisch, als dass die zielgerichtete Kombination unterschiedlicher Theorien nur bedingt einen *Test* dieser Theorien darstellt.

Natürlich wird der Vergleich von Theorien und Erklärungen seit langem diskutiert. Die bestehende methodische Literatur ist jedoch durch zwei Defizite gekennzeichnet. Erstens findet ein Großteil der relevanten Diskussion nur *innerhalb* der „kulturell“ (Goertz/Mahoney 2012) abgeschotteten Lager quantitativer und qualitativer Forschung statt. Im ersten Lager werden Aspekte des Theorievergleichs zum Beispiel in Bezug auf spezifische statistische Tests diskutiert (Clarke 2007b; Imai/Tingley 2012), im zweiten unter einer schwer überschaubaren Vielzahl von Bezeichnungen wie zum Beispiel „pattern matching“ (Campbell 1975, 182; Lange 2013, 53). Eine übergreifende Diskussion über die Bedeutung und die Anforderungen des Theorievergleichs wird dadurch erschwert; sie existiert bisher kaum. Dieser Punkt leitet über zum zweiten Defizit.

Der Wert der Unterscheidung zwischen x- und y-zentrierten Designs liegt unter anderem darin, dass sie quer zur qualitativ/quantitativ-Unterscheidung liegt (Ganghof 2005, 93). Sie hat dadurch einen Lager-übergreifenden Methodendiskurs ermöglicht (Gschwend/Schimmelfennig 2007). In diesem Diskurs wird indes das kontrastive Design noch nicht hinreichend präzisiert und gewürdigt. Eine verbreitete Sicht ist, dass eine Grundentscheidung zwischen X- und Y-Zentrierung getroffen werden muss, aber in beiden Designs „stets mögliche alternative Erklärungen [zu] berücksichtigen und kontrollieren“ sind (Gschwend/Schimmelfennig 2007, 19–21; vgl. auch Dür 2007, 285). Diese Forderung ist meines Erachtens zu vage, da sie nicht zwischen konkurrierenden (d.h. logisch widersprüchlichen) und komplementären Theorien unterscheidet. Ich argumentiere im Folgenden, dass ein Test konkurrierender Theorien weder in idealtypisch x-zentrierten noch idealtypisch y-zentrierten Designs eine Rolle spielt. In x-zentrierten Designs liegt der Fokus auf einer Theorie; andere Theorien werden nur in der Form von Kontrollvariablen einbezogen und müssen daher komplementär sein. Y-zentrierte Designs wollen Theorien kombinieren, so dass diese ebenfalls komplementär sein müssen. Allein kontrastive Designs fokussieren auf den Vergleich *mehrerer konkurrierender Theorien* und stellen insofern einen eigenen Grundtyp dar.

Der nächste Abschnitt entwickelt die Typologie. Die folgenden Abschnitte erläutern die Charakteristika sowie Vor- und Nachteile der drei Designs anhand beispielhafter Studien. Kontrastive Designs werden zunächst von x-zentrierten Designs abgegrenzt (Abschnitt 3) und dann von y-zentrierten (Abschnitt 4). Dabei werden jeweils auch kombinierte Designs thematisiert. Der letzte Abschnitt gibt einen Ausblick.

2. Drei idealtypische Forschungsdesigns

Wir müssen zunächst die Reichweite der zu entwickelnden Typologie klären. Sie behandelt nur den grundlegenden Fokus des Forschungsdesigns in Bezug auf die relevanten Theorien und die relevante empirische Evidenz. Sie umfasst explizit nicht unterschiedliche philosophische Konzeptionen von Kausalität. Diese Einschränkung ist aus folgendem Grund wichtig. Die Unterscheidung zwischen x- und y-zentrierten Designs wird häufig als deckungsgleich zur englischsprachigen Unterscheidung zwischen „causes-of-effects“ und „effects-of-causes“ behandelt (z.B. Kubbe 2016, 44). Letztere wird aber nicht nur von unterschiedlichen Autoren unterschiedlich definiert, sie umfasst häufig auch weit mehr als das Forschungsdesign (vgl. Morton/Williams 2010; Goertz/Mahoney

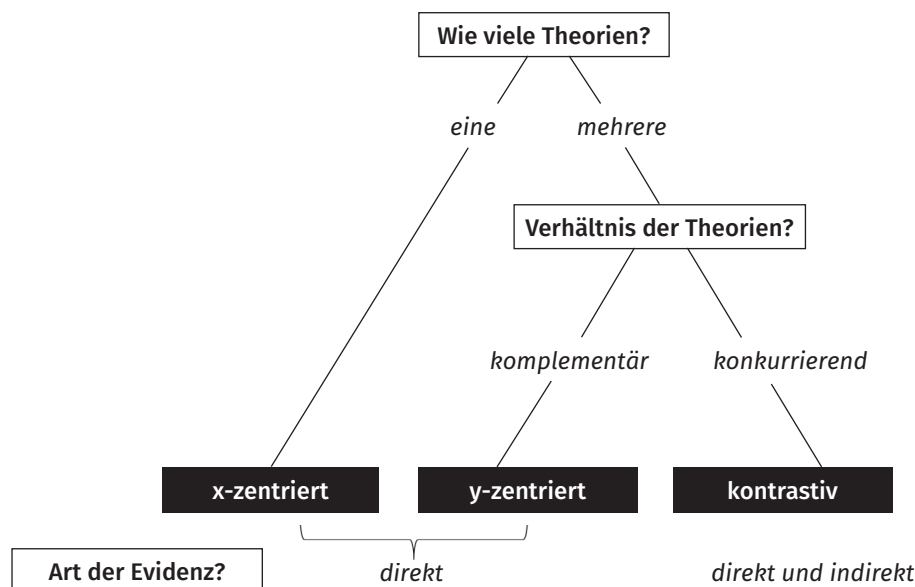
2012). So schreiben etwa Kittel und Morton (2012, 2): „In the effects-of-causes perspective, based in the Rubin causal model, a causal effect is observed if the difference in causal outcomes that can be attributed to the manipulation of one particular factor is statistically significant (King et al. 1994, 76–85).“ In diesem Verständnis kausaler Perspektiven geht es also auch um eine philosophische Konzeption von Kausalität. Die im Folgenden entwickelte Typologie zielt dagegen nur auf das grundlegende Forschungsdesign ab (vgl. auch Ganghof 2005), welches meines Erachtens mit unterschiedlichen Kausalitätskonzeptionen verbunden werden kann. Für die Zwecke dieses Beitrags beschränke ich mich auf Beispielstudien, die mit der von Kittel und Morton zitierten Standarddefinition kausaler Effekte im Einklang stehen (King et al. 1994, 76). Das Ziel ist es zu zeigen, dass sich auch auf der Basis dieser Definition drei grundlegend unterschiedliche Forschungsdesigns unterscheiden lassen.¹

Die vorgeschlagene Typologie ist in Abbildung 1 dargestellt. Sie schlägt eine Abgrenzung der drei Designs auf der Basis dreier Fragen vor. Die erste Frage lautet, wie viele Theorien im Fokus der Untersuchung stehen: eine oder mehrere. Mit „Theorie“ ist hier nicht unbedingt ein sehr umfassendes System von Annahmen gemeint. Schon die Annahme über den Kausalmechanismus, der zwei Variablen miteinander verknüpft, oder auch nur

über die Richtung der Kausalität zwischen zwei Variablen, ist eine Theorie im hier unterstellten Sinne (King et al. 1994, 19).

