

Empirische sozialwissenschaftliche Methoden in den Internationalen Beziehungen*

Der Ruf nach mehr Systematik und mehr Transparenz

Wie in allen sozialwissenschaftlichen Teildisziplinen muss empirisch-analytische Forschung in den IB möglichst systematisch und transparent sein, um für andere nachvollziehbar und überprüfbar zu sein. Sozialwissenschaftliche Methoden sind hierfür ein wichtiges Handwerkszeug. Um den Anschluss an die international hochwertige Forschung nicht zu verlieren, muss sich die deutsche IB stärker am State of the Art der sozialwissenschaftlichen Methoden orientieren. Dazu werden gute Kenntnisse der quantitativen und qualitativen Analyseverfahren in der Aus- und Weiterbildung des wissenschaftlichen Nachwuchses zunehmend wichtig. Zudem kann eine Orientierung an den Kriterien sozialwissenschaftlicher Forschung helfen, die Qualität von Untersuchungen zu verbessern und unsere Ergebnisse für andere nachvollziehbar zu machen. In Bezug auf Transparenz und Systematik haben vor allem die qualitativen Forschungsarbeiten Nachholbedarf. Eine wichtige Orientierung kann hier der Replikationsstandard geben.

1. Einleitung¹

Während im deutschsprachigen Publikationsraum der Eindruck entstehen kann, dass die politikwissenschaftliche Teildisziplin der Internationalen Beziehungen (IB) in methodischer Hinsicht unterbelichtet sei, sind Publikationen in international renommierten IB-Zeitschriften methodisch hochwertig fundiert. Zudem werden in Fachzeitschriften der IB wichtige methodische Weiterentwicklungen veröffentlicht. Da sich die internationale IB auf methodisch sehr hohem Niveau bewegt und sich ständig weiterentwickelt, läuft die deutsche IB immer stärker Gefahr, hier nicht mehr den Anschluss finden zu können. Wenn von Politikwissenschaft als einer »Disziplin auf der Suche nach Wissenschaftlichkeit« (Kittel 2009: 577) die Rede sein kann, dann

* Die Autorin und die Autoren dieses Forums sind von der ZIB-Redaktion gebeten worden, sich in ihren Beiträgen mit folgenden fünf Leitfragen auseinanderzusetzen: 1. Welcher Stellenwert sollte methodischen Fragen in den Internationalen Beziehungen (IB) zukommen?; 2. An welchen Kriterien und Traditionen sollten sich die IB in methodischer Hinsicht orientieren?; 3. Was sollten Methoden in den IB leisten?; 4. In welchem Verhältnis stehen Methode und Theorie?; 5. In welchem Verhältnis stehen Methode und Untersuchungsgegenstand? Es stand der Autorin und den Autoren jedoch frei, einzelne Fragen in den Vordergrund zu rücken, andere Fragen nur am Rande zu thematisieren und so eigene Schwerpunkte zu setzen.

1 Ich danke den beiden anonymen Gutachterinnen und Gutachtern für hilfreiche und anregende Kommentare und Michael Steinke für seine kompetente Unterstützung in der Aufarbeitung der in der ZIB publizierten Artikel der letzten Jahre.

scheint deren Teildisziplin der IB sich ganz besonders stark auf die Suche begeben zu müssen, um sich internationalen Standards in der empirisch-analytischen Forschung anzunähern. Wissenschaftliche Forschungsarbeiten sollten systematisch und transparent sein, damit unsere Ergebnisse für andere nachvollziehbar sind und kritisch bewertet werden können. Hierzu sind fundierte Methoden ein wichtiges Handwerkszeug.

2. Stellenwert der Methoden in den Internationalen Beziehungen

In den meisten Staaten Europas ist der Mainstream der Politikwissenschaft entweder sehr stark der interpretativ-hermeneutischen Vorgehensweise verhaftet oder weitgehend, wie Benoit Rihoux (2006: 333) es nennt »superjournalistic«, also ohne methodischen Anspruch. Weiterhin sind innerhalb der empirisch-analytischen Politikwissenschaften quantitative Forscherinnen und Forscher in der Minderheit, eine Situation, die in der europäischen IB – ganz ähnlich wie in Deutschland – noch stärker ausgeprägt sein dürfte.² Dies steht im krassen Gegensatz zum anglo-amerikanischen Sprachraum. Ein Überblick über die verwendeten Methoden in den führenden internationalen, englischsprachigen Zeitschriften der IB³ von 1975-2000 zeigte, dass rein deskriptiv historische Aufsätze zunehmend weniger publiziert werden, wohingegen methodisch fundierte Artikel inzwischen klar dominieren. Dabei bilden bereits Ende der 90er Jahre quantitative Artikel mit über 40% die Mehrzahl; die Verwendung formaler Modelle und Fallstudien findet sich in jeweils 14% der Artikel wieder (Sprinz/Wolinsky-Nahmias 2004: 5-8).

Ob die deutsche IB tatsächlich ein »weniger methodisch informierter Bereich« (Jahn 2007: 17) der Politikwissenschaft ist, lässt sich am besten anhand einer Bestandsaufnahme der Methodenorientierung in Publikationen bewerten. Hier fällt das Urteil für die deutsche IB leider nicht sehr positiv aus. So zeigt Bernhard Kittels (2009) Analyse, dass die Publikationen der IB in vier führenden deutschen politikwissenschaftlichen Zeitschriften weder rein theoretische noch theoriegeleitete oder deskriptiv empirische Aufsätze, sondern in überwältigender Mehrheit sachkundig erörternd sind. In den Jahren 1969-2008 wurden 89,2% aller Beiträge zur internationalen Politik und den internationalen Beziehungen mit wenig systematischen wissenschaftlichen Standards publiziert, was sich in der jüngsten Vergangenheit nur sehr geringfügig verbessert hat (Kittel 2009: 586).

