

Die Geschichte des Streitkräftevergleichs in der internationalen Politik: Machtvergleiche und die Macht des Vergleichens

Der Beitrag leistet eine historische Rekonstruktion von Praktiken des Streitkräftevergleichs an der Schnittstelle zwischen Internationalen Beziehungen (IB) und Geschichtswissenschaft. Konzeptionell baut er auf einem Verständnis von Gleichgewicht als zentralem Beobachtungsschema und symbolisch generalisiertem Kommunikationsmedium in der Konstruktion und Reproduktion von Weltpolitik auf. Der Hauptteil des Beitrages zeichnet die Entwicklung von Streitkräftevergleichen als eine zentrale Beobachtungs- und Vergleichspraxis von und zwischen Staaten vom späten 18. Jahrhundert bis in die Gegenwart nach. Wir fragen wer vergleicht, wer verglichen wird, was verglichen wird, und wie die Vergleiche durchgeführt werden. Es wird herausgearbeitet, dass es sich bei Praktiken des Streitkräftevergleichs um Orte der fortwährenden Aushandlung und Kontestation über das „Wer“, „Was“ und „Wie“ handelt. Diese historische Rekonstruktion leistet dabei einen Beitrag zum besseren Verständnis, wie Systemeigenschaften auf der Makroebene des internationalen Systems durch Praktiken widergespiegelt und konstituiert werden, und damit zur historisch-soziologischen Forschung zur Entwicklung des modernen weltpolitischen Systems.

1. Einleitung¹

„Was Friedrich Nietzsche vor 150 Jahren als Signum einer neuen Epoche diagnostiziert hatte, ist mittlerweile zum Normalfall geworden. Im heutigen ‚Zeitalter der Vergleichung‘ wird praktisch alles einem Vergleich unterzogen und oft auch in eine Rangfolge gebracht: Krankenhäuser, Schulen, Unternehmen und Staaten, Sportler, Politikerinnen, Wissenschaftler und Angestellte, Produkte, Dienstleistungen und Feriendestinationen“ (Heintz 2016: 305).

Auch in den Internationalen Beziehungen (IB) ist das Vergleichen allgegenwärtig. Die außenpolitischen Interessen verschiedener Staaten, historische Zustände des internationalen Systems, die Konfliktaffinität unterschiedlicher Regierungsformen etc.: ein Großteil der Untersuchung internationaler Beziehungen beruht darauf, dass ein Vergleichsmerkmal (*tertium comparationis*) dazu dient, bestimmte Vergleichsgegenstände (*comparata*) synchron oder diachron aufeinander zu beziehen –

1 Dieser Beitrag entstand im Rahmen des von der Deutschen Forschungsgemeinschaft (DFG) geförderten Bielefelder Sonderforschungsbereichs 1288, „Praktiken des Vergleichens. Die Welt ordnen und verändern“. (Teilprojekt A01, „Comparing Forces and the Forces of Comparison“). Für wichtige Hinweise zur Überarbeitung einer früheren Fassung dieses Aufsatzes danken wir Thomas Müller sowie den GutachterInnen der zib.

sei es in umfangreichen Datensätzen, sei es in wenigen Fallstudien – um hierdurch Einsichten in Regelmäßigkeiten oder Veränderungsprozesse zu gewinnen.

Was mit dem *Vergleich als Methode* in den IB jedoch kaum oder meistens nur implizit untersucht wird, sind *Vergleiche als ubiquitäre soziale Praxis* selbst. Dabei kann im Anschluss an eine zunehmend umfangreiche Forschung in benachbarten Disziplinen davon ausgegangen werden, dass das Vergleichen nicht nur eine mehr oder weniger arbiträr eingesetzte Operation darstellt. Vielmehr handelt es sich beim Vergleichen selbst um eine grundlegende kognitive, soziale Relationierung überhaupt erst ermöglichende Operation, während die Herausbildung komplexer und institutionalisierter Vergleichspraktiken ein wesentliches Charakteristikum der Moderne darstellt, von dem kein Bereich der sozialen Welt ausgenommen bleibt. Es besteht insoweit überhaupt kein Anlass zu der Vermutung, dass es sich nicht auch bei den internationalen Beziehungen um einen von Vergleichspraktiken durchdrungenen und in weiten Teilen sogar konstituierten Bereich handeln sollte.

Der vorliegende Beitrag hat sich zum Ziel gesetzt, anhand des Streitkräftevergleichs die Geschichte des modernen weltpolitischen Systems in Form der Entwicklung einer bestimmten Form von Vergleichspraktiken zu rekonstruieren. Diese Rekonstruktion ist insofern exemplarisch, als dass sie nur eine von vielen solcher Formen untersucht, dabei allerdings mit dem Streitkräftevergleich auf eine Vergleichspraxis zurückgreift, ohne die insbesondere etwa die für das moderne weltpolitische System prägende vergleichende Beobachtung von Staaten untereinander im Hinblick auf ihre relative Macht (bzw. ein *Gleichgewicht der Mächte*) kaum denkbar wäre. Der Beitrag argumentiert dabei, dass der analytische Wert der (historischen) Untersuchung von Vergleichspraktiken in den internationalen Beziehungen zum einen darin besteht, einen anderen Blickwinkel auf die Entwicklung des weltpolitischen Systems zu werfen. Zum anderen besteht der analytische Wert insbesondere auch darin zu beleuchten, dass es sich bei den in Vergleichspraktiken sichtbar werdenden und/oder hergestellten Relationierungen keineswegs nur um Formen einfachen Beobachtens handelt, sondern diese überhaupt nur als Zusammenspiel und Entwicklung unterschiedlicher Vergleichstechniken und –institutionen, der damit verbundenen Vergleichsinteressen, sowie konkreter Auseinandersetzungen und Aushandlungsprozesse um das Vergleichen zu verstehen sind. Neben dem Anliegen, anhand einer historischen Rekonstruktion von Praktiken des Streitkräftevergleichs einen Beitrag zum Verständnis eines wichtigen Aspekts der Entwicklung des modernen weltpolitischen Systems zu leisten, möchte der vorliegende Aufsatz auch einen Beitrag zum besseren Verständnis von Streitkräftevergleichen als einer oft ausgeübten, aber kaum in ihrer Funktion reflektierten Praxis leisten.

Der folgende Abschnitt führt kurz in die aktuelle Forschung zu Vergleichspraktiken ein, um diese dann auf Fragen der historischen Konstitution und Entwicklung des weltpolitischen Systems zu beziehen. Hierbei werden Streitkräftevergleiche als besondere Form der Vergleichspraxis im Rahmen von Machtvergleichen identifiziert, wobei letztgenannte insbesondere seit dem Wiener Kongress vor allem die Form von systembezogenen, routinierten und teilweise institutionalisierten Gleichgewichtsbeobachtungen annehmen.

Vor diesem Hintergrund unternimmt der darauffolgende Teil des vorliegenden Beitrags auf Grundlage intensiver Archivrecherchen eine schlaglichtartige historische Rekonstruktion der Entwicklung von Streitkräftevergleichen. Bei dieser Rekonstruktion kann es selbstverständlich nicht darum gehen, eine umfassende Geschichte des Streitkräftevergleichs zu schreiben. Vielmehr soll die Entwicklung und die Rolle des Streitkräftevergleichs zunächst anhand dreier Leitfragen vorgenommen werden: Wer hat verglichen? Wer bzw. was wurde verglichen? Wie wurde verglichen? Der dann folgende Abschnitt bezieht diese historische Rekonstruktion auf die vorangegangenen theoretisch-konzeptionellen Überlegungen zur Rolle von Vergleichspraktiken in der Entwicklung des modernen weltpolitischen Systems. Dabei geht es zum einen darum, den aus der Bearbeitung der *Wer*-, *Was*- und *Wie*-Fragen gewonnenen Überblick einerseits theoriegeleitet zur Frage danach zu verdichten, warum verglichen wurde und wird. Zum anderen geht es darum darzustellen, dass es sich bei Praktiken des Streitkräftevergleichs um Orte der fortwährenden Aushandlung und Kontestation über das *Wer*, *Was* und *Wie* handelt. Der abschließende Abschnitt resümiert hieraus folgend, dass Vergleichspraktiken nicht nur an sich sowie in ihrer Rolle im Kontext (der Entwicklung) des weltpolitischen Systems ein vielversprechendes Forschungsfeld darstellen, sondern insbesondere auch als geeignet erscheinen, die Verschränkungen von Strukturbildungs- bzw. Systemwandlungsprozessen auf verschiedenen Ebenen darzustellen.