Haben wir es mit mehreren Theorien zu tun, lautet die zweite Frage, ob diese Theorien konkurrierend oder komplementär sind. Sie sind konkurrierend, wenn sie logisch inkonsistente Annahmen über die relevanten (kausalen) Mechanismen treffen, und somit nicht gleichzeitig „bestätigt“ werden können. Konkurrierende Theorien können also gar nicht in kohärenter Weise kombiniert werden. Komplementäre Theorien betonen hingegen unterschiedliche (kausale) Mechanismen, die auch gleichzeitig wirken können. Diese Theorien sind kombinierbar. Ein Beispiel für konkurrierende Theorien – das wir später wieder aufnehmen – ist an dieser Stelle hilfreich. Nehmen wir an, wie wollen den kausalen Effekt der Agenda-Setzer-Macht der Regierung auf Politikergebnisse untersuchen (Becher 2010). Das Problem ist, dass dieser kausale Effekt *theoretisch unterschiedlich spezifiziert werden kann*. Wir wissen nämlich nicht genau, wer innerhalb der Regierung die Agenda-Macht ausübt: die Premierministerin, die Finanzministerin oder die zuständige Fachministerin. Wenn wir annehmen, dass sich diese drei Spezifikationen der Agenda-Macht logisch ausschließen, haben wir es mit konkurrierenden Theorien zu tun.

Abbildung 1: Drei grundlegende Forschungsdesigns



Quelle: Eigene Darstellung.

¹ Ich denke, dass sich die vorgeschlagene Typologie auch auf andere Kausalitätskonzeptionen anwenden lässt, zum Beispiel solche, die in der „Qualitative Comparative Analysis“ Verwendung finden (Goertz/Mahoney 2012, Schneider/Wagemann 2012). Ich kann dies hier aus Platzgründen nicht näher begründen.

Da dieselben Theorien unterschiedlich spezifiziert werden können, ist es in der Politikwissenschaft nicht immer eindeutig ob zwei Theorien konkurrierend oder komplementär sind. Die Typologie hebt deshalb hervor, dass mit der Spezifikation eines Forschungsdesigns eine *Entscheidung* darüber verbunden ist, wie das Verhältnis der relevanten Theorien zueinander verstanden wird. Diese Entscheidung sollte bewusst getroffen und klar kommuniziert werden.

Die dritte Frage der Typologie ist für die Abgrenzung der Forschungsdesigns streng genommen nicht mehr notwendig, sie vertieft aber unser Verständnis: Welche Art von Evidenz wird für die Unterstützung einer Theorie angeführt? Ich unterscheide mit dem Wissenschaftstheoretiker Reiss (2015, 347) zwischen *direkter* und *indirekter* Evidenz. Eine Theorie wird direkt durch Daten gestützt, wenn das Datenmuster so ist, wie man es auf der Basis der Theorie erwarten kann. Sie wird im Gegensatz dazu indirekt gestützt, wenn das Datenmuster nicht so ist, wie man es auf der Basis einer konkurrierenden Theorie erwarten kann.

Betrachten wir nur eine Theorie, dann kann indirekte Evidenz definitionsgemäß keine Rolle spielen. Dasselbe gilt, wenn unser Ziel in der Kombination *komplementärer* Theorien besteht. In idealtypischen x- und y-zentrierten Designs ist deshalb nur direkte Evidenz von Bedeutung. Anders in kontrastiven Designs: hier ist auch indirekte Evidenz bedeutsam. Es kann zum Beispiel sein, dass zwei Theorien bestimmte Beobachtungen *gleich gut* erklären können. Wir hätten dann zwar direkte Evidenz für beide Theorien, aber keine indirekte Evidenz, auf Basis derer wir eine Theorie gegenüber der anderen vorziehen können. Das ideale Szenario in einem kontrastiven Design ist natürlich, dass nur eine von mehreren Theorien die gesammelten Daten erklären kann. Diese Theorie erhielte dann direkte Unterstützung (aufgrund ihrer eigenen Erklärungskraft) und indirekte Unterstützung (aufgrund des Scheiterns der konkurrierenden Theorien). Zumindest sollte eine Theorie die Evidenz besser erklären können als die Konkurrenz. Klar ist deshalb auch, dass indirekte Evidenz *allein* niemals ausreichend ist: Wenn eine Theorie nichts erklärt, dann wird sie natürlich auch nicht dadurch gestützt, dass eine andere Theorie nichts erklärt.

Geht man von der Standardkonzeption kausaler Effekte aus (King et al. 1994), lässt sich der Unterschied zwischen den drei Designs folgendermaßen zusammenfassen:

1. X-Zentrierte Designs versuchen einen theoretisch spezifizierten kausalen Effekt zu isolieren (experimentell oder durch Kontrollvariablen) und seine Größe möglichst genau und ohne Verzerrungen zu schätzen.
2. Kontrastive Designs sind dann relevant, wenn es für einen kausalen Effekt widersprüchliche theoretische

Spezifizierungen gibt. Das Ziel ist es, die Beste der vorliegenden Spezifizierungen zu bestimmen.

3. Y-zentrierte Designs untersuchen, wie mehrere komplementäre Theorien über kausale Effekte kombiniert werden können um bestimmte Ereignisse oder Varianzen zu erklären.

Im Folgenden möchte ich anhand von Beispielarbeiten zeigen, dass wir durch diese Abgrenzung der drei Idealtypen die Unterschiede sowie Vor- und Nachteile von Forschungsdesigns besser in den Blick bekommen.

3. X-zentrierte versus kontrastive Forschungsdesigns

Plädoyers für kontrastive Designs sind manchmal mit einer expliziten Kritik an x-zentrierten Designs verbunden. Letztere können nämlich zu einem logischen Fehlschluss verleiten: Wenn wir aus einer Theorie eine Hypothese ableiten und direkte Evidenz für diese Hypothese finden, folgt daraus *keine* Bestätigung der Theorie. Der Grund ist, dass es viele andere Theorien geben kann, die genauso gut oder besser mit der vorhandenen Evidenz im Einklang sind. Viele Methodiker haben daraus gefolgert, dass logisch valide Hypothesentests vergleichend oder „kontrastiv“ sein müssen: sie müssen eine Theorie gegen mindestens eine konkrete Alternative testen (Sober 1990; Clarke 2007a). Aus dieser Sicht kann eine Theorie niemals absolut bestätigt werden. Sie kann nur besser bestätigt sein als eine genau bestimmte Menge bestehender Konkurrenztheorien.