Betrachten wir die Beiträge der *Zeitschrift für Internationale Beziehungen* (ZIB) der letzten fünf Jahre zeigt sich allerdings, dass es nicht ganz so schlimm steht um

2 In quantitativen Untersuchungsformen werden numerisch gemessene Konzepte und Hypothesen unter Verwendung statistischer Verfahren überprüft. Mithilfe qualitativer Methoden werden zumeist wenige Fälle detaillierter analysiert, wobei nicht numerische Informationen die Grundlage bilden, wie etwa historisches Archivmaterial oder nicht-standardisierte Interviews (King et al. 1994).

3 Die in der Studie bewerteten Zeitschriften waren *American Political Science Review*, *International Organization*, *International Security*, *International Studies Quarterly*, *Journal of Conflict Resolution* und *World Politics* (Sprinz/Wolinsky-Nahmias 2004: 5).

die deutschsprachige IB.⁴ Die ZIB hat zum Ziel, die »theoretisch und methodisch reflektierte IB-Forschung in Deutschland zu repräsentieren« (Zeitschrift für Internationale Beziehungen 2011). Diese besondere Betonung auf Theorie und Methode resultiert in der Tat in der Veröffentlichung von nur wenigen erörternden Arbeiten. Eine Durchsicht der Forschungsaufsätze (nicht der Beiträge zu Symposien oder Foren) der Jahre 2006-2010 hat gezeigt, dass sich die Mehrzahl der Forschungsartikel auf empirische sozialwissenschaftliche Methoden beziehen. Die verwendeten Methoden decken ein breites Spektrum ab. Sie reichen von Experimenten, über statistische Analysen bis hin zu vergleichenden Fallstudien und Inhaltsanalysen. In den letzten fünf Jahren konnten zehn der 30 Artikel der quantitativen Forschung zugerechnet werden.⁵ Erwartungsgemäß verwenden die meisten Artikel ein qualitatives Fallstudien-design, wobei die Systematik in der Fallauswahl stark variiert und die genauen methodischen Kriterien für die Durchführung und Datenerhebung ebenfalls mehr oder minder transparent sind. Die Fallstudien werden verstärkt zur Theorienbildung herangezogen bzw. dazu verwendet, Hypothesen mit Beispielen zu illustrieren oder zu plausibilisieren. Sehr wenige Publikationen sind rein theoretischer Natur. Insgesamt kann konstatiert werden, dass, anders als in den von Kittel (2009: 582) untersuchten Zeitschriften, sich die Artikel der ZIB der letzten Jahre an theoretischen und methodischen Aspekten orientieren, wobei die methodischen Anwendungen allerdings nicht immer sehr systematisch und transparent erscheinen und es fraglich ist, ob sie einer Replikation standhalten würden.

Zusammenfassend zeigen die verschiedenen Bestandsaufnahmen der Publikationen, dass in der deutschen IB quantitative Studien zwar präsent, jedoch – anders als im anglo-amerikanischen Publikationsraum – nach wie vor weit entfernt vom Mainstream sind (vgl. Kittel 2009). Die meisten Wissenschaftlerinnen und Wissenschaftler in der deutschen IB arbeiten qualitativ; Hypothesen werden zudem nicht getestet, sondern in erster Linie generiert und plausibilisiert (vgl. Deitelhoff/Wolf 2009: 461-462). Dem Mainstream der deutschen IB wird allgemein eine empirisch-analytische sozialwissenschaftliche Orientierung bescheinigt (Deitelhoff/Wolf 2009: 467; Kessler 2010). Fraglich ist jedoch, wie wissenschaftlich fundiert der Umgang mit Methoden tatsächlich ist und wie stringent sozialwissenschaftliche Kriterien angewandt werden. In Kombination mit den Befunden von Kittels (2009) Untersuchung muss die Schlussfolgerung gezogen werden, dass sozialwissenschaftliche Methoden nach wie vor keinen hohen Stellenwert in der deutschen IB einnehmen.

4 Kittel (2009: 582, Fn. 2) weist darauf hin, dass die führende IB-Zeitschrift, die ZIB, nicht in seiner Studie untersucht wurde.

5 Darunter ist ein Artikel mit Experimenten, vier Artikel verwenden multivariate Regressionsanalysen, die restlichen teilen sich auf quantitative Inhaltsanalysen, die Beschreibung von systematisch erhobenen Datensammlungen und deskriptive Statistiken auf.

3. Internationale Beziehungen als empirisch-analytische Sozialwissenschaft

Die meisten Forscherinnen und Forscher, die die IB als eine empirisch-analytische Disziplin verstehen, orientieren sich an den Kriterien sozialwissenschaftlichen Arbeitens.⁶ Der (sozial-)wissenschaftliche Ansatz setzt auf die Vergleichbarkeit und Generalisierung, d.h. er versucht, wiederholt auftretende Ereignistypen und Muster zu identifizieren und für diese verallgemeinernde Aussagen zu treffen. Was diesen Ansatz auch auszeichnet, ist die Überzeugung, dass zwar soziales Verhalten (*behavior*) wissenschaftlich analysiert werden kann, aber nicht unbeobachtbare oder gar übernatürliche Phänomene. Forscherinnen und Forscher, die sich an sozialwissenschaftlichen Kriterien orientieren, fühlen sich verpflichtet, aus klaren theoretischen Argumenten abgeleitete Hypothesen empirisch zu überprüfen. Dazu werden empirische Daten systematisch erhoben und analysiert. Obwohl in den Sozialwissenschaften nicht damit zu rechnen ist, deterministische Gesetze wie in den Naturwissenschaften zu finden, sind Vertreterinnen und Vertreter dieses Ansatzes der Überzeugung, dass Studien trotzdem wissenschaftlich sein können, indem nach probabilistischen Erklärungen gesucht wird (vgl. Ellis 1994; Russett et al. 2006: 34; Bueno de Mesquita 2010). Empirische sozialwissenschaftliche Forschung sollte zudem replizierbar sein, wenn sie dem wissenschaftlichen Anspruch genügen will (King 1995).⁷ Während manche noch so weit gehen möchten, die wissenschaftliche Methode auf statistische Verfahren zu beschränken (Ellis 1994), sehen andere die oben genannten Kriterien wissenschaftlichen Arbeitens durchaus auch anwendbar auf qualitative Studien (King et al. 1994). Viele qualitative Forscherinnen und Forscher verwehren sich jedoch dagegen, dass die quantitative Schablone der qualitativen Forschung übergestülpt wird (Brady et al. 2006). Besonders unterstreichen möchte ich im Folgenden daher nur die Kriterien der Systematik und Transparenz, die sowohl auf quantitative und qualitative Forschungsarbeiten zutreffen.