2. Vergleichsbeobachtungen und Machtvergleiche im modernen weltpolitischen System

Wie einleitend bereits mit Bezug auf den Aufsatz von Bettina Heintz erwähnt, sind Vergleiche und deren Verbreitung ein wesentliches Kennzeichen der Moderne. Unterschiedliche Praktiken des Vergleichs stellen in diesem Sinne einen spezifisch modernen Weltzugang dar (vgl. Epple/Erhart 2015), was sich seit einigen Jahren in einer Zunahme entsprechender Forschungen in den Geistes- und Sozialwissenschaften, insbesondere auch im deutschsprachigen Raum, widerspiegelt.² Von zentraler Bedeutung dabei ist jedoch, dass Vergleichspraktiken und die in ihnen operierenden Vergleiche auf bestimmte Kriterien bezogene Relationen bzw. sich darin widerspiegelnde Rangordnungen nicht nur abbilden, sondern diese in weiten Teilen ebenfalls reproduzieren und damit konstituieren: Sportler, Staaten, Universitäten etc. nehmen ihren Platz in einer Tabelle gemeinhin nicht nur zur Kenntnis, sondern suchen diesen zu verbessern (Werron 2012). Diese gleichzeitig soziale Strukturen dar- wie herstellende Funktion von Vergleichen wurde dabei in den IB bislang vor

2 Vergleiche und Vergleichsbeobachtungen waren das Thema eines federführend von Bettina Heintz entworfenen, in der Exzellenzrunde 2011 nur sehr knapp gescheiterten Clusterantrags mit einer primär soziologischen und geschichtswissenschaftlichen Forschungsagenda. Seit 2017 setzt sich an der Universität Bielefeld der Sonderforschungsbereich, in dessen Rahmen der vorliegende Aufsatz entstanden ist, mit „Praktiken des Vergleichens“ auseinander.

allem in Bezug auf die Zunahme numerischer Vergleiche (vgl. etwa Broome/Quirk 2015) sowie in Bezug auf Status und Hierarchien im internationalen System untersucht. Während innerhalb der Statusliteratur die enorme Relevanz von Vergleichen und vom Vergleichen der AkteurInnen für die Bestimmung von Status sowie die Rückwirkung dieses Vergleichens auf die Statushierarchie betont wird, bilden die Praktiken des Vergleichens der AkteurInnen selbst keinen prominenten Untersuchungsgegenstand (vgl. etwa Jaschob 2018; Renshon 2017). Wenngleich insgesamt unbestritten sein dürfte, dass die Bedeutung numerischer Vergleiche in allen Lebensbereichen zunimmt, läuft ein enger Fokus nur auf die Anzahl dieser Vergleiche Gefahr, zweierlei zu übersehen: *Einerseits* gehen mit der Einführung und Etablierung bedeutsamer numerischer Vergleiche üblicherweise längerfristige Aushandlungsprozesse und Durchsetzungspolitiken einher. Nur in den seltensten Fällen werden wichtige Referenzgrößen einfach als Maßzahl in einem technisch-bürokratischen Akt festgelegt, nur um dann eine allseits akzeptierte Bezugsgröße abzugeben. Die jüngere Forschung hat diese komplexen Prozesse etwa in Bezug auf die Einführung und Etablierung des Bruttoinlands- bzw. Bruttosozialproduktes in historisch-soziologischen Arbeiten sichtbar gemacht (vgl. Speich 2013). *Andererseits* lässt sich zwar leicht in vielen Lebensbereichen eine Zunahme von Quantifizierung und numerischen Vergleichen beobachten, diese stehen aber nicht an und für sich, sondern sind immer in komplexere Vergleichspraktiken eingebunden. Üblicherweise stellen numerische Vergleiche zusätzlich und jenseits zu einer spezifischen Vergleichsbeobachtung partielle Aspekte in Bezug auf einen komplexeren Vergleichszweck dar: so mag der Medaillenspiegel bei den Olympischen Winterspielen zwar durchaus im engeren Sinne auch die jeweiligen nationalen Trainingsbedingungen und staatlichen Unterstützungsleistungen etwa für EiskunstläuferInnen oder BobfahrerInnen widerspiegeln; er vergleicht aber auch Staaten insgesamt in Bezug auf nationales Prestige und stellt dieses dabei teilweise auch mit her.³

Die im Folgenden vorgenommene historische Rekonstruktion der Entwicklung von Streitkräftevergleichungen erfolgt vor allem im letztgenannten Sinne, nämlich als Ausdruck *komplexer* vergleichender Machtbeobachtungen im weltpolitischen System. Die dabei herauszuarbeitenden Aushandlungsprozesse und Kontestationen im erstgenannten Sinne weisen dabei besonders darauf hin, dass es auch bei Streitkräftevergleichungen nie nur um die Darstellung von *real existierenden* Machtverhältnissen geht, sondern immer auch um deren Herstellung. Obwohl kaum umstritten ist, dass es sich bei Macht um ein *relationales* Konzept handelt und obwohl Machtvergleiche im Allgemeinen und Streitkräftevergleiche im Besonderen bereits seit geraumer Zeit durchgeführt werden – einschließlich vielfältiger Diskussionen darüber, wie sie am besten durchzuführen sind – waren die eigentlichen Vergleichspraktiken

3 Es geht hier allgemein um die Vergleichsbeobachtung von Nationalstaaten in Bezug auf Leistungsprestige, das letztendlich von den einzelnen Elementen des Vergleichs fast vollständig entkoppelt ist: der überwiegenden Mehrheit der den Medaillenspiegel Rezipierenden und darauf nationale Prestigedifferenzen Zuschreibenden (vgl. Emrich u.a. 2015) wird mutmaßlich nur schwerlich ein gesteigertes Interesse für die Eigenheiten etwa des Viererbobwettbewerbs unterstellt werden können.

und ihre Entwicklung bislang nicht Gegenstand der geschichtswissenschaftlichen bzw. der Literatur in den IB.⁴

Wir suchen mit der Rekonstruktion dieser Vergleichspraktiken damit nicht nur einen Beitrag zur Geschichte von Streitkräftevergleichen selbst zu leisten, sondern vor allem auch dazu, wie sie Machtbeobachtungen als eine zentrale Vergleichspraxis im weltpolitischen System mit hergestellt haben. Der konkrete analytische Mehrwert besteht dabei darin sichtbar zu machen, wie strukturelle Entwicklungen auf einer Makroebene des weltpolitischen Systems untrennbar mit Vergleichspraktiken als wichtigen Formen sozialer Praxis auf einer Mesoebene verknüpft sind, ohne dabei der einen oder anderen Ebene einen analytischen oder kausalen Vorrang zuzusprechen zu können oder zu müssen.⁵ Wir sehen die Geschichte von Streitkräftevergleichen in diesem Sinne als Baustein einer noch zu leistenden, historisch-soziologischen Detailrekonstruktion der Geschichte des modernen weltpolitischen System, wobei wir fraglos annehmen, dass es sich beim Vergleich von Streitkräften als zentralem Ausdruck militärischer Machtverhältnisse um einen der in diesem Zusammenhang größten Bausteine handeln dürfte: Das Vergleichen von Streitkräften erscheint als zentraler Bestandteil der Art und Weise, wie Staaten sich selbst in Bezug auf andere Staaten im weltpolitischen System verorten und vorstellen. Hierdurch tragen sie aktiv zur Konstitution und Ausgestaltung dieses Systems bei. Die *Evolution* dieser Vergleichspraxis vollzieht dabei die Komplexitätsentwicklung des modernen politischen Systems auf eigene Weise nach: Welten liegen zwischen der Aufzählung der Anzahl von Soldaten und Kanonen in den Briefen Napoleons auf der einen Seite und den komplexen Modellen zur Berechnung strategischer Stabilität und von Zweitschlagfähigkeiten im Kalten Krieg auf der anderen Seite – ganz zu schweigen von den gegenwärtigen Herausforderungen, militärische Streitkräfte hinsichtlich ihrer Fähigkeiten zur Kriegsführung im Cyberspace zu vergleichen (Willett 2019).⁶

Als eine sowohl im System generalisierte als auch für das System konstitutive spezifische Form zur Beobachtung und Kommunikation stellt sich dabei die Figur des

4 Im Rahmen einer überbordenden Literatur zu Macht in den Internationalen Beziehungen (IB) sei als guter Überblick und anspruchsvoller konzeptioneller Zugang hier nur Guzzini (2013) genannt.

5 Wir verwenden diese Ebenenunterscheidung als gegenstandsbezogene: im Kontext des weltpolitischen Systems stellt insofern die System- die Makroebene dar, die Handlungsbildet die Mikroebene. Institutionen, Praktiken, Organisationen konstituieren die darauf bezogene Mesoebene.

6 Wir möchten an dieser Stelle betonen, dass der vorliegende Beitrag vor allem auf eine historische Rekonstruktion von Praktiken des Streitkräftevergleichs an der Schnittstelle zwischen IB und Geschichtswissenschaft abzielt. Es geht dabei ausdrücklich nicht um einen Beitrag dazu, wie am besten zu vergleichen wäre, noch geht es um einen Beitrag zu Debatten um die Definition von *Macht* in den IB. In rein forschungspraktischer Hinsicht soll der Beitrag dabei durchaus sowohl die Möglichkeiten als auch die Schwierigkeiten einer wirklichen Kooperation zwischen GeschichtswissenschaftlerInnen und IBlerInnen verdeutlichen (im Gegensatz zur Aneignung von Ergebnissen der einen durch die jeweils andere Disziplin).

Machtgleichgewichts (Balance of Power) dar.⁷ *Balance of Power* ist nicht alles – aber es stellt ein zentrales Charakteristikum des weltpolitischen Systems dar, dass Staaten sich selbst und in Bezug auf andere *immer auch* in Termini eines Gleichgewichts beobachten (Albert 2016: 107-116). Die *Balance of Power* in diesem Sinne ist Beobachtungsform und Kommunikationsmedium (und nicht gleichsam Newtonsches Naturgesetz, wie etwa im politischen Realismus teilweise dargestellt). Unbeschadet der Tatsache, dass die Literatur über das *balancing* als Staatskunst bereits in der zweiten Hälfte des 18. Jahrhunderts Hochkonjunktur hatte und unbeschadet der Tatsache, dass Gleichgewichtspolitik in einem breit verstandenen Sinne mutmaßlich Bestandteil aller historisch bekannten Interaktionen zwischen verschiedenen Herrschaftsräumen ist (Buzan/Little 2000), erlangt die *Balance of Power* erst mit dem Wiener Kongress die Form einer routinierten und (anfänglich etwa im statistischen Ausschuss des Kongresses) institutionalisierten Form *systemischer* Selbstbeobachtung. Damit bleibt unbestritten, dass die vergleichende (Selbst-)Beobachtung im Hinblick auf Macht immer schon ein (auch semantisch) konstitutives Attribut von *Mächten* darstellt. Allerdings ging es dabei über lange Zeit selten zentral und kaum systematisch um das *Zählen* von Fähigkeiten eigenständiger Einheiten. Obwohl grundsätzlich schon in der Antike vorzufinden, spielen numerische Vergleiche vor der Moderne eine durchgängig eher untergeordnete Rolle. Vergleichsbeobachtungen zwischen *HerrscherInnen* waren vor allem in zeremoniell und rituell institutionalisierte Ordnungen eingebunden und durch diese begrenzt. Obwohl etwa auch in Vergleichsbeobachtungen unter HerrscherInnen im Alten Reich *harte* (sprich: materielle) Faktoren eine Rolle spielten, beruhten sie vor allem auf der ausführlichen Herausbildung von Machtritualen und –zeremonien, deren Entwicklung und Differenzierung sie umgekehrt antreiben. Im Anschluss an Barbara Stollberg-Rilinger (2015) könnte man in diesem Sinne durchaus argumentieren, dass die vergleichende Beobachtung symbolischer Bedeutungen sogar konstitutiv für politische Ordnung war. Erst die sich über einen langen Zeitraum hinziehende Konsolidierung territorialer Staatlichkeit, zusammen mit der Bindung territorialer Staatlichkeit an ein exklusives Souveränitätsverständnis sowie der Entwicklung rationaler Bürokratie als Mittel zur Organisation weitläufiger Besteuerung und Aushebung im modernen Staat ab Mitte des 18. Jahrhunderts (ausführlich bei Kluegling 1986), lässt quantitativen Maßzahlen eine herausragende Bedeutung für Vergleichsbeobachtungen zukommen (ohne dabei Ritual und Zeremonie, etwa im Rahmen diplomatischen Protokolls, je ganz zu ersetzen).