Konkurrenztheorien müssen strikt von Kontrollvariablen unterschieden werden. Konkurrierende Theorien treffen logisch inkompatible Annahmen darüber, was in der Welt passiert. Die Theorien hinter Kontrollvariablen modellieren dagegen unterschiedliche Ausschnitte der Welt. Diese Theorien sind in x-zentrierten Designs selbst von keinerlei Interesse (Sieberer 2007, 263). Sie werden lediglich einbezogen, um eine verzerrte Schätzung des interessierenden kausalen Effekts zu vermeiden (King et al. 1994). Die konkurrierenden Theorien in kontrastiven Designs sind hingegen von zentralem Interesse.²

3.1 Beispielstudien

Der Unterschied zwischen den beiden Designs wird am klarsten, wenn man dieselbe Theorie betrachtet. Die Vetospielertheorie ist ein gutes Beispiel (Tsebelis 2002; Ganghof 2003). Sie ist unter anderem eine Theorie darüber, wie Koalitionsregierungen in parlamentarischen

² Statistische Tests arbeiten häufig mit einer Nullhypothese, die aber nicht unbedingt für eine konkurrierende Theorie steht. Dies kann allerdings der Fall sein, wie etwa bei Martin und Vanberg (2014).

Regierungssystemen zu Entscheidungen kommen. Eine ihrer zentralen theoretischen Annahmen ist, dass jede Koalitionspartei bei jeder Entscheidungsfrage ein Vetorecht besitzt, unabhängig von ihrer Größe oder ihres Zugriffs auf ein bestimmtes Ministerium. Es gibt zwei bekannte Konkurrenztheorien. Der Theorie der Ministerialregierung (Laver/Shepsle 1996) zufolge bestimmt die Regierungspartei, die ein Ministerium kontrolliert, die Politik im jeweiligen Geschäftsbereich allein; es gibt keine Vetospieler. Und gemäß der Theorie des „gewichteten Kompromisses“ hat zwar jede Partei Einfluss auf die Regierungspolitik, aber dieser ist proportional zum Anteil der Mandate, die von ihr in die Regierung eingebracht werden. Größere Parteien haben also, anders als in der Vetospielertheorie, einen größeren Einfluss (zum Beispiel Martin/Vanberg 2014).

Eine x-zentrierte Beispielstudie ist Tsebelis (1999). Der Autor leitet aus seiner Theorie die Hypothese ab, dass die ideologische Distanz zwischen Vetospielern einen negativen Effekt auf die Anzahl bedeutsamer Gesetze hat. Diese Hypothese wird anhand eines Datensatzes über arbeitsrechtliche Gesetze in europäischen Demokratien zwischen 1981 und 1991 getestet. Die empirische Korrelation zwischen den gemessenen Vetospieler-Distanzen und der Anzahl der Arbeitsgesetze liefert *direkte* Evidenz für die Vetospielertheorie. Konkurrierende Theorien über die Entscheidungsfindung in Regierungen spielen in der Studie dagegen keine Rolle. Tsebelis (1999, 603) bezieht zwar einige kurz diskutierte Kontrollvariablen ein (für die er keine Evidenz findet), diese stehen jedoch in keinem theoretischen Widerspruch zu seiner Theorie.

Eine kontrastive Beispielstudie ist Boockmann (2006). Der Autor analysiert die Ratifizierung von Konventionen der internationalen Arbeitsorganisation (ILO) in 17 Demokratien zwischen 1960 und 1996. Er nimmt an, dass es innerstaatliche Akteure gibt, die für die Ratifizierungsentscheidung besonders kritisch sind. Er betrachtet zwei konkurrierende Theorien darüber, wer die kritischen Akteure sind: die Vetospielertheorie und die Theorie der Ministerialregierung. Boockmann schätzt deshalb zwei getrennte statistische Modelle, die sich nur durch die Modellierung des politischen Ratifizierungsprozesses unterscheiden. Im ersten wird die inhaltliche Position der Regierung durch den inhaltlich entferntesten Vetospieler bestimmt, im zweiten durch die Position des Arbeitsministers. Boockmann benutzt dann den statistischen Vergleichstest von Clarke (2007b), um die Erklärungskraft der beiden Modelle zu vergleichen. Die Theorie der Ministerialregierung erklärt die Daten besser als die Vetospielertheorie (Boockmann 2006, 170).³

Folgt man streng der kontrastiven Logik, so sollte sie den Forschungsprozess von Beginn an anleiten – insbesondere schon bei der Auswahl oder der Erzeugung von Daten. Denn die Vorhersagen von konkurrierenden Theorien überlappen sich häufig in großem Maße. Anders formuliert: Auch konkurrierende Theorien sind in weiten Teilen „beobachtungsäquivalent“. Es sollte deshalb genau spezifiziert werden, in Bezug auf welche Beobachtungen tatsächlich konkurrierende Vorhersagen vorliegen. Dieser Punkt lässt sich besonders gut anhand von experimentellen Studien illustrieren, in denen die relevanten Daten im Forschungsprozess selbst erzeugt werden. Ein berühmtes Beispiel aus der Biologie ist die Arbeit von Meselson und Stahl (1958). Die Autoren testeten drei widersprüchliche Theorien über die genauen Mechanismen der Vervielfältigung des Erbguts. Dazu manipulieren sie *Escherichia coli*-Bakterien auf eine Weise, die einen kritischen Test der drei Theorien darstellt. Das Experiment bestätigt die Theorie der semi-konservativen Replikation und verwirft ihre Konkurrentinnen.

Ähnliche politikwissenschaftliche Studien finden sich zum Beispiel in der Literatur zum themenspezifischen Wählerverhalten. Hier stehen sich (vermeintlich) widersprüchliche Theorien gegenüber, zum Beispiel die Theorien der Nähewahl (Downs 1957) und der Richtungswahl (Rabinowitz/Macdonald 1989). Diese und andere Theorien sind schon häufig mit Beobachtungsdaten getestet worden. Durch experimentell erzeugte Daten lassen sich indes gezielter „kritische Tests“ durchführen, die eindeutig zwischen den konkurrierenden Theorien unterscheiden. Ein gutes Beispiel ist die Studie von Tomz und Van Houweling (2008). Die Autoren bestimmen exakt die spezifischen Bedingungen, unter denen die konkurrierenden Theorien zu unterschiedlichen Vorhersagen führen, und erzeugen diese Bedingungen dann gezielt durch das experimentelle Design.

3.2 Methoden des Theorievergleichs

Bei kontrastiven Designs mit Beobachtungsdaten lassen sich nicht immer „kritische Tests“ erzeugen. Dann stellt sich die Frage, wie die relative Erklärungskraft der konkurrierenden Theorien bestimmt wird. Während hier nicht im Detail auf statistischen Debatten eingegangen werden kann, lassen sich zwei einflussreiche Ansätze unterscheiden.

Der erste ist sehr umstritten. Er besteht darin, unterschiedliche Theorien in einem statistischen „Supermodell“ zusammenzufassen, das alle erklärenden Variablen enthält, die aus den konkurrierenden Theorien abgeleitet werden können (Clarke 2001, 730; vgl. Greene 2012, 175). Ein gutes Beispiel ist die Studie von Volden und Carrubba (2004, 533), in der es um die Erklärung von „übergroßen“ Koalitionen in 24 Demokratien zwischen 1955 und 1998 geht. Die theoretische Anlage der

³ Für weitere kontrastive Tests der Vetospielertheorie, siehe Ganghof und Schulze (2015) sowie Ganghof (2016).