Theorien werden nicht nur zur Erklärung der Realität entwickelt, wir brauchen auch systematische Beweismittel/Zeugnisse (Evidenz), um diese Theorien zu testen. Eine systematische Vorgehensweise ist ein wichtiges Merkmal wissenschaftlicher Forschung (vgl. Ellis 1994). Systematik im Forschungsprozess verlangt nach sozialwissenschaftlichen Methoden. Systematisch forschen heißt, nach einem Plan oder nach einer Methode vorzugehen, wenn möglich nach objektivierbaren Kriterien; vor allem heißt es, nicht willkürlich Informationen zu sammeln und auszuwerten. Zu einer systematischen Vorgehensweise werden wir durch eine genaue Dokumentation und Begründung der ausgewählten Forschungsmethoden angeleitet. Mit einer systematischen Auswahl und Erhebung der Daten, der Interviewpartnerinnen und Interview-

6 Vgl. King et al. (1994); Jahn (2007: 12); Russett et al. (2006: 32-34); Ellis (1994: 9-12).

7 Der wissenschaftliche Ansatz ist in Grundzügen in vielen nordamerikanischen Einführungslernbüchern der IB beschrieben (z.B. Bueno de Mesquita 2010: 388-415; Russett et al. 2006: 27-48) und findet schon frühzeitig Eingang in die Ausbildung des wissenschaftlichen Nachwuchses. Die institutionelle Vertretung dieses Ansatzes findet sich u.a. in der *International Studies Association*-Sektion »Scientific Study of Social Processes«, in der sich vor allem die quantitativ orientierte IB-Forschungsgemeinschaft wiederfindet, und in der *Peace Science Society International*.

partner und der auszuwertenden Quellen verringern wir zudem die Gefahr, dass die Ergebnisse unserer Forschung verzerrt sind.

Das Zusammentragen der empirischen Beweisführung sollte dabei unbedingt transparent sein, damit eine Bewertung für andere möglich ist. Diese Forderung nach Transparenz und Nachvollziehbarkeit findet sich im Replikationsstandard wieder. Dies bedeutet, dass im Forschungsbericht die methodische Vorgehensweise so detailliert und präzise dokumentiert sein sollte, dass eine andere Forscherin oder ein anderer Forscher die Studien replizieren könnte, also mit genau derselben methodischen Vorgehensweise und denselben Daten zu denselben Ergebnissen kommen müsste. Nur wenn die methodische Vorgehensweise transparent dokumentiert ist, können andere Forscherinnen und Forscher die Qualität empirischer Analysen bewerten. Und nur so erzielen wir eine hohe Glaubwürdigkeit unserer Befunde (vgl. King 1995). Empirische Evidenz sollte möglichst objektiv, d.h. die Ergebnisse nicht entsprechend unserer Erwartungen verzerrt sein. Auch wenn Forschung nie gänzlich objektiv und frei von ethischen Werten sein wird, so wird die wissenschaftliche Methode von vielen als die am wenigsten subjektive Vorgehensweise bezeichnet (Bueno de Mesquita 2010: 388), um alternative Erklärungen für ein Phänomen zu bewerten.

Der Replikationsstandard ist in der quantitativen Forschung inzwischen weit verbreitet, so dass Zeitschriften zunehmend die Autorinnen und Autoren verpflichten, ihre Datensätze öffentlich zugänglich zu machen. Replikationen durch andere Forscherinnen und Forscher können somit Mängel im Forschungsdesign und die fehlende Robustheit der Ergebnisse aufdecken, aber auch Fehler in den Daten selbst offenlegen. Zwar beschränken sich die meisten Replikationsstudien auf die Re-Analyse mit den bestehenden Datensätzen, aber auch eine Transparenz der Datensammlung ist in den meisten Fällen gewährleistet und eine Replikation der Datenerhebung ist, wenn auch arbeitsaufwendig, im Prinzip möglich. Bei vielen qualitativen Studien gibt es diesbezüglich immer noch Nachbesserungsbedarf (vgl. King et al. 1994: 26-27). Ein transparenter Forschungsbericht müsste genaue Angaben zu den erhobenen Daten (z.B. den ausgewählten Quellen oder bei durchgeführten Interviews zur Form, zu den Fragen und zur Auswahl der Interviewpartnerinnen und Interviewpartner) enthalten. Leider ist dies, auch in den Artikeln der ZIB, eher selten der Fall. Häufig wird zwar beschrieben, dass Fallstudien oder Experteninterviews durchgeführt wurden, ohne aber die Vorgehensweise transparent zu machen. Somit lässt sich für andere nicht bewerten, ob die Quellen oder Interviewpartnerinnen und Interviewpartner systematisch oder mit Bias ausgewählt wurden.⁸

Welche Methoden ausgewählt werden sollten, ob formales Modellieren, statistische Verfahren oder ein Fallstudiendesign, hängt in erster Linie von der zu erforschenden Fragestellung ab, etwa ob wir eine verallgemeinernde oder eine auf nur wenige Fälle beschränkte These testen möchten. Welche Fragestellung wir untersuchen bzw. deren

8 Die Bedeutung von Replikationsstudien ist auch insofern wichtig, da wir nie eine einzelne Studie als Beweis für eine Hypothese betrachten dürfen. Erst mit vielen Studien, die möglichst auch mit unterschiedlichen Daten und Methoden durchgeführt werden, sollten wir Vertrauen in ein theoretisches Argument entwickeln dürfen, das selbstverständlich durch neuere Investitionen korrigiert werden könnte (Ellis 1994: 11).