Wie bereits erwähnt stellen zunehmend systematische, zunehmend detaillierte und zunehmend quantifizierende Vergleichsbeobachtungen ein wesentliches Charakteristikum der Moderne bzw. der modernen Gesellschaft dar (vgl. Heintz/Werron 2011). Hinter oftmals vordergründig scheinbar *technischen* Vergleichen verbergen sich dabei jedoch oftmals sowohl eine lange Geschichte wie auch eine spezifische Funktion: Streitkräftevergleiche sind Praktiken, welche die Art und Weise, wie zeit-

7 Mit *formal* meinen wir hier, dass *Macht* im hier gebrauchten Sinne unabhängig von der Anzahl verfügbarer inhaltlicher Bestimmungen operiert und eher als generalisiertes Medium sozialer Interaktion (im Sinne Talcott Parsons) oder als symbolisch generalisiertes Kommunikationsmedium (in Niklas Luhmanns Anpassung) zu verstehen ist.

genössische BeobachterInnen die jeweilige internationale Ordnung wahrnehmen, sowohl widerspiegeln als auch selbst beeinflussen. Sie stellen eine wichtige Basis für spezifische Machtzuschreibungen an AkteurInnen dar (einschließlich, im Falle von Staaten, Selbstzuschreibungen). Die historische Entwicklung von Streitkräftevergleichen kann insofern als lange und komplizierte Geschichte von Bemühungen und Auseinandersetzungen darüber angesehen werden, was gezählt werden soll und wie es gezählt werden soll, wie klassifiziert und kategorisiert werden soll, wer als autoritative Stimme und neutrale Instanz in Bezug auf Vergleichsstandards akzeptiert wird, und an wen – und zu welchem Zweck – Vergleiche adressiert werden.

Gerade vor diesem Hintergrund erfordert es die Untersuchung der Entwicklung von Streitkräftevergleichen jedoch, sie nicht einfach „nur“ als ein Beispiel für einen allgemeinen Trend aufzufassen, sondern bereits an dieser Stelle auf eine Reihe von besonderen Rahmenbedingungen und Umständen von Streitkräftevergleichen hinzuweisen. So mag es auf den ersten Blick plausibel sein zu vermuten, dass Vergleichsbeobachtungen hin zu *harten* und vermeintlich einfacher zu quantifizierenden Daten tendieren. Da aber gerade im militärischen Bereich viele nur schwer quantifizierbare Größen eine Rolle spielen, wie etwa Taktik, Disziplin, Ausbildung, oder auch Status,⁸ ist hier ein besonders intensiver Blick darauf erforderlich, wie und zu welchem Zweck verglichen wird. Darüber hinaus kommt es nicht nur darauf an, die letztendlich vergleichenden AkteurInnen sowie die sich im Laufe der Zeit wandelnden Akteurskonstellationen zu beleuchten, sondern auch zu fragen, für welche spezifischen Publika Vergleiche durchgeführt werden (und zu welchen spezifischen Zwecken diese Publika adressiert werden). So macht es mutmaßlich einen Unterschied, ob Streitkräftevergleiche zum Zweck einer wissenschaftlichen Analyse oder zum Zweck der Legitimierung einer Erhöhung von Militärausgaben durchgeführt werden. Schließlich sind bei Streitkräftevergleichen noch Besonderheiten hinsichtlich der Materialbasis zu berücksichtigen. Im Gegensatz zu anderen Vergleichsgegenständen sind verlässliche Daten zu Streitkräften oftmals weder öffentlich verfügbar noch etwa werden sie (wie etwa das Bruttosozialprodukt) wesentlich zum Zwecke des Vergleichs produziert. Erhebliche Unterschiede bestehen in diesem Zusammenhang zwischen Demokratien und Nicht-Demokratien, da es vor allem die Rolle von Parlamenten in der Verabschiedung des Staatshaushaltes bzw. in der Ratifikation von internationalen Verträgen ist, welche dazu führt, dass entsprechende Daten öffentlich verfügbar werden. Streitkräftevergleiche zeichnen sich also sowohl durch ein gleichsam intrinsisches Problem der Validität und Reliabilität der Daten als auch durch ein Kategorisierungs- bzw. Standardisierungsproblem aus (herausragend sichtbar etwa in Auseinandersetzungen im Rahmen von Rüstungskontrollverhandlungen bei der Frage, welche Waffensysteme mit welchen Waffensystemen gegengerechnet werden sollen).

8 Für das, was heute unter *Status* in den IB verstanden wird (sowohl Platz innerhalb als auch Zugehörigkeit zu einer Gruppe), werden in den hier verwendeten Quellen - vor allem denen des 18., 19. und größtenteils auch denen des 20. Jahrhunderts - verschiedene Wörter wie *Prestige*, *Rang* oder *Stellung* benutzt. Der in der IB-Statusliteratur teilweise diskutierte Unterschied zwischen Status und Prestige ist hier nicht von Bedeutung, da *Status* im Untersuchungskontext für den größten Teil des Untersuchungszeitraums keine Quellsprache ist. Zur Diskussion um Status und Prestige vgl. Jaschob (2018: 53-55).

3. Praktiken des Streitkräftevergleichs

3.1 Wer hat verglichen?

Die Praktiken des Streitkräftevergleichs weisen über die Zeit einen steigenden Institutionalierungsgrad auf. Ein großer Teil dieser Institutionalisierung bestand im Wandel von einer unsystematischen und sporadischen zu einer systematischen Praxis des Datensammelns, sowohl in Bezug auf die Beobachtung der eigenen als auch in Bezug auf die Beobachtung fremder Streitkräfte. Zu regelmäßigen Verzerrungen in der Datenbasis kam es dabei bei der Beobachtung der eigenen Streitkräfte vor allem infolge von Informationsdefiziten im bürokratischen Prozess, während bei der Beobachtung anderer Streitkräfte naturgemäß die Verfügbarkeit der Rohdaten selbst ein größeres Problem darstellt.

Es können im Wesentlichen drei mit Streitkräftevergleichen und dem vorausgehenden Beobachten und Aggregieren von Daten befasste Akteursgruppen unterschieden werden: staatliche Akteure, internationale Organisationen und nichtstaatliche AkteurInnen. Vor allem die Kategorie der nichtstaatlichen AkteurInnen ist dabei breit und umfasst ein Spektrum von Think Tanks über einzelne WissenschaftlerInnen bis hin zu den Massenmedien.⁹

Bis ins 18. und frühe 19. Jahrhundert fanden sich Streitkräftevergleiche vor allem in den Schriften der Souveräne (d.h. Monarchen), die noch selbst an der Planung und Durchführung von Feldzügen beteiligt waren, etwa bei König Friedrich II von Preußen oder bei Erzherzog Joseph II von Österreich (vgl. etwa Friedrich II 1775: 203-204; Joseph II 1761). Eine frühe und wichtige Datenquelle für Streitkräftevergleiche waren dabei das diplomatische Personal sowie Offiziere, die in fremden Streitkräften dienen oder gedient hatten (Kloosterhuis/Bahl 2015: 407-412, 479).

Ab Ende des 18. Jahrhunderts begannen darüber hinaus sogenannte *Staatenkundler*, als Teil der Etablierung statistischer Wissensbestände über Staaten, auch Daten über die Streitkräfte europäischer Staaten zu publizieren (vgl. z.B. Sprengel 1793; vgl. ausführlich Kluebing 1986) und diese teilweise zu interpretieren und zu vergleichen (vgl. z.B. Horsley 1744; Kolb 1860; Coxe 1806). Datenlieferanten waren dabei amtliche Mitteilungen, Korrespondentennetzwerke, sowie Staatsbedienstete selbst (vgl. Kluebing 1986: 40-42, 63-64).