Studie ist kontrastiv und fokussiert auf fünf Theorien. Einige Theorien werden von bis zu fünf Variablen repräsentiert, andere von nur einer Variablen. Ein Beispiel für letztere ist die Theorie von Axelrod (1970), der zufolge übergroße Koalitionen gebildet werden, wenn die „überflüssigen“ Parteien keinen Einfluss auf die ideologische Spannweite des Kabinetts haben. Die Autoren schätzen zunächst fünf getrennte Modelle für jede Theorie und fassen sie dann in einem Supermodell zusammen. Auf der Basis dieses Supermodells kommen sie zum Beispiel zu dem Schluss, dass es für Axelrods Theorie „no support“ (Volden/Carrubba 2004, 533) gibt, obwohl dies in dem getrennten Modell durchaus der Fall war.

Diese Form des quantitativen Theorievergleichs ist derselben Kritik ausgesetzt wie y-zentrierte Studien: die Spezifikation des Supermodells ist „atheoretical“ (Clarke 2001, 731). Denn wenn die Modellspezifikation theoretisch angeleitet sein sollte, dann bräuchten wir auch eine „Supertheorie“, die erklärt, wie die verschiedenen Theorien zusammenpassen. Die Möglichkeit einer solchen Supertheorie widerspricht aber der notwendigen Ausgangsannahme kontrastiver Designs: *konkurrierende Theorien*.

Der zweite Ansatz ist, Modelle auf der Basis der einzelnen Theorien zu spezifizieren und dann ihre Erklärungskraft anhand statistischer Maßzahlen zu vergleichen. Dies war die Vorgehensweise von Boockmann, der den Vergleichstest von Clarke (2007) benutzt hat. Dieselbe Methodik wendet Eppner (2014, Kapitel 6) auf die Fragestellung von Volden und Carrubba (2004) an. Dabei verwendet er den Friedman-Test, der es erlaubt, mehr als zwei Theorien zu vergleichen (Clarke 2011). Ein weiterer bekannter Test ist der *J*-Test (Davidson/Mackinnon 1981), der auch für die vergleichende Evaluation der Vetospielertheorie verwendet wurde (Franzese 2002; Bräuninger 2005).

Die methodische Diskussion über den systematischen Vergleich von Theorien steht aber in der quantitativen Politikwissenschaft noch eher am Anfang (Imai/Tingley 2012). Und auch neuere Studien in führenden Zeitschriften greifen weiterhin auf „Supermodelle“ zurück – zumindest *zusätzlich* zu statistischen Vergleichstests. Ein Beispiel ist die Studie von Martin und Vanberg (2014). Die Forschungsfrage der Autoren lautet, wessen Präferenzen eigentlich in den Politik-Entscheidungen von Koalitionsregierungen repräsentiert werden. Ihr Forschungsdesign ist beispielhaft kontrastiv. Anhand von Gesetzgebungsdaten aus Dänemark, Deutschland und den Niederlanden werden drei Theorien gegeneinander getestet: die bereits erwähnten Theorien der Ministerialregierung und des gewichteten Kompromisses sowie eine Theorie, die auf dem Medianwähler-Theorem basiert (Black 1948). Die Autoren verwenden dabei sowohl den Vergleichstest von Clarke als auch ein kombiniertes Supermodell. Beide Methoden identifizieren die

Theorie des gewichteten Kompromisses als die beste der drei Theorien.

3.3 Wann welches Design?

Wann sollten wir welches Design wählen? Die Antwort ergibt sich im Prinzip aus der Typologie. Wenn die theoretische Spezifikation eines kausalen Effekts relativ unumstritten ist, bietet sich ein x-zentriertes Design zur Schätzung seiner Größe an. Ist diese Spezifikation indes kontrovers und gibt es mehrere klar definierte Konkurrenten, so mag ein kontrastives Design die bessere Wahl sein.

Zu beachten ist allerdings, dass Forschung ein kollektiver Prozess ist. Nicht jede Studie kann alles auf einmal leisten. Darüber hinaus ist direkte Evidenz ebenfalls entscheidend. Zwischen zwei Theorien den Sieger zu ermitteln, bringt uns relativ wenig, wenn beide die Daten nur schlecht erklären können. Es kann deshalb durchaus sinnvoll sein, sich zur Stützung einer kontroversen Theorie *zunächst* auf direkte Evidenz zu konzentrieren. In diesem Sinne hat die oben angesprochene x-zentrierte Studie von Tsebelis (1999) dazu beigetragen seine Theorie als erklärungskräftig zu etablieren, worauf nachfolgende Vergleichstests aufbauen konnten (Ganghof 2016).

Dennoch sollte eine kritische Rezeption x-zentrierter Tests immer fragen, ob es nicht plausible Konkurrenztheorien gibt, welche die relevanten Daten genauso gut oder besser erklären können. Ein gutes Beispiel ist die quantitative Studie von Vanberg (2005). Sie testet eine Theorie des Verhaltens von Verfassungsrichtern, die der Autor entwickelt hat. Manow und Ganghof (2005, 26ff.) halten die X-Zentrierung dieses Tests für vorbildlich. Sie wird jedoch problematisch, wenn wir annehmen, dass es konkurrierende Theorien gibt, die Vanbergs statistische Befunde ähnlich gut oder besser erklären können. Genau dies behauptet Hüller (2014). Folgt man ihm, könnte man einerseits Vanbergs Design kritisieren. Andererseits kann man dessen x-zentrierte Studie aber auch als fruchtbaren Impuls verstehen, im nächsten Schritt die vermeintlich konkurrierende Theorie systematisch auszuformulieren und explizit gegen Vanbergs zu testen.

3.4 Kombinierte Designs

Natürlich können x-zentrierte und kontrastive Designs auf unterschiedliche Weisen kombiniert werden. Ein Beispiel ist eine Studie von Becher (2010) zur Arbeitsmarktpolitik von 20 OECD Ländern zwischen 1973 und 2000. Der Autor argumentiert, dass die Vetospielertheorie und die Theorie der Ministerialregierung durch entsprechende Modifikationen *komplementär gemacht* werden können. Die Vetospielertheorie hat nämlich keine Antwort auf die Frage, wer innerhalb der Regierung der Agenda-Setzer ist (Tsebelis 2002, 3). Becher nimmt an,

dass dies die jeweiligen Fachminister sind. Von der Vetospielertheorie übernimmt er aber die Überlegung, dass der Handlungsspielraum des Agenda-Setzers durch die ideologische Heterogenität der Vetospieler begrenzt wird. Der Autor entwickelt also eine kleine „Supertheorie“ zur Kombination zweier vermeintlich konkurrierender Theorien. Seine zentrale Kausalhypothese lautet, dass der ideologische Einfluss der Minister-Partei von der ideologischen Distanz zwischen den Vetospielern abhängig ist.