Eingrenzung wird uns von der Theorie oder Teilaspekten einer solchen vorgegeben, die wir entwickeln, adaptieren und vor allem überprüfen möchten.

4. Methoden als Handwerkszeug

Methoden sollten in der IB nichts anderes leisten als in anderen sozialwissenschaftlichen Disziplinen auch. Sie haben die Aufgabe, Strukturen und systematische Verfahren vorzugeben, sie liefern das Handwerkszeug für die Analyse empirischer und theoretischer Problemstellungen und geben den Forschungsprozess vor. Indem wir methodisch fundiert an Fragestellungen herangehen, gewährleisten wir eine systematische Vorgehensweise. Dadurch wird unsere Arbeit für andere nachvollziehbar und somit überprüfbar. Eine klare und transparente Anwendung von Methoden diszipliniert unsere Forschungsarbeit in gewisser Weise (Morrow 2010: 299).

Dabei dürfte es unerheblich sein, ob wir quantitative oder qualitative Methoden verwenden. Statistische Methoden haben laut James D. Morrow (2010: 299-300) jedoch einen gewissen Vorteil. Sie halten uns dazu an, möglichst viele vergleichbare Fälle mit einem einheitlichen Forschungsprozess zu untersuchen. Statistisches Vorgehen zwingt uns zudem, Maßzahlen für (komplexe) Konzepte zu entwerfen. Die Validität der Indikatoren kann unterschiedlich hoch sein und wird leider häufig von bestehenden Datensätzen vorgegeben. Zwar sind die Maße oft verbesserungsfähig und stoßen gelegentlich auch an Grenzen, aber sie vermeiden, dass wir Konzepte vage halten und von Fall zu Fall unterschiedlich interpretieren können. Statistische Tests geben eine systematische Bewertung von Evidenz vor, also klare Richtlinien, nach denen wir uns verständigen können, ob die getestete Hypothese Unterstützung findet oder nicht.

Es gibt keine sozialwissenschaftlichen Methoden, die IB-spezifisch sind. Kennzeichnend hierfür ist die Tatsache, dass die meisten Methodenbücher auf die Sozialwissenschaften allgemein (z.B. Ellis 1994; King et al. 1994) bzw. auf die Politikwissenschaften (z.B. Gschwend/Schimmelfennig 2007; Van Evera 1997) und nur wenige speziell auf die IB ausgerichtet sind (Sprinz/Wolinsky-Nahmias 2004; Klotz/Prakash 2008).

Viele internationale IB-Forscherinnen und Forscher setzen sich jedoch sehr stark mit methodischen Fragestellungen auseinander. Dass methodischen Fragen unter den Fachkolleginnen eine besondere Aufmerksamkeit gilt, lässt sich schon daran erkennen, dass der am häufigsten zitierte Artikel im Bereich der Konfliktforschung ein methodischer Artikel von Nathaniel Beck, Jonathan Katz und Richard Tucker (1998) mit dem wenig attraktiven Titel »Taking Time Seriously in Binary Time-Series-Cross-Section Analysis«, ist. Der am zweithäufigsten genannte Artikel stellt den Datensatz des *Correlates of War*-Projektes zu *Militarized Interstate Disputes* (Jones et al. 1996) vor (ESI-Topics 2011). Als weiteres Beispiel für die Aufmerksamkeit, die methodische Artikel zuweilen erfahren, ist der Beitrag von Michael Ward, Brian Greenhill und Kristin Bakke (2010) zu Modellbewertungen mittels Prognosen zu

nennen, der vom *Journal of Peace Research* mit dem Preis für den besten Artikel des Jahres 2010 ausgezeichnet wurde.

Methoden entwickelnde Beiträge finden eher im Organ *Political Analysis* einen Veröffentlichungsraum. Ausgewiesene Methodenexpertinnen und Methodenexperten stellen aber doch des Öfteren neue Ansätze auch in den Zeitschriften der IB vor. In der internationalen IB beschäftigen sich Forscherinnen und Forscher sehr intensiv mit methodischen Aspekten, auch auf technisch sehr hohem Niveau, wie z.B. ein Symposium in *International Organization* zur Verwendung von *fixed-effects*-Panel-schätzungen, um ungemessene Heterogenität aufzugreifen (z.B. Green et al. 2001; Beck/Katz 2001), oder die Beiträge zum Sonderheft *Conflict Management and Peace Science* zur Modellspezifizierung und Auswahl von Kontrollvariablen (z.B. Achen 2005; Oneal/Russett 2005) belegen. Kurzum, in internationalen Fachzeitschriften werden sehr wohl stark methodisch entwickelnde Arbeiten prominent in der IB veröffentlicht.

Die quantitative Forschergemeinschaft in der IB war lange Zeit verschiedenen Kritikpunkten ausgesetzt, worauf sie mit methodischen Weiterentwicklungen und Verbesserungen reagierte. Eine gute Kenntnis der neueren Methoden ist insofern wichtig, da sie international zum Standard gehören. Ohne deren Beherrschung ist es schwierig, in international renommierten Fachzeitschriften zu publizieren. Insbesondere die neueren statistischen und ökonometrischen Verfahren werden aber unter deutschen IB-Forschern und Forscherinnen nur von einer Minderheit beherrscht. Im Folgenden werden ein paar Beispiele wichtiger methodischer Fragen in der quantitativen IB dargelegt: die Thematik der dyadischen Analyseeinheit, der Umgang mit der statistischen Signifikanz in großen Samples und der Umgang mit Endogenität. Vor allem die Besprechung des letzten Punktes soll der häufig geäußerten Kritik, *large-n*-Studien könnten nur Aussagen über Korrelationen aber nicht Kausalitäten treffen, entgegenreten.