Mit der zunehmenden Bürokratisierung und Professionalisierung des modernen Staates und der daraus folgenden Gründung von General- bzw. Admiralstäben wurden diese (oder äquivalente Institutionen/Abteilungen in den Ministerien) zunehmend zu den relevanten staatlichen Akteuren des Streitkräftevergleichs (während

9 In der jüngeren Vergangenheit scheinen Privatfirmen (wie etwa IHS Jane's) eine größere Rolle zu spielen, die sich auf die Analyse von Rüstungsmärkten spezialisieren, um diese Analysen öffentlich anzubieten oder im Auftrag von Rüstungsfirmen und Regierungen anzufertigen. Der Zugang auch zu den öffentlich verfügbaren Analysen zu Forschungszwecken ist jedoch aufgrund der damit üblicherweise verbundenen extrem hohen Kosten praktisch nicht gegeben.

die Rolle der Souveräne in dieser Hinsicht unwichtiger wurde):¹⁰ dabei produzierten diese Institutionen nicht nur Streitkräftevergleiche, sondern stellen sie auch anderen Einrichtungen des Staates bzw. des Militärs auf Anforderung zur Verfügung. Ihre regelmäßige Beobachtung und Berichterstattung fungierte dabei als Haupttriebkraft hinter der Institutionalisierung von Vergleichen und begründete zum Teil langlebige Traditionen (siehe etwa das britische *Parliamentary Paper* – später *Command Paper* – über die *Fleets of Great Britain and Foreign Countries*,¹¹ oder die jährlichen österreichischen Denkschriften des Generalstabes des k. u. k. Heeres (Landstreitkräfte Österreich-Ungarns) von 1888 bis 1914 (vgl. Conrad Chef des Generalstabes 1893–1914)). Auch das vorgelagerte Datensammeln erlebte im Laufe des 19. Jahrhunderts einen Institutionalisierungsschub. Der deutliche Anstieg in der Anzahl dauerhafter diplomatischer Delegationen in anderen Ländern während des 19. Jahrhunderts beförderte hier die gegenseitigen Beobachtungskapazitäten. Von besonderer Bedeutung war hierbei die zunehmende Praxis der Entsendung von Militär- und Marineattachés als Teil des diplomatischen Korps (Persson 2010: 46-65; Vagts 1967; Tachikawa 2015). Die Datenquellen, aus denen die Attachés ihre Informationen bezogen, waren entweder offizielle (und offiziöse)¹² Publikationen über die nationalen Streitkräfte (etwa in Jahrbüchern oder Dokumenten in Bezug auf Haushaltsplanungen) und damit öffentlich zugänglich,¹³ oder aber informelle persönliche Netzwerke und Treffen (unterstützt durch einen *esprit de corps* oder durch Adelsbeziehungen). Daneben wurden Daten (vor allem in Krisen- und Kriegezeiten) durch verschiedene Formen der Spionage gesammelt, welche wieder-

-
- 10 Der preußische Generalstab mit seiner spezifischen Organisation wurde ab der Mitte des 19. Jahrhunderts von vielen anderen nationalen Streitkräften als Vorbild gesehen. Frankreich, Russland, Japan und andere Staaten versuchten ihre Generalstäbe nach diesem Modell zu gestalten. Auch dies war die Folge intensiven gegenseitigen Beobachtens und Vergleichens (vgl. Herrera/Mahnken 2003: 224-225).
 - 11 Der Titel des Papiers wechselte einige Male. Es wurde zwischen 1886 und 1939 (außer im Ersten Weltkrieg) jeweils vor der Debatte über die *Navy Estimates* im Parlament veröffentlicht. Es diente nach Aussage der britischen Admiralität als „the official document upon which M.P.'s (and others) rely for the latest information regarding British and foreign fleets“ (vgl. Admiralty 1933).
 - 12 Beispielsweise veröffentlichte das *Jahrbuch für Deutschlands Seeinteressen* (auch bekannt als *Nauticus*) sowohl Daten bezüglich der deutschen Seeinteressen (Außenhandel, Handelsschifffahrt, etc.) als auch zu den Seestreitkräften der „größeren Seemächte“ (unter anderem zu den Marinebudgets, den Budgets für Landesverteidigung, dem Personalbestand und der Anzahl und Spezifikationen der Kriegsschiffe) einschließlich einer „vergleichenden Zusammenstellung der Kriegsflotten der größeren Seemächte“ (vgl. *Nauticus* 1908: 501-595). Es wurde von einem anonymen *Nauticus* herausgegeben und suggerierte dadurch, eine private Publikation zu sein. Allerdings stand dahinter nicht nur finanziell sondern auch als Herausgeber das Nachrichtenbüro des Reichsmarineamtes. Auch die Artikel wurden von Mitarbeitern des Büros verfasst (vgl. Bönker 2012: 209).
 - 13 1831 fragte der preußische Generalstab offiziell beim preußischen Außenministerium um Hilfe bei der Sammlung von Informationen bezüglich der Streitkräfte fremder Staaten an. Um das diplomatische Personal zu instruieren, wurde eine detaillierte und systematische Liste mit Informationen verteilt, an denen der Generalstab interessiert war. In Antwort auf die Anfrage des Generalstabs erwiderte der preußische Botschafter in London, dass alle gewünschten Informationen in militärischen Zeitschriften und den Anhängen zu den Parlamentsdebatten zu finden seien. Alle Bände der letzteren seien im Besitz des preußischen Innenministeriums (vgl. von Krauseneck 1831).

rum selbst durch die Entstehung der (militärischen) Geheimdienste institutionalisiert wurde (Seligmann 2006; Jackson/Siegel 2005; Warner 2014; Grawe 2017: 27-108).

Im Feld der nicht-staatlichen AkteurInnen beschäftigten sich ab dem letzten Drittel des 19. Jahrhunderts im Bereich der Wissenschaft auch Rechtsgelehrte mit Streitkräftevergleichen. Sie verglichen zwar nicht selbst, äußerten sich aber in den aufkommenden Diskussionen über die Möglichkeiten von Abrüstung und Rüstungsbeschränkung im Völkerrecht auch zu den Herausforderungen der Bestimmung relativer militärischer Macht und relativer Rüstungsstände (Keefer 2016: 99-104). Mit Publizisten und Journalisten betrat Ende des 19. Jahrhunderts ein weiterer nicht-staatlicher Akteurstypus die Bühne. Streitkräftevergleiche wurden nunmehr in Tageszeitungen, Zeitschriften, Pamphleten und Büchern veröffentlicht. Es ist kaum rekonstruierbar, inwieweit Publizisten und Journalisten die Vergleiche selbst produzierten bzw. von staatlicher Seite verfügbare in einem anderen Format reproduzierten. Oft wurden Daten etwa aus statistischen Jahrbüchern oder aus Angaben in Parlamentsdebatten zusammen- bzw. direkt von den entsprechenden staatlichen Stellen zu Verfügung gestellt (vgl. Biefang 2009: 65-96; Bönker 2012: 208-209).

Nach dem Ende des Ersten Weltkrieges befasste sich der Völkerbund als erste internationale Organisation mit Streitkräftevergleichen. Der Völkerbund sammelte und veröffentlichte im jährlichen Turnus offizielle Daten über die Streitkräfte seiner Mitglieder. Diese wurden zum einen 1922 und 1923, in Vorbereitung auf eine mögliche Abrüstungskonferenz, in einer umfassenden statistischen Erhebung der Rüstung verschiedener Staaten zusammengeführt (*Statistical Enquiry*; vgl. League of Nations 1922, 1923). Zum anderen wurden sie jährlich (allerdings in weniger komplexer Form als im *Statistical Enquiry*) im *Armaments Year-Book* veröffentlicht. Die Jahrbücher enthielten außerdem Angaben über die Militärausgaben der Großmächte (und zogen daraus den Schluss, dass diese sich auf Krieg vorbereiteten) und für einige Länder auch Angaben über die „Exports of Arms and Implements of War“ (vgl. League of Nations 1939/1940: 69; Webster 2005: 501).

Nach dem Ende des Zweiten Weltkrieges waren es auf der einen Seite die *North Atlantic Treaty Organization* (NATO) und auf der anderen Seite der Warschauer Pakt, welche Streitkräftevergleiche sowohl produzierten als auch als wichtige Instrumente in (öffentlichen) Debatten und Verhandlungen einsetzten (Soviet Union 1982–1987; NATO 1982, 1984, 1988; Krakau/Diehl 1989; Herzog/Wildgen 1986). Dabei wurden die von der NATO verwendeten Daten vor allem von US-Nachrichtendiensten bereitgestellt (vgl. Garthoff 1990). Nichtsdestotrotz verwendete die NATO für die Veröffentlichung von Streitkräftevergleichen in den 1970er Jahren nur öffentlich verfügbare externe Daten, denn: „for security reasons, the figures appearing in the following tables have been taken from independent sources which have already been made public“ (NATO Information Service 1972: 20). Erst in den 1980er Jahren begann die NATO, in diesem Zusammenhang eigene Daten zu verwenden, während die Sowjetunion in veröffentlichten Streitkräftevergleichen auf einen Mix von Datenquellen setzte, vor allem auch auf Daten des *International In-*

stitute for Strategic Studies (IISS): „As before, besides the data supplied by competent Soviet authorities, the book cites information given by the London and other institutes for strategic studies, and also by official American sources“ (Soviet Union 1982–1987; hier: 2. Aufl. 1984: 5).

Die *United Nations Organization*/Vereinten Nationen (UNO) als Nachfolgeorganisation des Völkerbundes begann erst im Jahre 1981 mit der Sammlung und Veröffentlichung von Daten zu den Militärausgaben ihrer Mitglieder.¹⁴ Ähnlich wie beim Völkerbund im Jahre 1924 bestand hier der Hauptzweck darin, Transparenz im Interesse von Spannungsabbau und der Unterstützung von Abrüstungsbemühungen herzustellen (vgl. United Nations 1989: 15).

Während des Kalten Krieges betrat mit den Think Tanks – und vor allem mit dem *Stockholm International Peace Research Institute* (SIPRI) und dem IISS – eine neue Kategorie von Akteuren die Bühne des Streitkräftevergleichs. Während diese Think Tanks ihre eigenen Streitkräftevergleiche veröffentlichten und die Veröffentlichungen etwa als Datenquellen in Veröffentlichungen des Warschauer Paktes dienten, blieben diese Vergleiche nicht unangefochten. So bestand etwa ein Hauptgrund dafür, dass sich die NATO in den 1980er Jahren auf eigene Daten stützte, in einer expliziten Ablehnung der IISS-Daten (vgl. Auswärtiges Amt 1979; Bundesministerium der Verteidigung 1979; Ostrich/Green 1981-2).