Das übergeordnete Design der Studie ist somit x-zentriert, fokussiert auf die Frage: „Do veto players constrain ministerial discretion?“ (Becher 2010, 45). Es gibt indes auch zwei kontrastive Elemente. Erstens will Becher nicht nur seine kombinierte Theorie durch direkte Evidenz stützen. Er will auch zeigen, dass sie die Daten *besser* erklärt als die ursprüngliche Ministerialregierungs-Theorie ohne Berücksichtigung von Vetospielern (Becher 2010, 47). Zweitens gesteht Becher explizit ein, dass seine theoretischen Annahmen über die Verteilung von Agenda-Macht in der Regierung kontrovers sind: diese Macht könnte auch beim Regierungschef oder Finanzminister statt beim Arbeitsminister liegen. Becher testet seine Theorie deshalb auch im Vergleich zu den konkurrierenden Theorien. Dieser Vergleich führt zu keinem eindeutigen Sieger, wenngleich Becher (2010, 54) zumindest in Bezug auf eine seiner beiden abhängigen Variablen gewisse Vorteile für seine Theorie sieht.

4. Y-zentrierte versus kontrastive Forschungsdesigns

Der Unterschied dieser beider Designs betrifft die Frage, ob mehrere Theorien in einer bestimmten Studie als konkurrierend oder komplementär betrachtet werden (Abbildung 1). Als y-zentriert werden also hier solche Designs bezeichnet, die mehrere komplementäre Theorien kombinieren.

Es gibt zwei alternative Definitionen y-zentrierter Designs, die zur Klarheit hier kurz abgegrenzt werden sollen. Die erste versteht diese Designs als Versuche Phänomene „so genau und vollständig wie möglich zu erklären“ (Gschwend/Schimmelfennig 2007, 21). Diese Definition mag als idealtypische Zuspitzung plausibel sein, allerdings ist es schwer Forscher zu finden, die tatsächlich an die Möglichkeit vollständiger Erklärungen glauben (Ganghof 2005, 84; Sieberer 2007, 278). Die zweite Alternativdefinition y-zentrierter Designs lautet, dass es bei ihnen um die Untersuchung mehrerer theoretischer Implikationen eines einzigen umfassenden theoretischen Modells geht (vgl. Kittel/Morton 2012, 2-3). Im Gegensatz zu diesen beiden Definitionen liegt der Fokus hier auf der *Kombination komplementärer, aber unterschiedlicher Theorien*.

Meines Erachtens hilft uns die hier vorgeschlagene Definition y-zentrierter Designs deren Rechtfertigung besser zu verstehen. Das Plädoyer für solche Designs kann nämlich mit einer expliziten Kritik kontrastiver Designs verbunden werden. Diese lautet, dass Theorien häufig als konkurrierend stilisiert werden, obwohl sie auch als komplementär verstanden und sinnvoll kombiniert werden können. Interessanterweise wird diese Kritik am klarsten von prominenten Protagonisten vergleichender Theorietests formuliert:

Choosing between explanations takes up an enormous amount of researcher's time and pages in journals. It is not always necessary, and both time and pages could be put to more productive use. There are no precise rules for determining whether to engage in explanation comparison, but we have to ask ourselves, "What exactly is to be gained from comparatively testing these explanations? Instead of taking for granted that one explanation must always be the "winner," we should learn all that we can from the explanations we have. (Clarke/Primo 2012, 166)

Der letzte Satz formuliert die Essenz der y-zentrierten „Philosophie“. Er macht deutlich, dass ein idealtypisch y-zentriertes Design *kein* kontrastives Element enthält. Y-zentrierte Studien sind nicht wie Wettkämpfe, bei denen es nur einen Sieger geben kann, sondern wie Choreografien, in denen jeder eine Rolle zu spielen hat. Ein gutes Beispiel für einen y-zentrierten Forschungsansatz ist meines Erachtens die „Heidelberger Schule“ der vergleichenden Policy-Forschung. Empirische Analysen dieser Schule streben eine „möglichst gute Erklärung von Staatstätigkeit“ an; und aus diesem Grund kombinieren sie „zumeist mehrere Theorieansätze miteinander und ergänzen sie durch politikfeldspezifische Einflussfaktoren“ (Wenzelburger/Neumann 2015, 255).

Die Metapher der Choreographie macht allerdings auch den Nachteil y-zentrierter Studien deutlich. Kontrastive Designs können als echte *Tests* verstanden werden, da sie zumindest unter den verfügbaren Theorien *die beste* ermitteln. Die flexible Kombination von Theorien zur befriedigenden Erklärung eines Phänomens stellt indes nur sehr bedingt einen Test dieser Theorien dar (vgl. Ganghof 2005, 84). Reine y-zentrierte Studien haben entweder einen *explorativen*, hypothesen-generierenden Charakter (Sieberer 2007, 277) oder sie können als *Anwendungen* politikwissenschaftlichen Wissens verstanden werden. Wenn ein Sachverständiger einen Verkehrsunfall rekonstruiert, muss er dabei auch verschiedene physikalische Modelle und Gesetzmäßigkeiten kombinieren – er testet diese dabei aber nicht. Y-zentrierte Studien können Anwendungen in demselben Sinne sein, wobei der Test der verwendeten theoretischen Modelle andernorts erfolgen muss.

Genau hier liegt freilich auch der Grund, warum der Unterschied zwischen kontrastiven und y-zentrierten Designs häufig verwischt wird, gerade in qualitativen Studien. Denn durch die Einbeziehung kontrastiver Elemente kann eine y-zentrierte Erklärung zumindest teilweise getestet werden. Es findet dann aber bereits eine Kombination idealtypischer Forschungsdesigns statt.

4.1 Beispielstudien

Ich diskutiere zunächst zwei quantitative Studien, bei denen die oben genannte Kritik an kontrastiven Designs deutlich wird. Als erstes Beispiel kann erneut die experimentelle Arbeit von Tomz und van Houweling (2008) dienen. Diese stellt abschließend fest: „Our statistical model *assumes* that each voter follows *one of the three* canonical issue voting rules. An alternative perspective holds that all voters employ a common, mixed decision rule“ (ebd., 314, Hervorhebung durch mich, S.G.). Es gibt noch eine dritte Perspektive, der die Autoren auf der Basis ihrer Evidenz letztlich zuneigen: die vermeintlich konkurrierenden Theorien seien insofern komplementär, als dass *unterschiedliche Wählergruppen unterschiedlichen Entscheidungsregeln* folgen. Die zu Beginn kontrastiv angelegte Studie nimmt somit am Ende gleichsam eine y-zentrierte Wendung.

Eine ähnliche Wendung zeigt sich bei Franzese (2002, 2010) über mehrere Studien hinweg. Diese Studien basieren auf Beobachtungsdaten zur Finanzpolitik in entwickelten Demokratien. In einer früheren Arbeit hatte Franzese (2002, 156) einen kontrastiven statistischen Test der Vetospielertheorie durchgeführt. Die wichtigste Konkurrenztheorie war dabei die oben erwähnte Theorie des gewichteten Kompromisses. Er kam auf der Basis von *J*-Tests zu dem Schluss, dass die Vetospielertheorie die Daten besser erklärt als die Kompromisstheorie. In einer neueren Studie verändert der Autor jedoch sein Verständnis der beiden Theorien. Franzese (2010, 353) argumentiert nun, dass eine „theoretische Synthese“ von drei Theorien – also eine Art Supertheorie – möglich ist: von den beiden genannten sowie der Theorie kollektiven Handelns. Das Forschungsdesign verschiebt sich dadurch von kontrastiv zu y-zentriert. Der Autor entwickelt ein komplexes statistisches Modell, in das alle drei theoretischen Perspektiven einfließen. Er findet auch für alle drei Perspektiven direkte empirische Evidenz. So wie bei Tomz und van Houweling (2008) unterschiedliche Theorien für unterschiedliche Wählergruppen Unterstützung erhalten, so werden sie bei Franzese (2010) für unterschiedliche Aspekte der Finanzpolitik gestützt.