Da internationale Beziehungen Interaktionen zwischen staatlichen und nichtstaatlichen Akteuren sind, wird in vielen quantitativen Studien der dyadische Ansatz verwendet. Angefangen mit der Studie von Stuart A. Bremer (1992) haben die gängigen Modelle zur Erklärung von internationalen militärischen Konflikten eine Dyade, also ein Staatenpaar, als Untersuchungseinheit gewählt (z.B. Oneal/Russett 2005). Methodische Entscheidungen, die getroffen werden mussten, beschäftigten sich mit der Frage, ob alle Dyaden oder nur politisch relevante Staatenpaare in die jeweiligen Studien aufgenommen werden sollten (Lemke/Reed 2001). D. Scott Bennett und Alan C. Stam (2000) führten die Unterscheidung zwischen gerichteten und nicht-gerichteten Dyaden ein und ermöglichten somit die statistische Analyse von Fragen, wie z.B. welcher der beiden Staaten in der Dyade der Initiator und welcher das Ziel militärischer Aggression ist.⁹

In den quantitativen Studien der IB werden häufig alle Staaten des internationalen Systems untersucht bzw. alle international anerkannten Staaten mit einer Mindest-

9 Dyadische Studien sind in der Konfliktforschung, aber auch in den Studien zu internationalen Wirtschaftsbeziehungen üblich (z.B. Morrow et al. 1998).

größe von 500.000 Einwohnerinnen und Einwohnern. Bei dieser großen Zahl an Beobachtungen, besonders bei dyadischen Studien, werden die Varianzen so gering, dass wir beinahe immer statistisch signifikante Ergebnisse erzielen. Vor allem deswegen wird der Sternchen-Fokus, also die Bedeutung, die der statistischen Signifikanz zugeschrieben wird, seit Längerem als unangemessen kritisiert (z.B. Ward et al. 2010).¹⁰ Die statistische Signifikanz sagt nichts darüber aus, ob die Ergebnisse auch inhaltlich bedeutsam sind. Deswegen ist es inzwischen üblich, dass in quantitativen Studien zusätzlich die substantiellen Effekte berechnet und interpretiert werden (King et al. 2000) und in Zukunft werden wir vermutlich verstärkt eine Bewertung der Modelle anhand ihrer Prognosekraft sehen können. So stellen Ward und Koautoren (2010) eine Methode vor, mit der wir Modelle in Bezug auf ihre Vorhersagekraft vergleichend bewerten können. Vor dem Hintergrund, dass wir verallgemeinernde Aussagen anstreben, haben wir letztendlich natürlich den Anspruch, dass unsere Modelle in hoffentlich nicht allzu ferner Zukunft so gut sein werden, dass wir über unser *sample* hinaus möglichst zutreffende Vorhersagen treffen können (*out-of-sample prediction*), um gewissen Ereignissen wie militärischen Konflikten rechtzeitig entgegenwirken zu können. Die quantitative IB-Forschung im Bereich der Konfliktforschung hat angefangen, sich methodisch in diese Richtung weiter zu entwickeln (z.B. Bueno de Mesquita 2011). Bereits jetzt nehmen Publikationen zu wissenschaftlichen Prognosen zu.¹¹

Die quantitativ-sozialwissenschaftliche Forschung muss sich regelmäßig dem Vorwurf stellen, dass mit Regressionsanalysen nur Korrelationen, aber nicht Kausalitäten analysiert werden könnten. Einige quantitative Forscherinnen und Forscher haben jedoch sehr wohl den Anspruch, mithilfe kausaler Inferenzen die Ursachen empirischer Phänomene zu erklären. Ähnlich wie in Experimenten soll geklärt werden, welche Bedingungen vorhanden sein müssten, um ein bestimmtes Ereignis oder Verhalten zu beobachten. Der Vergleich findet zwischen dem Vorhandensein und Nichtvorhandensein der Bedingung statt. Quasi-experimentelle Verfahren und Regressionsanalysen sind ein geeigneter Ersatz, um kausale Inferenz zu ziehen, selbst dann, wenn keine Randomisierung der Fallauswahl möglich sein sollte. Wichtig ist, dass sich die Forscherinnen und Forscher der damit verbundenen Ungewissheit bewusst sind und entsprechende Unsicherheitsmaße im Forschungsbericht angeben (Angrist/Pischke 2009; King et al. 1994).

Zudem greifen Sozialwissenschaftlerinnen und Sozialwissenschaftler seit geraumer Zeit auf ökonometrische Analyseverfahren zurück, insbesondere auf Instrumentalvariablenansätze, um kausale Beziehungen zu untersuchen. Im Gegensatz zur Statistik ist die Ökonometrie nicht schüchtern im Umgang mit kausaler Inferenz (Angrist/Pischke 2009). So verfügt die Ökonometrie über einen ganzen Werkzeugkasten, mit

10 Gelegentlich haben wir in der IB-Forschung zwar viele Fälle, aber gewisse Ereignisse, wie z.B. Kriege treten selten auf, was die Analysen teilweise erschwert. Für die Schätzung seltener Ereignisse haben Methodikerinnen und Methodiker jedoch besondere statistische Verfahren entwickelt (King/Zeng 2001).

11 Siehe den Überblicksartikel zu Entwicklungen auf diesem Gebiet von Schneider et al. (2011).

dem Endogenitätsprobleme und Fragen der Kausalität angegangen werden können. Durch Instrumentalvariablen-Schätzverfahren und simultane Gleichungssysteme können wir die Richtung der Kausalität zweier Variablen schätzen oder mit Granger-Kausalitätstests die zeitliche Dimension untersuchen.¹²

Während die quantitative IB-Forschung eine beeindruckende methodische Weiterentwicklung erfährt, hat auch die qualitative Forschung methodisch keinen Stillstand zu verzeichnen. Angestoßen durch das Buch von Gary King und Koautoren (1994), das bei qualitativen Forscherinnen und Forschern auf gemischte Reaktionen traf, haben sich qualitative Methodikerinnen und Methodiker intensiv mit Fragen des häufig vorgeworfenen Selektionsbias bei Fallstudien beschäftigt. Vergleichende Fallstudien-*designs* gehen inzwischen sehr bewusst mit der Fallauswahl um. Verfahren wie das *process tracing* sind geeignet, um Kausalmechanismen innerhalb eines Falles nachzuvollziehen und *fuzzy-set*-Analysen verfolgen das Ziel der Verallgemeinerung und bedienen sich der Aufdeckung von notwendigen und hinreichenden Bedingungen, um nur ein paar Beispiele zu nennen (vgl. Bennett/Elman 2006: 459-460, 468-472; Brady et al. 2006; Levy 2008). Zunehmend werden Rufe nach der Kombination von quantitativen und qualitativen Methoden, einem *mixed-methods*-Ansatz, laut, was den Vorteil haben könnte, Schwächen in den quantitativen und qualitativen Methoden auszugleichen. Dazu bedarf es jedoch Forscherinnen und Forscher, die beide Methoden auf gutem Niveau beherrschen, und davon gibt es bisher nur wenige Beispiele (Sprinz/Wolinsky-Nahmias 2004: 375-378).