Zusammenfassend ist festzuhalten, dass im Laufe der Zeit drei Akteursgruppen in der Streitkräftevergleichen zugrunde liegenden Datensammlung sowie in der Herstellung von Vergleichen und in deren Gebrauch eine Rolle spielen:

- Staatliche Akteure wie Souveräne (Monarchen), General- und Admiralstäbe, Nachrichtendienste, Militär- oder Marineattachés, etc.;
- Internationale Organisationen wie der Völkerbund, die Vereinten Nationen, NATO, der Warschauer Pakt;
- Nichtstaatliche Akteure aus Wissenschaft und Think Tanks sowie JournalistInnen und PublizistInnen.

Wichtige Faktoren, die einen Wandel innerhalb der Akteursgruppen bedingten bzw. neue AkteurInnen hervortreten ließen, waren neben Professionalisierung und Institutionalisierung innerhalb staatlicher Organisationen auch gesellschaftliche Wandlungsprozesse wie die zunehmende Bedeutung der Presse für die internationale Politik, die Entstehung internationaler Organisationen, sowie die zunehmende Bedeutung von Think Tanks im Verwissenschaftlichungstrend im Kalten Krieg.

3.2 *Comparata und Tertia*

Wahrnehmungen von Machtverteilungen und –verschiebungen, vorherrschende Konfliktkonstellationen sowie geteilte Vorstellungen darüber, was als wichtiges militärisches Machtmittel gilt, bildeten und bilden eine wichtige Grundlage dafür, wer oder was in Streitkräftevergleichen verglichen wird und wurde – ob sie etwa dazu

14 Siehe <http://www.un-arm.org/MilEx/Home.aspx>; 25.03.2020.

dienen, globale oder nur regionale Verteilungen zu beurteilen, ob es um den Vergleich von Land-, Luft-, oder Seestreitkräften geht, oder aber im Kalten Krieg darum, das nukleare oder das konventionelle Gleichgewicht zu beurteilen. Streitkräftevergleiche wurden (und werden) immer auch vom Wandel der oben genannten Faktoren beeinflusst: in Zeiten von Machtverschiebungen, Veränderungen der Konfliktkonstellationen und (militär-)technologischen Neuerungen veränderten sich auch die Praktiken des Streitkräftevergleichs.

3.2.1 Wer wurde verglichen?

In Streitkräftevergleichen geht es nicht um einen Vergleich *jeder mit jedem*, sondern um die Erfassung von relativen militärischen Fähigkeiten und von Macht innerhalb einer „relevant comparison group“ (Renshon 2017: 3). Diese Gruppe wurde (und wird) sowohl durch den Status eines Staates (Groß-, Mittel- oder Kleinmacht) bestimmt, als auch durch geographische, meist regionale Bezugsräume, die sich entweder aus dem Kontext militärischer Operationsplanung im Blick auf zukünftige Kriege, oder aber aus den Subsystemen des internationalen Systems, auf die sich die Vergleiche bezogen, ergaben. Daher variiert die Definition einer Region je nach Vergleich, insbesondere gelten hier andere Maßstäbe für Land- und für Seestreitkräfte.

Bei der Untersuchung der Praktiken des Streitkräftevergleichs von Großmächten (bzw. nach 1945 von beiden Militärblöcken) zeigt sich, dass diese immer die Hauptvergleichsgruppe füreinander bildeten, während Mittelmächte und Kleinstaaten eher in regionalen Vergleichen der Großmächte eine Rolle spielten, oder aber in direkten gegenseitigen Vergleichen.

Streitkräftevergleiche sind deshalb immer auch ein Abbild dessen, wer als system- bzw. teilstreitkraftrelevante Großmacht betrachtet wurde. Dabei wurden Streitkräftevergleiche nicht genutzt, um zu bestimmen, wer Teil der Großmachtgruppe war oder werden sollte. Wer jedoch als Großmacht von den anderen anerkannt wurde,¹⁵ war in den meisten Fällen auch Teil von systemweiten Vergleichen. Streitkräftevergleiche sind daher Spiegel für Machtverschiebungen, da letztere zu Veränderungen in den Vergleichen führen. Teilweise sind sie aber auch Indikator für anstehende systemweite Machtverschiebungen, da Staaten in Streitkräftevergleichen als relevante Militärmächte auftauchen können, bevor sie allgemein als Großmächte anerkannt wurden. Streitkräftevergleiche verdeutlichen daher nicht nur, wer Teil der Gruppe ist, sondern auch, wie die Machtverteilung und die Hierarchie innerhalb dieser Gruppe wahrgenommen werden.

15 Jost Dülffer (2003: 56) bringt dies auf den Punkt: „Großmacht war nur die Macht, die von den anderen Großmächten als solche akzeptiert wurde“. Und dies unabhängig davon, ob entsprechende Faktoren wie militärische Macht faktisch Großmachtstatus bedeuten würden (vgl. auch Jaschob 2018: 38; Ward 2017: 35-47).

Bis zum Ende des 19. Jahrhunderts standen im Mittelpunkt von Streitkräftevergleichen zumeist jeweils mindestens zwei der Großmächte, d.h. Frankreich, Großbritannien, Russland, Preußen/das Deutsche Reich, oder Österreich(-Ungarn) (ab ca. 1867 auch Italien; vgl. Dülffer 2003: 57), manchmal unter Berücksichtigung von deren jeweils kleineren Verbündeten. Diese Berücksichtigung spiegelte dabei bestehende Allianzen wider und verdeutlichte, wer als dominierende Großmacht einer Region betrachtet wurde. So verglich beispielsweise Friedrich II im Jahr 1775 Frankreich, Spanien, Neapel und Sardinien auf der einen Seite mit Preußen, Großbritannien, Österreich und den Niederlanden auf der anderen Seite (Friedrich II 1775: 203-204). Ende des 19. Jahrhunderts wurden wiederum in einer französischen Publikation Frankreich und Russland auf der einen Seite mit dem Dreibund (Deutsches Reich, Österreich-Ungarn, Italien) auf der anderen Seite verglichen (Hayard 1892).¹⁶

Auch bei Seestreitkräftevergleichen tritt der Bezug auf bestimmte Mächtegruppen sowie regionale Bezugsräume in den Vergleichen deutlich hervor. Während sich die maritimen Großmächte Frankreich und Großbritannien als antizipierte Gegner miteinander verglichen (Corry 1858), verglichen sich die maritimen Mittelmächte mit stärkerem Regionalbezug. Für Preußen lag beispielsweise der Fokus in den 1860er Jahren auf der Ostsee, mit Dänemark und Russland als möglichen zukünftigen Gegnern (Roon 1862).

Die unterschiedlichen regionalen Bezüge sowie die lange bestehende Differenzierung zwischen Vergleichen von Land- und von Seestreitkräften verdeutlichen den Umstand, dass europäische Großmächte nicht notwendigerweise auch maritime Großmächte waren. Preußen etwa zählte seit Mitte des 18. Jahrhunderts zu den europäischen Großmächten, eine maritime Großmacht wurde das Deutsche Reich aber erst zu Beginn des 20. Jahrhunderts.¹⁷

Der kolonial-imperialistische Wandel von einem europazentrierten zu einem globalen System Ende des 19. Jahrhunderts, verbunden mit der Aufnahme Japans und der Vereinigten Staaten (USA) in die Gruppe der Großmächte (etwa Dülffer 2003), beinhaltete dann auch einen signifikanten Wandel hinsichtlich der Frage, wer verglichen wurde. Allerdings bestand der Unterschied zwischen Land- und Seestreitkräftevergleichen auch danach zu einem gewissen Grad weiter. Landstreitkräftevergleiche, vor allem die der europäischen Großmächte, blieben weiterhin europazentriert,¹⁸ während zum Beispiel der U.S.-Generalstab schon zu Beginn des 20. Jahr-

16 Einen guten Überblick über die Geschichte der Großmächte und die Allianzsysteme des 18. Jahrhunderts bietet Heinz Duchhardt (1997). Zum 19. Jahrhundert siehe Bridge/Bullen (2005).

17 Der Großmachtstatus Preußens ging bei Reichsgründung auf das Deutsche Reich über, das als Rechtsnachfolger Preußens gilt (Jaschob 2018: 149).

18 Ende des 19. Jahrhunderts bildeten sich, so Jost Dülffer (2003: 61-63), zwei neue Subsysteme innerhalb des globalen internationalen Systems („Weltstaatsensystem“) heraus, die „nicht als Ausweitung des europäischen Mächtesystems“ (ebd.: 63) gesehen werden können. Man kann demnach argumentieren, dass auch das europäische System als Subsystem innerhalb des globalen Systems bestand und dies im Hinblick auf militärische Macht durch Landstreitkräfte für die europäischen Großmächte weiterhin entscheidend blieb.

hunderts in Publikationen die Armeen der europäischen (das russische Kaiserreich, das Deutsche Reich, Frankreich, Österreich-Ungarn, Italien und Großbritannien), sowie außereuropäischen Großmächte (USA und Japan) miteinander verglich (U.S. War Department 1911).

Auf der anderen Seite spiegelte die zunehmend prominente Praxis systemweiter Vergleiche von Seemächten sowohl die Globalisierung des weltpolitischen Systems als auch die damit einhergehende Vergrößerung der Gruppe der maritimen Großmächte wider. Ausschlaggebend für letzteres war zunächst nicht die Erweiterung der Großmächtegruppe per se, sondern dass im imperialistischen Wettbewerb um Einfluss, Rohstoffe, Territorien und den Zugang zu Märkten die Bedeutung von Seemacht für den Großmachtstatus zunahm und Schlachtfлотen zu einem relevanten „Gradmesser für Großmacht“ wurden (Dülffer 2003: 52; vgl. Bönker 2012: 24-27). Daraus folgte der Ausbau der Schlachtfлотen in Staaten wie dem Deutschen Reich, den USA und Japan, die zuvor keine signifikanten Seestreitkräfte besaßen (etwa Friedberg 1988: 136-139; 152-167).