Auch Franzese strebt keineswegs nach einer vollständigen Erklärung finanzpolitischer Länderunterschiede, sondern nur nach der Kombination unterschiedlicher theoretischer Ansätze. Den fehlenden Theorievergleich

hebt er dabei explizit als Begrenzung seiner Studie hervor. Zum Beispiel:

[T]he empirical specification *assumes* that the partisanship reflected in fiscal policy is a weighted average of the preferences of government members, indicative of some bargained compromise. It does not *test* the empirical efficacy of that assumed bargaining structure against alternatives (Franzese 2010, 362, Hervorhebung im Original).

Als drittes, qualitatives Beispiel möchte ich die Studie von Schnietz (2000) und ihre Kategorisierung durch Dür (2007) diskutieren. Dür trennt nicht klar zwischen kontrastiven und y-zentrierten Designs. Er definiert letztere im qualitativen Kontext als Versuch „bestimmte Ereignisse zu erklären“ (Dür 2007, 285). Diese Definition lässt aber offen, ob unterschiedliche Theorien bei der Erklärung dieser Ereignisse konkurrieren oder kombiniert werden sollen. Der Autor entscheidet sich für ersteres Verständnis und deshalb implizit für die Diskussion kontrastiver Designs. Seine Diskussion basiert auf folgender Annahme: „Die Erklärung eines Ereignisses oder einer Reihe von Ereignissen ist jedoch nur dann überzeugend, wenn eine Studie die Überlegenheit einer Erklärung über alle anderen zeigen kann“ (ebd.).

Als Beitrag über kontrastive Designs sind Dürs Ausführungen überaus lehrreich. Er zeigt, welche hypothesentestende Kraft auch qualitative Studien haben können, wenn sie kontrastiv angelegt sind und sich auf Datenmuster konzentrieren, für die konkurrierenden Theorien unterschiedliche „Vorhersagen“ treffen. Allerdings blendet Dür (2007, 292) die kritischen Fragen an ein kontrastives Design explizit aus: Sind die diskutierten Theorien wirklich konkurrierend? Warum können sie nicht kombiniert werden? Was haben wir von der Ermittlung eines „Siegere“? Mehr noch: Seine Diskussion tendiert an einigen Stellen dazu, Studien in ein „kontrastives“ Schema zu pressen, obwohl sie eher y-zentriert sind.

Zumindest bei der Studie von Schnietz (2000) ist dies der Fall. In dieser geht es um unterschiedliche theoretische Erklärungen eines wichtigen Gesetzes der US-amerikanischen Handelspolitik: des *Reciprocal Trade Agreements Act* (RTAA) von 1934. Eine dieser Theorien erklärt das Gesetz als Resultat eines Lernprozesses. Dür (2007, 298) sieht diese Theorie in Konkurrenz zu anderen Theorien und schreibt auch Schnietz eine entsprechende Sicht zu:

Karen E. Schnietz (2000: 420) *widerlegte* diese und ähnliche Erklärungen, die die amerikanische Handelsliberalisierung seit 1934 als Resultat eines Lernprozesses sehen, jedoch in überzeugender Manier. ... Schnietz (2000) selbst schlägt eine *alternative Er-*

klärung für amerikanische Handelsliberalisierung vor, der zufolge die demokratische Partei den RTAA verwendete, um niedrigere Zölle langfristig festzuschreiben (Dür 2007, 298–9, Hervorhebung hinzugefügt).

Dür (2007, 299) selbst sieht wiederum Inkonsistenzen in der Erklärung von Schnietz. Auf den ersten Blick scheint die gesamte Diskussion also als „Wettbewerb“ zwischen konkurrierenden Theorien gut beschrieben. Bei genauerer Betrachtung entspricht dies aber gar nicht der Sicht von Schnietz selbst. Diese ist *y*-zentriert:

While it is *true* that Democrats hoped to stimulate the economy with the RTAA, as the lesson hypothesis asserts, and that Roosevelt provided strong policy leadership, as the crisis explanation posits, these explanations for the RTAA are *incomplete*. ... This article argues that the RTAA was an attempt to finally provide durability to Democrat's preferred low-tariff policy after seven decades of Republican tariff policy dominance (Schnietz 2000, 438, Hervorhebung hinzugefügt).

Schnietz sieht die drei Theorien eindeutig als komplementär an. Aspekte des Lernens, der Krisendynamik und der Parteipolitik können sinnvoll kombiniert werden. Die Argumentationsstrategie der Autorin hat freilich auch kontrastive Elemente. Aber wie das Zitat deutlich macht, zeigt Schnietz die Erklärungslücken anderer „Theorien“ nicht auf, um diese zu *widerlegen*, sondern um zu zeigen, dass sie *ergänzt* werden sollten.

4.2 Kombinierte Designs

Die Studie von Schnietz (2000) steht somit beispielhaft für eine Möglichkeit, *y*-zentrierte und kontrastive Elemente zu verbinden: letztere sollen zeigen, dass bestimmte Theorien der Ergänzung bedürfen. Bei dieser Kombinationsstrategie ist jedoch ein wichtiger Punkt zu beachten. Wenn wir *y*-zentrierte Studien über das Ideal „vollständiger“ Erklärungen definieren, dann erscheint die Ergänzung einer Erklärung durch weitere Theorien fast immer angebracht – frei nach dem Motto: je mehr, desto besser. Verwerfen wir aber dieses Ideal als unerreichbar, dann müssen wir bei der Hinzufügung weiterer Theorien immer zwischen Kosten und Nutzen abwägen. Vor allem sollten wir das Sparsamkeitsprinzip berücksichtigen, welches auch als „Ockhams Rasiermesser“ bezeichnet wird (Sober 2015). Nach diesem Prinzip sollten wir generell einfachere gegenüber komplizierteren Erklärungen vorziehen. Statistische Tests des Theorievergleichs berücksichtigen dieses Prinzip in der Regel, während qualitative Forscher es nach meinem Eindruck häufiger vernachlässigen (vgl. Ganghof 2016). Das Spar-

samkeitsprinzip ist deshalb eines der Themen, zu denen ein Lager-übergreifender Methodendiskurs sehr fruchtbar sein könnte.

In qualitativen Studien ist noch eine andere Art der Kombination von *y*-zentrierten und kontrastiven Designs weit verbreitet (Ganghof 2005, 83–85). Dabei setzen Forscher verschiedene Theorieansätze *y*-zentriert zu einer historischen Erklärung zusammen. Sie machen dabei zwar keinen Versuch, diese Gesamterklärung mit einer konkurrierenden Erklärung zu vergleichen, aber einzelne *Komponenten* der Gesamterklärung – zum Beispiel theoretische Annahmen über die Präferenzen oder Weltansichten bestimmter Akteure – werden dadurch gerechtfertigt, dass sie die beobachteten Daten besser erklären als die relevanten Konkurrenten.