Auch in den hermeneutisch-interpretativen Analyseverfahren findet sich zu unterschiedlichem Grad Systematik und Transparenz. Im Forschungsfeld zu den Fragestellungen nach normativen Werten und kommunikativem Handeln hat sich die Methode der Diskursanalyse etabliert. Dabei werden aus möglichst vielen Texten Repräsentationen der Realität herausgearbeitet (vgl. Neumann 2008). Normen und Werte selber sind zwar nicht direkt beobachtbar, aber sie können sich empirisch übersetzen in Sprache und Schriften, die dann diskursanalytisch und mit entsprechender Systematik ausgewertet werden könnten. Probleme in der Textauswahl und der Abbildung und Identifizierung von Repräsentationen erinnern an Probleme mit der operationalen Definition von Konzepten und der systematischen Auswahl von Quellen und Daten anderer Methoden.

5. Verhältnis zwischen Theorie und Methode

Einig sind sich die Forscherinnen und Forscher nicht nur in der Bedeutung von empirischen Analysen, sondern auch darin, dass sozialwissenschaftliche Forschung immer theoriegeleitet sein sollte. Theorien als Erklärungen der Zusammenhänge von empirischen Phänomenen bestehen aus einer Vielzahl an Annahmen und Hypothesen,

12 In dieser Werkzeugkiste befinden sich auch Verfahren zum Umgang mit dem systematischen Bias bei nicht zufällig ausgewählten Fällen (Heckman-Selektionsmodelle) und vieles mehr (vgl. Schneider/Ruoff 2010).

die einer Überprüfung bedürfen. Eine gute Theorie sollte nicht nur sparsam sein und eine große Erklärungskraft haben, sondern auch falsifizierbar sein (Van Evera 1997; King et al. 1994). Eine nicht widerlegbare These bringt uns in wissenschaftlicher Hinsicht nicht weiter, da die empirische Evidenz immer passend interpretiert werden könnte.

In den IB gibt es eine Vielzahl an Theorien, die noch unzureichend bewertet sind, sowohl in Bezug auf deren interne Konsistenz als auch deren empirische Erklärungskraft. Ob Theorien logisch konsistent sind, kann mithilfe von formalen mathematischen Modellen überprüft werden. Die mathematische Sprache ist sehr präzise und logisch und kann uns so helfen, Ungereimtheiten in der Theorie aufzudecken (vgl. Snidal 2004: 233-237). Eine Theorie sollte aber auch nützlich sein und das erklären, wozu sie entwickelt wurde, sollte also mit der beobachtbaren Realität, der empirischen Evidenz übereinstimmen. Um die empirische Erklärungskraft zu testen, liefern die verschiedenen quantitativen und qualitativen Methoden wiederum das Handwerkszeug.

In vielen Fällen steht eine gründliche Überprüfung von Theorien und Hypothesen noch aus. Trotzdem werden laufend neue theoretische Argumente erarbeitet. Dies ist einerseits dem Trend geschuldet, dass ein einzelner Artikel im besten Fall ein formal modelliertes Argument mit empirischen Überprüfungen verbinden sollte, was häufig in oberflächlichen Tests resultiert. Dieser Anspruch, jeder Analyse ein formales Modell voranzustellen, verhindert zudem, dass Hypothesen durch mehrere Analysen mit unterschiedlichen Forschungsdesigns und unterschiedlichen Methoden gründlich überprüft werden. Andererseits liegt einer der Gründe für die kontinuierliche Entwicklung neuer Theorien darin, dass viele qualitative Forscherinnen und Forscher Fallstudien benutzen, um Hypothesen zu generieren, statt bestehende Hypothesen rigoros zu testen. Zwar lassen sich mit Fallstudien, insbesondere der Methode des *process tracing*, kausale Mechanismen sehr gut entwickeln, aber diese müssen dann auch an anderen (möglichst vielen) Fällen überprüft werden, damit wir Theorien mit zu geringer Erklärungskraft und Reichweite verwerfen können. Ein Plädoyer für mehr empirische (vor allem aufeinander aufbauende und vergleichbare) Tests von theoretischen Annahmen und Hypothesen ist dringend geboten.

6. Verhältnis von Methode und Untersuchungsgegenstand

Die Auswahl des geeigneten methodischen Handwerkszeugs hängt von mehreren Faktoren ab. Zunächst wird die Anzahl der zu untersuchenden Fälle erste Vorgaben machen, ob wir mit *large-n*-Studien und quantitativ vorgehen. Bei ca. zehn bis 60 Fällen bietet sich möglicherweise die Anwendung der *fuzzy-set*-Mathematik an, bei sehr wenigen Fällen die herkömmlichen vergleichenden Fallstudiendesigns oder bei einem einzelnen Fall z.B. das *process tracing*, um kausale Mechanismen nachzuziehen (vgl. Bennett/Elman 2006: 470; Levy 2008). Die Entscheidung wird auch dadurch bestimmt, ob wir verallgemeinernde Aussagen treffen oder uns auf einzelne oder wenige Fälle konzentrieren möchten. Allerdings kann eine vergleichende Studie weniger

Länder, was wir in der IB häufig als Fallstudien bezeichnen, oder eine Einzelfallstudie auch mit statistischen Verfahren durchgeführt werden. So können auch Staaten in Individuen, in Bezirke oder in eine Vielzahl an Ereignissen und somit in ausreichend viele Beobachtungen für eine statistische Analyse untergliedert werden (z.B. Hegre et al. 2009).