Diese Machtverschiebung beeinflusste zum einen, bedingt durch neue Konfliktkonstellationen, wer in regionalen Vergleichen berücksichtigt wurde. Beispielsweise wurde das Deutsche Reich um 1906 zum antizipierten Hauptgegner Großbritanniens in einem möglichen Krieg in der Nordsee und löste darin Frankreich ab, mit dem Großbritannien nach jahrhundertelanger Rivalität 1904 ein Abkommen schloss, welches die Beziehungen nachhaltig besserte. Dies bedeutete nicht, dass man Frankreich nicht weiterhin beobachtete. Doch der *two-power-standard*, welcher der britischen Marinerüstungspolitik zugrunde lag, wandelte sich zunächst dahingehend, wer mit einbezogen wurde (von Frankreich und dem russischen Kaiserreich zu Frankreich und dem Deutschen Reich) und wurde dann im Laufe der 1910er Jahre durch einen Machtstandard abgelöst, der, unter anderem bezogen auf die Nordsee, nur noch das Deutsche Reich berücksichtigte (vgl. Parkinson 2008: 45, 70-71, 76-79; Friedberg 1988: 144-152, 172-174, 199-207; Jaschob 2018: 235; Churchill 1912: 1555). Zum anderen tauchten, bedingt durch den Wandel zum globalen System, neben regionalen Vergleichen nunmehr systemweite globale Vergleiche der maritimen Großmächte hinsichtlich ihrer Seemacht auf, häufig auch in der Form von Rankings. Großbritannien bezog die U.S. Navy beispielsweise spätestens ab 1896 und die japanische Marine ab 1898 in maritime Machtvergleiche mit ein (Awdry 1896; Sittart Neale 1898). Geht man davon aus, dass Streitkräftevergleiche Indikatoren für Machtverschiebungen sind, so zeichnete sich mit der Berücksichtigung Japans und der USA in diesen Vergleichen eine Entwicklung ab, die im Falle Japans erst 1904/05 abgeschlossen war: die volle Anerkennung als Mitglied der Gruppe der Großmächte (vgl. Dülffer 2003: 61).

Das Beispiel der Seestreitkräfte verdeutlicht auch die Machtverschiebung durch den Ersten Weltkrieg. Noch 1904 verglich die britische Admiralität in *Estimated Distribution of Naval Power for 1904* Frankreich, Russland, Großbritannien, Japan, die USA und das Deutsche Reich (Horowitz 2010: 148), während sich die *Comparison of Strengths of British and Foreign Navies 1919/1923* auf Großbritannien, Frankreich, Italien, Japan und die USA bezog (Navy Board and Admiralty 1919).

Das Deutsche Reich und das Russische Kaiserreich (ab 1917 die Sowjetunion) hatten ihren Status als maritime Großmächte durch den Ersten Weltkrieg verloren. Italien, das zuvor zwar europäische Großmacht gewesen war, aber nicht immer in systemweiten Seestreitkräftevergleichen berücksichtigt wurde, wurde ab 1919, trotz eines primär regionalen Einflusses, zu einem ständigen Teil dieser Vergleiche. Auch die sich andeutende Rückkehr der Sowjetunion und des Deutschen Reichs in die Gruppe der Großmächte spiegelte sich bereits ab 1921 in britischen Seestreitkräftevergleichen wider (vgl. Murray 1921).

Die nächste Machtverschiebung, die das grundlegende Design von Streitkräftevergleichen beeinflusste, fand nach dem Zweiten Weltkrieg statt. Der Wandel von einem multipolaren zu einem bipolaren System, dominiert von den USA und der Sowjetunion, führte dazu, dass häufig nur diese beiden Länder bzw. die NATO und der Warschauer Pakt verglichen wurden. Dabei ist auch hier zwischen regionalen und global angelegten Vergleichsräumen zu unterscheiden, wiederum bedingt durch den jeweils spezifischen Vergleichskontext. Während beispielsweise im Zuge nuklearer Rüstungskontrolle die Sowjetunion und die USA hinsichtlich ihrer systemweiten nuklearen Macht verglichen wurden, handelte es sich im Kontext konventioneller Rüstungskontrolle um regionale Vergleiche, in denen hauptsächlich die NATO und der Warschauer Pakt verglichen wurden (vgl. Goldblat 2002). In der vom IISS herausgegebenen *Military Balance* wurden im Kalten Krieg zwar Daten zu den Streitkräften aller Staaten aufgelistet. Die Kapitel, die dezidiert die *Balance of Power* behandelten, bezogen sich allerdings auf Ost-West-Vergleiche. Dabei wurden für das konventionelle Gleichgewicht (meist auf Europa bezogen) häufig die Streitkräfte der NATO und des Warschauer Paktes, für das strategisch-nukleare Gleichgewicht (global) die USA und die Sowjetunion verglichen (etwa International Institute for Strategic Studies 1976; 1979; 1981; 1983).

3.2.2 Was wurde verglichen?

Während Streitkräftevergleiche vor dem Zweiten Weltkrieg vor allem dahingehend unterschieden werden konnten, ob es sich um Vergleiche von Land- oder Seestreitkräften (bzw. später auch von Luftstreitkräften) handelte, entstanden im Kalten Krieg schnell unterschiedliche Vergleichsformen, die sich jeweils auf konventionelle bzw. nukleare militärische Fähigkeiten bezogen. Dies beeinflusste nicht nur wer verglichen wurde, sondern danach unterschied sich auch, was verglichen wurde. Darüber hinaus hing das *Was* innerhalb dieser einzelnen Gruppen von drei Faktoren ab: der für die zu einem spezifischen Zeitpunkt für die jeweilige Teilstreitkraft (oder militärischen Fähigkeiten) als am wichtigsten betrachteten Waffensysteme, dem Zweck des Streitkräftevergleichs sowie der angelegten Dimensionen militärischer Macht. In Bezug auf letztere lassen sich drei Dimensionen unterscheiden. Erstens, die quantitative Dimension, in der sich etwa Truppenstärke oder die Anzahl von Schiffen oder Gefechtsköpfen finden. Zweitens, die qualitative Dimension: hier findet sich eine Reihe von relativ einfach messbaren Elementen, die sich

vor allem auf spezifische Charakteristika von den in der quantitativen Dimension berücksichtigten Elemente beziehen, wie etwa die Tonnage oder Bewaffnung von Schiffen, das Kaliber von Geschützen, die Reichweite und Geschwindigkeit von Raketen und Artillerie, etc. Deutlich schwieriger quantifizierbar sind in dieser Dimension Elemente wie Taktik, Organisation, Rekrutierungssystem, Ausbildungsstand, Disziplin oder Moral. In der dritten Dimension militärischer Macht finden sich jene Elemente, die nicht direkt den Streitkräften selbst zuzuordnen sind, wie etwa Staatsfinanzen, geostrategische Position, oder aber Größe von Territorium oder Bevölkerung.

Während sich die (Wahrnehmung der) Bedeutung verschiedener Dimensionen militärischer Macht nur graduell über die Zeit veränderte, hing die Einschätzung der Bedeutsamkeit von bestimmten Waffensystemen immer von den Interessen der jeweils involvierten AkteurInnen, der vorherrschenden Strategie, der antizipierten relevanten Statusattribute (vgl. Jaschob 2018: 169-170; Renshon 2017: 119-120), sowie vom technologischen Wandel ab. Das *Was* in Streitkräftevergleichen ist also rückgebunden an und eingebettet in verschiedene, nicht nur das Militär prägende, Wandlungsprozesse, wie technologische Entwicklungen oder Vorstellungen über die Status und Macht bedingenden bzw. darstellenden Faktoren.

In der Gesamtschau einer historischen Rekonstruktion lässt sich feststellen, dass im 18. Jahrhundert und der ersten Hälfte des 19. Jahrhunderts hauptsächlich Truppenstärken sowie die Anzahl an Kanonen und Schiffen verglichen wurden, wobei der Vergleich auch qualitativer Elemente (vor allem Truppenstruktur und Taktik) verstärkt nach den napoleonischen Kriegen hinzukommt (vgl. Herrera/Mahnken 2003: 205-217). Ab der Mitte des 19. Jahrhunderts wurde das, was verglichen wurde, zunehmend von technologischen Innovationen beeinflusst. Dies betraf zum einen neue und verschiedene Waffentypen und Transportmittel (Eisenbahn), zum anderen die technologische Qualität von Waffen. Die rasanten technologischen Entwicklungen waren dabei auf zwei Ebenen relevant: zum einen stand die Befürchtung, ohne angemessene Reaktion auf technologische Entwicklungen militärisch ins Hintertreffen zu geraten, da zunehmend der technologische Stand als auch die Qualität der Waffen den Ausgang zukünftiger Kriege bestimmen würden (ausführlich bei Förster 2016; vgl. auch Herrera/Mahnken 2003: 217-241). Zum anderen führte der rasante Wandel dazu, dass – so die Überzeugung – nur diejenigen Großmacht sein konnten, die die ökonomischen und industriellen Fähigkeiten besaßen, um mit der technologischen Entwicklung Schritt zu halten (Buzan/Lawson 2013: 629). Dabei war man sich in militärischer Hinsicht immer bewusst, dass die qualitativ-technologische Überlegenheit allein nicht notwendigerweise zum Sieg führen würde, zumal sich die Qualität der Waffen phasenweise weitgehend angeglichen hatte. Ein Sieg würde nur mit entsprechender Ausbildung, Disziplin und Moral, sowie mit einer der technologischen Entwicklung angepassten Taktik möglich sein und auch die Truppenstärke blieb weiterhin der wichtigste, wenn auch in Teilen relativierte Faktor im Vergleich von Heeresmacht (Storz 2003: 217-220; 227-228; Wettstein 2016: 216-218).

Die Anzahl der in den Vergleichen zu berücksichtigenden Faktoren stieg somit zunehmend an, mit der Folge einer steigenden Komplexität von Vergleichen, einschließlich damit einhergehender Vergleichbarkeitsprobleme (vgl. Seligmann 2016: 33-34; Wettstein 2016: 168-180).