Auch in quantitativen Studien können *y*-zentrierte und kontrastive Elemente verbunden werden. Dabei sollte allerdings der Versuchung widerstanden werden, durch die erwähnten statistischen „Supermodelle“ *gleichzeitig* bestimmte Theorien als nicht erklärungskräftig auszusortieren *und* dann aus den vermeintlich erklärungskräftigen Theorien eine kombinierte Erklärung zu bauen. Zumindest ist es genau dieses Stöbern in Datenmustern, gegen das sich die Vertreter *x*-zentrierter und kontrastiver Designs gleichermaßen wenden. Was aber durchaus sinnvoll sein kann, ist die explizite Entwicklung einer kombinierten „Supertheorie“ – zum Beispiel Bechers (2010) Synthese aus Vetospieler- und Ministerialregierungs-Theorie – und das kontrastive Testen dieser Theorie gegen eine Konkurrentin wie etwa die Theorie des gewichteten Kompromisses.

5. Schluss

Der Beitrag unterscheidet drei idealtypischen Forschungsdesigns: *x*-zentriert, kontrastiv und *y*-zentriert. Die präzise Abgrenzung dieser drei Designs führt zu einem tieferen Verständnis ihrer jeweiligen Vor- und Nachteile. Zum einen erlaubt uns die typologische Einbeziehung kontrastiver Designs einen klareren Blick auf die potentiellen Probleme *x*-zentrierter Designs. Diese sind vor allem problematisch, wenn die theoretische Spezifizierung der relevanten kausalen Effekte kontrovers ist – wenn es also plausible *konkurrierende* Theorien gibt. Zum anderen verdeutlicht der Vergleich mit dem kontrastiven Design auch die Spezifika und Vorteile des *y*-zentrierten Designs. Dieses betont die Komplementarität von Theorien und versucht mehrere relevante Theorien in befriedigende Erklärungen einzubringen. Die *y*-zentrierte Perspektive impliziert insofern auch eine potentielle Kritik an kontrastiven Designs, nämlich dann, wenn die vermeintlich konkurrierenden Theorien bei genauerer Betrachtung komplementär sind.

Da die drei Designs unterschiedliche Zielsetzungen verfolgen, können sie flexibel kombiniert werden. Dabei ist insbesondere der zentrale Nachteil rein y-zentrierter Designs zu beachten. Diese können nur sehr bedingt als *Tests* von Theorien oder Erklärungen verstanden werden. Sollen die in y-zentrierten Designs konstruierten Erklärungen – oder einzelne Komponenten davon – auch getestet werden, ist die Hinzufügung anderer Designelemente ratsam. An anderer Stelle habe ich die Kombination von X- und Y-Zentrierung behandelt (Ganghof 2005), hier lag der Fokus auf der Kombination y-zentrierter und kontrastiver Elemente.

Die Diskussion war bewusst an der Forschungspraxis orientiert und hat methodologische Großdebatten vermieden. Als Ausblick möchte ich jedoch den Bezug zu einer dieser Debatten herstellen: der zwischen King et al. (1994) und Goertz/Mahoney (2012). Erstere argumentieren, dass es eine einheitliche, *x-zentrierte* Logik jeder sozialwissenschaftlichen Forschung gibt, der auch qualitative Studien folgen sollten. Letztere insistieren dagegen darauf, dass quantitative und qualitative Forschung zwei unterschiedliche „Kulturen“ darstellten und in letzterer *y-zentrierte* Fragen zentral seien. Aus der hier entwickelten Perspektive ergeben sich zwei Anmerkungen zu dieser Kontroverse. Erstens haben wir gesehen, dass ein y-zentrierter Ansatz auch in der quantitativen Forschung nach wie vor verbreitet ist. Zweitens scheint das *kontrastive* Design in dieser Debatte zu wenig beachtet zu werden. Dabei ist gerade dieses Design in beiden „Kulturen“ gleichermaßen zentral. Die hier vorgelegte Typologie mag vielleicht dazu beitragen, diesem einigenden Band zwischen qualitativer und quantitativer Forschung stärker Rechnung zu tragen.

Literatur

- Axelrod, Robert (1970). *Conflict of Interest: A Theory of Divergent Goals with Applications to Politics*, Chicago.
- Becher, Michael (2010). Constraining Ministerial Power: The Impact of Veto Players on Labor Market Reforms in Industrial Democracies, 1973–2000, in: *Comparative Political Studies*, Vol. 43(1), 33–60.
- Black, Duncan (1948). On the Rationale of Group Decision-making, in: *Journal of Political Economy*, Vol. 56, 23–34.
- Boockmann, Bernhard (2006). Partisan politics and treaty ratification: The acceptance of International Labour Organisation conventions by industrialised democracies, 1960–1996, in: *European Journal of Political Research*, Vol. 45(1), 153–180.
- Bräuninger, Thomas (2005). A partisan model of government expenditure, in: *Public Choice*, Vol. 125(3–4), 409–429.
- Campbell, Donald T. (1975). “Degrees of Freedom” and the Case Study, in: *Comparative Political Studies*, Vol. 8 (2), 178–193.
- Clarke, Kevin A (2011). *A Nonparametric Approach to Testing Multiple Competing Models*, Unpublished work, University Of Rochester.
- Clarke, Kevin A/David M Primo (2012). *A model discipline: Political science and the logic of representations*, New York et al.
- Clarke, Kevin A. (2001). Testing Nonnested Models of International Relations: Reevaluating Realism, in: *American Journal of Political Science*, Vol. 45(3), 724–744.
- Clarke, Kevin A. (2007a). The Necessity of Being Comparative: Theory Confirmation in Quantitative Political Science, in: *Comparative Political Studies*, Vol. 40(7), 886–908.
- Clarke, Kevin A. (2007b). A Simple Distribution-Free Test for Nonnested Model Selection, in: *Political Analysis*, Vol. 15(3), 347–363.
- Davidson, R./J. G. Mackinnon (1981). Several tests for model specification in the presence of alternative hypotheses, in: *Econometrica*, Vol. 49, 781–793.
- Downs, Anthony (1957). *An Economic Theory of Democracy*, New York.
- Dür, Andreas (2007). Einige Anregungen zur Auswahl zwischen konkurrierenden Erklärungsansätzen in Y-zentrierter Forschung, in: Thomas Gschwend/Frank Schimmelfennig (Hg.): *Forschungsdesign in der Politikwissenschaft: Probleme - Strategien - Anwendungen*, Frankfurt /New York, 281–304.
- Franzese, Robert J (2010). The Multiple Effects of Multiple Policymakers: Veto Actors Bargaining in Common Pools, in: *Rivista italiana di scienza politica*, Vol. 40(3), 341–370.
- Franzese, Robert J. (2002). *Macroeconomic Policies of Developed Democracies*, Cambridge.
- Ganghof, Steffen (2003). Promises and Pitfalls of Veto Player Analysis, in: *Swiss Political Science Review*, Vol. 9(2), 1–25.
- Ganghof, Steffen (2005). Kausale Perspektiven in der vergleichenden Politikwissenschaft: X-zentrierte und Y-zentrierte Forschungsdesigns, in: Sabine Kropp/Michael Minkenberg (Hg.): *Vergleichen in der Politikwissenschaft*, Wiesbaden, 76–93.
- Ganghof, Steffen (2015). The Empirical Uses of Theoretical Models: The Case of Veto Player Theory, in: *Political Studies Review*, Forthcoming. DOI: 10.1111/1478-9302.12098
- Ganghof, Steffen/Kai Schulze (2015). Vetospieler und Institutionen, in: Georg Wenzelburger/Reimut Zohlnhöfer (Hg.): *Handbuch Policy-Forschung*, Wiesbaden, 113–148.