Innerhalb der quantitativen Forschung ist die Auswahl der angemessenen Methode sehr stark abhängig von der Datenstruktur und von der Ausprägung der zu erklärenden Variable. Das einfachste Schätzverfahren für Regressionsanalysen, die Methode der kleinsten Quadrate, findet Anwendung bei intervallskalierten abhängigen Variablen, leider eine eher seltene Ausprägung vieler Variablen in der IB. Stattdessen müssen wir häufig auf Methoden wie Logit oder Probit-Schätzung bei Variablen dichotomer Ausprägung (z.B. Krieg/Nicht-Krieg) oder auf Verfahren für Zählvariablen zurückgreifen (z.B. Anzahl der Todesopfer). Die Auswahl der adäquaten Methode ist hier eine Standardvorgehensweise. Die Datenstruktur bestimmt weiterhin, ob wir Verfahren für Zeitreihen, Querschnittanalysen oder spezielle Verfahren für Daten mit Panelstruktur verwenden. Zeitliche und räumliche Abhängigkeiten in den Daten lassen sich mit verschiedenen ökonometrischen Korrekturen aufgreifen. Hierzu gibt es Tests, die zum Standardrepertoire der Statistik und der Ökonometrie gehören, mit deren Hilfe wir bestimmen können, welche Verfahren wir genau zu verwenden haben. Welche Verfahren verwendet werden, hängt allerdings nicht selten von den methodischen Fähigkeiten der Forscherin und des Forschers ab. Auch wenn die Verwendung elaborierter Verfahren verführerisch sein mag, gilt jedoch klar die Maxime, dass weniger mehr ist und dass einfache Verfahren den komplexeren vorzuziehen sind. Dieser Grundsatz scheint teilweise in Vergessenheit geraten zu sein. Zunehmend lesen wir quantitative Forschungsartikel, die sich intensiv mit der Anwendung relativ komplizierter methodischer Verfahren befassen und dabei die theoretische Fragestellung und Fundierung aus dem Blick verlieren (vgl. Morrow 2010: 305).

7. Schlusswort

Zwar mag die bisherige Kritik der Methodenabstinenz in der deutschen IB in gewisser Weise zutreffen (z.B. Jahn 2007: 17), insgesamt setzt sich aber auch bei deutschen IB-Forscherinnen und Forschern zunehmend die Erkenntnis durch, dass Methoden – ähnlich wie Theorien – ein unerlässliches Handwerkszeug für sozialwissenschaftliche Forschung sind. Ein gutes Verständnis des *state of the art* der sozialwissenschaftlichen Methoden, das notwendig wäre, um auf internationalen Standards zu publizieren, ist in der deutschen IB nicht genügend breit vorhanden. Nur wenige Forscherinnen und Forscher beherrschen die entsprechenden Methoden auf einem ausreichenden Niveau, um in international renommierten Zeitschriften publizieren zu können. Das wachsende Interesse an einer stärkeren, gerade auch quantitativen Methodenorientierung lässt sich vor allem bei Nachwuchswissenschaftlerinnen beobachten und ist sicherlich auch dem Druck, in englischsprachigen Zeitschriften mit hohem Ranking zu publizieren, zu verdanken. Das Interesse an quantitativen Methoden in der IB steigt

aber in einem Tempo, das noch sehr langsam ist. Problematisch ist dabei besonders, dass sich die Methoden rasant weiterentwickeln und anspruchsvoller werden. Statistische Kenntnisse sind bei weitem nicht mehr ausreichend, sondern müssen durch Kenntnisse der Ökonometrie ergänzt werden. Trotz der ebenfalls weitreichenden Entwicklungen der letzten Jahre im Bereich der qualitativen Methoden herrscht hier vor allem noch Nachbesserungsbedarf in der Systematik und Transparenz des Forschungsprozesses.

Teilweise mögen die fehlenden Methodenkenntnisse der Grund sein, warum sich Forscherinnen und Forscher für eine nicht systematische Vorgehensweise entscheiden. Deswegen kann an dieser Stelle nur die Forderung anderer (z.B. Kittel 2009: 591-593) nach mehr Methodenausbildung aufs Dringlichste unterstützt werden. So sind nur wenige Fachbereiche diesbezüglich so gut aufgestellt, dass der wissenschaftliche Nachwuchs die notwendige Ausbildung erhält. Es gibt zwar vereinzelt Forscherinnen und Forscher, die sich mittels »*learning by doing*« und unterstützt durch Sommerschulen notwendige Methodenkenntnisse aneignen, doch ist dies, wenn auch nicht unmöglich, bedeutend anstrengender.