Im Bereich der Seestreitkräftevergleiche wandelte sich das, was verglichen wurde, im Laufe des langen 19. Jahrhunderts aufgrund verschiedener technologischer Entwicklungen sowie einhergehend mit sich verändernden Vorstellungen über Macht und Status im internationalen System. Die fünf wichtigsten technologischen Neuerungen, welche die Marinen prägten, waren der Dampfantrieb, die Schiffschraube, die Sprenggranate, die gezogenen Geschütze, sowie die Panzerung (Hobson 2004: 29). In einer Übergangsphase wurden noch Segel- und Dampfschiffe gleichzeitig verglichen (vgl. Corry 1858), doch jede neue Generation von Schiffen führte dazu, dass die vorhergehenden aus militärischer Sicht an Wert verloren (Buzan/Lawson 2013: 631). In den Vergleichen spiegelte sich dies darin, dass nicht nur die Tonnage einer Schiffskategorie spezifiziert wurde, sondern auch, wie alt die Schiffe maximal waren (vgl. etwa Reichsmarineamt 1897). Besonders deutlich wurde dies nach dem sogenannten *Dreadnoughtsprung* 1906. Die HMS *Dreadnought* war das erste Schlachtschiff mit rein schwerer Artillerie einheitlichen Kalibers und ein Teil der britischen Strategie, die maritime Überlegenheit nicht nur durch die Anzahl, sondern auch durch die Qualität der Schiffe zu bewahren (Seligmann 2016: 29-30). Ihr Stapellauf führte dazu, dass alle älteren Schlachtschiffe aufgrund ihrer qualitativen Unterlegenheit an Relevanz verloren und man auch in den Seestreitkräftevergleichen häufig begann, bei den Schlachtschiffen zwischen Schiffen der *Dreadnought*-Klasse und anderer, älterer Schiffsklassen zu unterscheiden (vgl. etwa Office of Naval Intelligence 1909).

Die zunehmende Bedeutung von Seemacht für den globalen Großmachtstatus beeinflusste ab den 1890er Jahren grundlegend, was verglichen wurde. Im militärisch-strategischen Aushandlungsprozess über die Frage, was einen zukünftigen Krieg und damit auch Seemacht bestimmen würde, setzte sich in den 1880er und 1890er Jahren die Strategie der Seeschlacht und Blockade durch. Schlachtschiffe und zunehmend auch Große Kreuzer (später auch sogenannte Schlachtkreuzer) waren nicht nur relevante militärische Faktoren von Seemacht, welche im Krieg eine Blockade und eine siegreiche Seeschlacht ermöglichen sollten, sowie im Frieden nationale territoriale und ökonomische Expansion durch abschreckende Wirkung absicherten, sondern Schlachtschiffe wurden auch zum Statussymbol einer Großmacht (ausführlich beispielsweise bei Hobson 2004; Bönker 2012). Daraus folgte, dass diese Schiffstypen die Hauptkategorien bildeten, die in systemweiten Seestreitkräftevergleichen verglichen wurden (vgl. etwa Reichsmarineamt 1897). Die enorme Relevanz vor allem von Schlachtschiffen blieb, trotz der Erfahrungen des Ersten Weltkriegs, auch in der Zwischenkriegszeit bestehen, was sich in den Rüstungskontrollverhandlungen der *Washington Naval Conference* sowie auch in internen Diskussionen über maritime Machtstandards, beispielweise der britischen Admiralität widerspiegelte (vgl. Tooze 2014: 21-23; Admiralty 1919-1921).

Mit der Einführung von Nuklearwaffen entstand dann nach dem Zweiten Weltkrieg eine grundlegende Zweiteilung hinsichtlich dessen, was verglichen wurde (vgl. Prados 1982; Freedman 2003). Es ging und geht nicht mehr nur um Waffensysteme, Flugzeugtypen oder Schiffsklassen, sondern um die Unterscheidung zwischen nuklearen und konventionellen Fähigkeiten. Verglichen wurden nun auch die Anzahl nuklearer Sprengköpfe sowie deren Trägersysteme als relevanter Faktor militärischer Macht (etwa Soviet Union 1987; NATO 1982). Entscheidend für das *Was* in Streitkräftevergleichen war im Kalten Krieg, ob die konventionelle oder die nukleare Balance beurteilt werden sollte. Die bereits im 19. Jahrhundert begonnene Komplexitätssteigerung setzte sich während des Kalten Krieges weiter fort (etwa Rohn 1990). Aktuell deutet das Aufkommen von Cyberkriegsführung auf einen weiteren Entwicklungssprung in der Komplexität von Streitkräftevergleichen und damit verbundene Vergleichbarkeitsprobleme hin.

3.3 Wie wurde verglichen?

Grundsätzlich können zwei Formen von Streitkräftevergleichen voneinander unterschieden werden: *einfache* Vergleiche und *komplexe* Vergleiche. Mit einfachen Vergleichen ist gemeint, dass entweder nur ein Element militärischer Macht verglichen wird, der Vergleich also innerhalb einer bestimmten Kategorie verbleibt, oder aber mehrere Kategorien verglichen, aber nicht miteinander verknüpft werden. Als klassische Darstellung kann beispielsweise die Zeitungsillustration von *roten* und *blauen* Soldaten, Panzern, Artillerie und Flugzeugen gelten. Auch wenn einfache Vergleiche zumeist das widerspiegeln, was allgemein als *Erbsenzählen* bezeichnet wird, handelt es sich dabei nicht notwendigerweise um numerische Vergleiche. So ist etwa der Vergleich der Moral verschiedener Armeen ein einfacher, allerdings kein üblicherweise in Zahlen ausgedrückter Vergleich. Im Gegensatz dazu werden in einem komplexen Vergleich zwei oder mehr Staaten hinsichtlich von zwei oder mehr Elementen militärischer Macht verglichen. Diese Vergleiche können sowohl rein numerische Vergleiche sein, oder numerische Vergleiche, die durch ein qualitatives Element relativiert werden. Komplexes Vergleichen beinhaltet nicht nur die Herausforderung zu entscheiden, was in einem bi- oder multilateralen Kontext zu berücksichtigen ist, sondern auch festzulegen, wie die unterschiedlichen Elemente gewichtet und verrechnet werden sollen (von Sass 2018: 2-3).¹⁹

Beide Vergleichsmethoden haben ihre Vor- und Nachteile, die immer wieder thematisiert wurden. Einfache numerische Vergleiche hatten den Vorteil, dass sie für

19 Hinsichtlich der Unterscheidung einfach/komplex ist anzumerken, dass diese *nicht* in eine Unterscheidung quantitativ/qualitativ übersetzt werden kann. Es existieren keine „rein“ quantitativen oder „rein“ qualitativen Vergleiche. So würde es beispielsweise keinen Sinn machen, nur die Moral von SoldatInnen zu vergleichen und dabei komplett die Frage zu ignorieren, wie viele SoldatInnen es sind. Auf der anderen Seite lassen sich etwa Militärausgaben sehr einfach in einer einzigen Zahl ausdrücken – aber die nicht enden wollende Debatte darüber was Militärausgaben sind, ist nicht nur Ergebnis fehlender haushaltsbuchhalterischer Standardisierung zwischen Staaten (vgl. Albert/Müller 2019).

ExpertInnen wie LaiInnen verständlich sind. Die die Vergleiche produzierenden und verwendenden AkteurInnen waren sich jedoch dessen bewusst, dass einfache Vergleiche nicht ausreichen, um relative militärische Fähigkeiten und Macht zu bewerten (Collins/Cordesman 1978: xvi-xix, 5-8; etwa auch Hall 1987; Kolb 1860). Es gab daher schon immer Versuche, einfache Vergleiche durch die Anreicherung mit weiteren Vergleichselementen aussagekräftiger zu machen, sie gleichsam hin zu einem komplexen Vergleich weiterzuentwickeln. So verglich Friedrich II. im 18. Jahrhundert beispielsweise nicht nur die Anzahl der Soldaten, sondern auch die Finanzen der Staaten, um den einfachen numerischen Vergleich hinsichtlich der tatsächlichen Kriegsfähigkeit zu relativieren. Darüber hinaus wurden nicht einfach zu quantifizierende qualitative Elemente oder nicht-militärische Elemente berücksichtigt: Diese wurden entweder versucht zu quantifizieren (vgl. Macnamara 1919) oder der Vergleich der Truppenstärke oder anderer quantitativer Elemente wurde durch den Verweis auf Moral, Disziplin, Ausbildung etc. relativiert (vgl. Kolb 1860; Goldhamer 1979; Wettstein 2016).

Die Grenzen einfacher numerischer Vergleiche wurden ab Ende des 19. Jahrhunderts immer häufiger thematisiert. Dies wurde zum einen durch die zunehmende Verwendung von numerischen Streitkräftevergleichen in den Rüstungsetatdebatten der Großmächte - vom deutschen Reichskanzler Leo von Caprivi (1891: 3112) als „Zahlenwut“ bezeichnet, die auch in Frankreich kritisch gesehen werde²⁰ - bedingt, zum anderen durch den Beginn multinationaler²¹ Rüstungskontrollbestrebungen.

In Vorbereitung auf Rüstungskontrollverhandlungen bemühten sich intern alle Staaten um komplexe Vergleiche, die versuchten, alle für die eigene nationale Sicherheit als notwendig erachteten Anforderungen und Elemente zu berücksichtigen, und „to discuss how a ratio of strength necessitated by each requirement can be established“ (Macnamara 1919: 38). Auch für nichtstaatliche Akteure, wie das IISS im Kalten Krieg, war es möglich, die Berechnung des konventionellen Gleichgewichts in Europa durch „comparison of the strengths of both men and equipment“ mit “consideration of qualitative characteristics, factors such as geographical advantages, deployment, training and logistic support, and of differences in doctrine and philosophy” zu ergänzen. Darüber hinaus mussten für eine aussagekräftige Beurteilung auch das strategische nukleare Gleichgewicht und die relative Stärke der Seestreitkräfte mitberücksichtigt werden (Institute for Strategic Studies 1976: 97).