- Goertz, Gary/James Mahoney (2012). *A Tale of Two Cultures: Qualitative and Quantitative Research in the Social Sciences*, Princeton.
- Greene, William H. (2012). *Econometric Analysis, Seventh (International) Edition.*, Essex.
- Gschwend, Thomas/Frank Schimmelfennig (2007). Forschungsdesign in der Politikwissenschaft: Ein Dialog zwischen Theorie und Daten, in: Thomas Gschwend/Frank Schimmelfennig (Hg.): *Forschungsdesign in der Politikwissenschaft: Probleme – Strategien – Anwendungen*, Frankfurt /New York, 9–38.
- Hüller, Thorsten (2014). Sind Bundesverfassungsrichter ‚rationale Trottel‘?, in: *Zeitschrift für Politikwissenschaft*, Vol. 24, 5–28.
- Imai, Kosuke/Dustin Tingley (2012). A Statistical Method for Empirical Testing of Competing Theories, in: *American Journal of Political Science*, Vol. 56(1), 218–236.
- King, Gary/Robert O. Keohane/Sidney Verba (1994). *Designing Social Inquiry: Scientific Inference in Qualitative Research*, Princeton.
- Kittel, Bernhard/Rebecca B. Morton (2012). Introduction: Experimental Political Science in Perspective, in: Bernhard Kittel/Wolfgang J. Luhan/Rebecca B. Morton (Hg.): *Experimental Political Science: Principles and Practices*, Houndmills, 1–14.
- Kubbe, Ina (2016). *Experimente in der Politikwissenschaft: Eine methodische Einführung*, Wiesbaden.
- Lange, Matthew (2013). *Comparative-Historical Methods*, London.
- Laver, Michael/Kenneth A. Shepsle (1996). *Making and Breaking Governments*, Cambridge.
- Manow, Philip/Steffen Ganghof (2005). Mechanismen deutscher Politik, in: Steffen Ganghof/Philip Manow (Hg.): *Mechanismen der Politik. Strategische Interaktion im deutschen Regierungssystem*, Frankfurt, 9–33.
- Martin, Lanny W./Georg Vanberg (2014). Parties and Policymaking in Multiparty Governments: The Legislative Median, Ministerial Autonomy, and the Coalition Compromise, in: *American Journal of Political Science*, Vol. 58(4), 979–996.
- Meselson, Matthew/Franklin W. Stahl (1958). The Replication of DNA in *Escherichia Coli*, in: *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 44(7), 671–682.
- Morton, Rebecca B. /Kenneth C. Williams (2010). Experimental Political Science and the Study of Causality: From Nature to the Lab, New York.
- Rabinowitz, George/Stuart Elaine Macdonald (1989). A Directional Theory of Issue Voting, in: *The American Political Science Review*, Vol. 83(1), 93–121.
- Reiss, Julian (2015). A Pragmatist Theory of Evidence, in: *Philosophy of Science*, Vol. 82(3), 341–362.
- Schnapp, Kai-Uwe (2015). x-zentriertes Design, in: Rainer Diaz-Bone/Christoph Weischer (Hg.): *Methoden-Lexikon für die Sozialwissenschaften*, Wiesbaden, 86.
- Schneider, Carsten Q./Claudius Wagemann (2012). *Set-Theoretic Methods for the Social Sciences: A Guide to Qualitative Comparative Analysis*, Cambridge.
- Schnietz, Karen E. (2000). The Institutional Foundation of U.S. Trade Policy: Revisiting Explanations for the 1934 Reciprocal Trade Agreements Act, in: *Journal of Policy History*, Vol. 12 (4), 417–444.
- Sieberer, Ulrich (2007). „Aber könnte es nicht auch sein dass...?“. Die Auswahl unabhängiger Variablen in X-zentrierten und Y-zentrierten Forschungsdesigns, in: Thomas Gschwend/Frank Schimmelfennig (Hg.): *Forschungsdesign in der Politikwissenschaft: Probleme – Strategien – Anwendungen*, Frankfurt-New York, 257–283.
- Sober, Elliot (1990). Contrastive empiricism, in: C. W. Savage (Hg.): *Minnesota Studies in the Philosophy of Science*, Minneapolis, MN, 392–410.
- Sober, Elliot (2013). Empiricism, in: Martin Curd/Stathis Psillos (Hg.): *The Routledge Companion to Philosophy of Science*. Second Edition, New York, 160–169.
- Sober, Elliot (2015). *Ockham's Razors: A User's Manual*, New York.
- Tomz, Michael/Robert P. Van Houweling (2008). Candidate Positioning and Voter Choice, in: *American Political Science Review*, Vol. 102(03), 303–318.
- Tsebelis, George (1999). Veto Players and Law Production in Parliamentary Democracies: An Empirical Analysis, in: *American Political Science Review*, Vol. 93(3), 591–608.
- Tsebelis, George (2002). *Veto Players. How Political Institutions Work*, New York.
- Vanberg, Georg (2005). Verfassungsgerichtsbarkeit und Gesetzgebung: Zum politischen Spielraum des Bundesverfassungsgerichtes, in: Steffen Ganghof/Philip Manow (Hg.): *Mechanismen der Politik. Strategische Interaktion im deutschen Regierungssystem*, Frankfurt, 183–213.
- Volden, Craig/Clifford J. Carrubba (2004). The Formation of Oversized Coalitions in Parliamentary Democracies, in: *American Journal of Political Science*, Vol. 48(3), 521–537.
- Wenzelburger, Georg/Sebastian Jäckle/Pascal König (2014). *Weiterführende statistische Methoden für Politikwissenschaftler: Eine anwendungsbezogene Einführung mit Stata*, München.
- Wenzelburger, Georg/Frieder Neumann (2015). Vergleichende Staatstätigkeitsforschung: Entwicklungslinien – Heidelberger Schule – integrative Erklärungsansätze, in: Uwe Wagschal/Sebastian Jäckle/Georg Wenzelburger (Hg.): *Einführung in die vergleichende Politikwissenschaft: Institutionen – Akteure – Policies*, Stuttgart, 251–264.

Autor

Steffen Ganghof ist Professor für Vergleichende Politikwissenschaft an der Universität Potsdam. Er beschäftigt sich mit Demokratieforschung, politischer Ökonomie, politischer Philosophie und sozialwissenschaftlicher Methodologie. Zu seinen Buchpublikationen gehören *The Politics of Income Taxation* (ECPR Press) und *Mechanismen der Politik* (Campus, hrsg. mit Philip Manow). Seine Artikel sind unter anderem im *British Journal of Political Science*, *Comparative Political Studies*, *Party Politics* und *West European Politics* erschienen. Sein derzeitiger Forschungsschwerpunkt sind die Muster demokratischer Mehrheitsbildung im internationalen Vergleich.