Literatur

- Achen, Christopher H. 2005: Let's Put Garbage-Can Regressions and Garbage-Can Probits Where They Belong, in: *Conflict Management and Peace Science* 22: 4, 327-339.
- Angrist, Joshua D./Pischke Jörn-Steffen 2009: *Mostly Harmless Econometrics*, Princeton, NJ.
- Beck, Neal/Katz, Jonathan N. 2001: Throwing out the Baby with the Bath Water: A Comment on Green, Kim, and Yoon, in: *International Organization* 55: 2, 487-495.
- Beck, Nathaniel/Katz, Jonathan N./Tucker, Richard 1998: Taking Time Seriously in Binary Time-Series-Cross-Section Analysis, in: *American Journal of Political Science* 42: 4, 1260-88.
- Bennett, Andrew/Elman, Colin 2006: Qualitative Research: Recent Developments in Case Study Methods, in: *Annual Review of Political Science* 9: 455-76.
- Bennett, D. Scott/Stam, Alan C. 2000: Research Design and Estimator Choices in the Analysis of Interstate Dyads, in: *Journal of Conflict Resolution* 44: 5, 653-685.
- Brady, Henry E./Collier, David/Seawright, Jason 2006: Towards a Pluralistic Vision of Methodology, in: *Political Analysis* 14: 3, 353-368.
- Bremer, Stuart A. 1992: Dangerous Dyads: Conditions Affecting the Likelihood of Interstate War, 1816-1965, in: *Journal of Conflict Resolution* 36: 2, 309-341.
- Bueno de Mesquita, Bruce 2010: *Principles of International Politics*, 4. Auflage, Washington, DC.
- Bueno de Mesquita, Bruce 2011: A New Model for Predicting Policy Choices: Preliminary Tests, in: *Conflict Management and Peace Science* 28: 1, 65-85.
- Deitelhoff, Nicole/Wolf, Klaus Dieter 2009: Der Widerspenstigen Selbst-Zähmung? Zur Professionalisierung der Internationalen Beziehungen in Deutschland, in: *Politische Vierteljahresschriften* 50: 3, 451-475.
- Ellis, Lee 1994: *Research Methods in the Social Sciences*, Madison, WI.
- ESI-Topics 2011: Armed Conflict, in: <http://www.esi-topics.com/armed-conflict/papers/a1.html>; 28.04.2011.
- Green, Donald P./Kim, Soo Yeon/Yoon, David H. 2001: Dirty Pool, in: *International Organization* 55: 2, 441-468.

- Gschwend, Thomas/Schimmelfennig, Frank 2007: Forschungsdesign in der Politikwissenschaft, Frankfurt a. M.
- Hegre, Håvard/Østby, Gudrun/Raleigh, Clionadh 2009: Poverty and Civil War Events: A Disaggregated Study of Liberia, in: *Journal of Conflict Resolution* 53: 4, 598-623.
- Jahn, Detlef 2007: Was ist Vergleichende Politikwissenschaft? Standpunkte und Kontroversen, in: *Zeitschrift für Vergleichende Politikwissenschaft* 1: 1, 9-27.
- Jones, Daniel M./Bremer, Stuart A./Singer, J. David 1996: Militarized Interstate Disputes, 1816-1992: Rationale, Coding Rules, and Empirical Patterns, in: *Conflict Management and Peace Science* 15: 2, 163-213.
- Kessler, Oliver 2010: Internationale Politische Theorie. Fehlentwicklung oder fehlende Entwicklung der IB?, in: *Zeitschrift für Internationale Beziehungen* 17: 2, 317-334.
- King, Gary 1995. Replication, Replication, in: *PS: Political Science and Politics* 28: 3, 444-452.
- King, Gary/Keohane, Robert O./Verba, Sidney 1994: *Designing Social Inquiry: Scientific Inference in Qualitative Research*, Princeton, NJ.
- King, Gary/Tomz, Michael/Wittenberg, Jason 2000: Making the Most of Statistical Analyses: Improving Interpretation and Presentation, in: *American Journal of Political Science* 44: 2, 347-361.
- King, Gary/Zeng, Langche 2001: Explaining Rare Events in International Relations, in: *International Organization* 55: 3, 693-715.
- Kittel, Bernhard 2009: Eine Disziplin auf der Suche nach Wissenschaftlichkeit. Entwicklung und Stand der Methoden in der deutschen Politikwissenschaft, in: *Politische Vierteljahresschriften* 50: 3, 577-603.
- Klotz, Audie/Prakash, Deepa (Hrsg.) 2008: *Qualitative Methods in International Relations: A Pluralist Guide*, New York, NY.
- Lemke, Douglas/Reed, William 2001: The Relevance of Politically Relevant Dyads, in: *Journal of Conflict Resolution* 45: 1, 126-144.
- Levy, Jack S. 2008: Case Studies: Types, Designs, and Logics of Inference, in: *Conflict Management and Peace Science* 25: 1, 1-18.
- Morrow, James D. 2010: The Next Wave: Bringing Together Modeling and Data, in: *Conflict Management and Peace Science* 27: 4, 299-307.
- Morrow, James D./Siverson, Randolph M./Tabares, Tressa E. 1998: The Political Determinants of International Trade: The Major Powers, 1907-90, in: *American Political Science Review* 92: 3, 649-661.
- Neumann, Iver B. 2008: Discourse Analysis, in: Klotz/Prakash 2008: 61-77.
- Oneal, John R./Russett, Bruce 2005: Rule of Three, Let it Be? When More Really is Better, in: *Conflict Management and Peace Science* 22: 4, 293-310.
- Rihoux, Benoit 2006: Two Methodological Worlds Apart? Praises and Critiques from a European Constructivist, in: *Political Analysis* 14: 3, 332-352.
- Russett, Bruce/Starr, Harvey/Kinsella, David 2006: *World Politics: The Menu for Choice*, 8. Auflage, Boston, MA.
- Schneider, Gerald/Gleditsch, Nils Petter/Carey, Sabine 2011: Forecasting in International Relations: One Quest, Three Approaches, in: *Conflict Management and Peace Science* 28: 1, 5-14.
- Schneider, Gerald/Ruoff, Gabriele 2010: Quantitative Methoden, in: Masala, Carlo/Sauer, Frank/Wilhelm, Andreas (Hrsg.): *Handbuch der Internationalen Politik*, Wiesbaden.
- Snidal, Duncan 2004: Formal Models of International Relations, in: Sprinz/Wolinsky-Nahmias 2004, 227-264.
- Sprinz, Detlef F./Wolinsky-Nahmias, Yael (Hrsg.) 2004: *Models, Numbers & Cases: Methods for Studying International Relations*, Ann Arbor, MI.
- Van Evera, Stephen 1997: *Guide to Methods for Students of Political Science*, New York, NY.
- Ward, Michael D./Greenhill, Brian D./Bakke, Kristin M. 2010: The Perils of Policy by p-Value: Predicting Civil Conflict, in: *Journal of Peace Research* 47: 4, 363-375.

Zeitschrift für Internationale Beziehungen 2011: Webseite, in: <http://www.zib.nomos.de/>;
29.04.2011.