In den Verhandlungen selbst war es jedoch kaum möglich, sich darauf zu einigen, was in den Vergleichen berücksichtigt werden sollte und wie die einzelnen Elemente zu gewichten wären (siehe 4. Vergleichszwecke und -debatten). Auch komplexe Vergleiche hatten demnach ihre Grenzen.

20 Er bezeichnet weiter „dieses Rechnen mit den Zahlen nicht ganz unbedenklich“. Die Zunahme numerischer Vergleiche wird von den Akteuren aktiv wahrgenommen und kommentiert. Caprivi beginnt beispielsweise den Teil seiner Rede mit dem Satz „man hat in der Welt sich vielfach gewöhnt, Armeen nach ihrer Zahl zu schätzen“.

21 Es gab bereits im 18. und 19. Jahrhundert Beispiele für bilaterale Rüstungsbeschränkung, wie das englisch-französische Marineabkommen von 1787 oder das Rush-Bagot Abkommen von 1817. Ausführlich bei Keefer (2016).

Ein weiterer Faktor, der die Vergleichsmethoden beeinflusste, waren technologische Innovationen, vor allem in Gestalt der Einführung von Computern. Dabei erleichterte die zunehmende Rechenleistung zumindest auf technischer Ebene komplexe Vergleiche. Durch neue, computerbasierte Gefechts- und Kriegssimulationen und Methoden des *operations research* entstanden neue Vergleichspraktiken, die hochkomplexe Simulationen und Modelle zur Grundlage hatten (Rohn 1990; Lutz 1986). Ein weiteres Beispiel für eine Methode komplexen Streitkräftevergleichens ist das während des Kalten Kriegs von Andrew Marshall entwickelte Verfahren des *Net Assessment*. Ziel dieses Verfahrens ist es, quantitative durch qualitative Einschätzungen zu ergänzen, um mittel- bis langfristige Trends zu antizipieren und deren Auswirkungen auf das militärische Kräfteverhältnis abzuschätzen (vgl. Roberts/Kaushal 2018).

4. Vergleichszwecke und -debatten: Warum wurde verglichen?

Die vorhergehende überblicksartige Rekonstruktion hat bereits an mehreren Stellen deutlich gemacht, dass es auch bei Praktiken des Streitkräftevergleichs nie um eine einfache empirische Bestandsaufnahme im Sinne einer *Inventur* ging, welche dann gleichsam als Grundlage für evidenzbasierte politische und militärische Diskurse herangezogen wurde. Dass, wie, was und von wem verglichen wurde hing von einer Reihe unterschiedlicher, nicht unbedingt eindeutig voneinander getrennter, Motive ab. Gerade die historische Entwicklung und Verquickung verschiedener Vergleichsformen zeigt dabei jedoch auf, dass Streitkräftevergleiche einen Strukturwandel des weltpolitischen Systems nie alleine nachzeichnen, sondern in Teilen auch mit beeinflussen und dass es sich vor allem auch um Praktiken handelt, welche verschiedene Ebenen von Ordnungsbildung zu verknüpfen in der Lage sind.

Dass Streitkräftevergleiche nie nur eine Realität nachzeichnen, sondern diese in vielen Teilen mit beeinflussen, zeigt sich vorrangig daran, dass Debatten um das „richtige“ Vergleichen untrennbarer Bestandteil der Geschichte von Streitkräftevergleichen sind. *Was* und *wie* gezählt wird beeinflusst selbstverständlich unmittelbar auch die Schlussfolgerungen, die aus diesen Vergleichen bezüglich von Machtverhältnissen gezogen werden (vgl. Seligmann 2016: 33-34; Widenmann 1909: 131-136; Rohn 1990: 4; Collins/Cordesman 1978: xviii, 5-8).²² Unterschiedliche

22 Unsicherheiten über die Korrektheit von Daten oder aber die (un-)bewusste Verwendung von falschen Zahlen (um etwa in Rahmen von Aufrüstungsdebatten eine bestimmte Agenda zu verfolgen), ob in einfachen oder komplexen Vergleichen, führten durchaus zu Fehleinschätzungen militärischer Fähigkeiten mit auch zeitgenössisch wahrgenommenen destabilisierenden Effekten (vgl. z.B. Rose 2013: 111-117). Vor diesem Hintergrund wird die bedeutende Rolle verständlich, welche dem offiziellen Datenaustausch als Ziel in Rüstungskontrollverhandlungen seit dem frühen 20. Jahrhundert zukam. So betrachtete nicht zuletzt der Völkerbund sein *Armaments Year-Book* als wichtiges Mittel zur Stabilisierung: die Datenpublikation „directly affects Disarmament and the maintenance of Peace“. Sie erschaffe „among the general staffs concerned that feeling of confidence which is a result of open dealings, instead of the mutual distrust engendered by indirect and surreptitious methods of obtaining information“ (League of Nations 1924: 4).

Schlussfolgerungen hinsichtlich der relativen militärischen Macht des potentiellen Gegners aufgrund von unterschiedlichen Kategorisierungen können etwa gerade in Phasen des Wettrüstens destabilisierende Effekte haben und die Spannungen zwischen Staaten erhöhen. Ein anschauliches Beispiel liefert die Auseinandersetzung, welche sich im Jahre 1909 der deutsche Marineattaché mit Sir John Fisher, dem damaligen *First Sea Lord* Großbritanniens lieferte: es ging darum, welche Schiffe in die Kategorie der Dreadnoughts gezählt werden sollten und welche Schiffe in welcher Kategorie wie klassifiziert werden sollten. Nachdem Fisher erklärte, dass „Blücher‘ ja 28cm Schnellfeuerkanonen [hat] und gleichwertig mit einem Panzerschiff [ist]“, entgegnete der Marineattaché:

„Es tut mir leid, wenn ich da zur Aufklärung ein Geheimnis verraten muß; aber ‚Blücher‘ hat nicht 28 cm, sondern nur 21 cm. Im Übrigen ist er, selbst wenn er 28 cm hätte, nur 15000 t groß, und wenn Sie ihn mitrechnen, müssen Sie mindestens Ihre acht ‚King Edwards‘ auch mitrechnen, und dann ist der von Mr. M‘ Kenna angestellte Vergleich zwischen der Anzahl der englischen und deutschen Dreadnoughts erst recht falsch“ (Widenmann 1909: 132-133).

Die Debatten über die richtigen Vergleiche wurden jedoch nicht nur über reine Zahlen oder Kategorisierungen geführt, sondern auch über die Frage, welche qualitativen Elemente in einen Vergleich mit einbezogen werden sollten. Beispielsweise verglich 1914 der *First Lord of the Admiralty*, Winston Churchill, deutsche und britische Dreadnoughts hinsichtlich ihrer Anzahl und ihrer Feuerkraft, während der *Second Sea Lord*, Lord Jellicoe, diese hinsichtlich ihrer Anzahl und ihrer Tonnage verglich. Während Churchill aus seinem Vergleich den Schluss zog, dass die Royal Navy einen Vorteil gegenüber der deutschen Marine habe, warnte Lord Jellicoe auf Grundlage seines Vergleichs davor, die deutsche Marine zu unterschätzen (Seligmann 2016: 33-34).

In Debatten um Kategorisierungen im Kontext von Rüstungskontrollverhandlungen ging es nicht nur um die Frage, was in welcher Kategorie zu zählen wäre, sondern auch darum, wie gegebenenfalls eine Kategorie durch die andere ausgeglichen werden kann. So wurde etwa vom U.S.-Delegationsvorsitzenden vor der *London Naval Conference* von 1930 vorgeschlagen

„that each ton displacement by a light cruiser up to 5000 tons be rated at 90 per cent of the value of each ton displaced by a destroyer, and that each ton displaced by a heavy cruiser of between 5000 and 10000 tons be rated 90 per cent of the value of each ton of a light cruiser“ (Hall 1987: 66).

Durchsetzen konnte sich dieser Vorschlag aufgrund unterschiedlicher Ansichten bezüglich der Gewichtung der einzelnen Schiffstypen jedoch nicht.

Darüber hinaus wurde nicht selten über die richtige Formel gestritten, die nicht nur militärische Fähigkeiten bewerten würde, sondern auch die spezifischen Sicherheitsbedürfnisse eines Staates berücksichtigte. Dieses Formelproblem stellte einen der Hauptstreitpunkte in praktisch jeder Rüstungskontrollverhandlung seit den Haager Konferenzen 1899 dar (Keefer 2016; Dülffer 1981: 13).

Im Rückblick auf die Haager Konferenzen von 1899 (im Kontext der Friedensinitiative von Papst Benedikt XV) stellte ein internes Memorandum des deutschen Auswärtigen Amtes von 1917 fest:

„Auf den Haager Friedenskonferenzen ist die Frage einer internationalen Beschränkung der militärischen Rüstungen eingehend erörtert worden; doch ist es damals nicht gelungen, eine Formel zu finden, die geeignet wäre, der verschiedenen politischen und geographischen Lage der einzelnen Länder sowie der Verschiedenheit ihrer militärischen Organisationen und wirtschaftlichen Interessen auch nur annähernd Rechnung zu tragen und demnach als Grundlage für ein internationales Abrüstungsabkommen zu dienen“ (Auswärtiges Amt 1917).

Neben der korrekten Bewertung militärischer Macht und der Berücksichtigung der Sicherheitsbedürfnisse spielte in diesen Debatten auch die Frage des wahrgenommenen Status der Staaten eine Rolle, welcher die vertraglich erlaubte Größe der Marine oder der Armee jedes Landes mit beeinflusste (vgl. von Bethmann Hollweg 1911: 6002; Keefer 2016: 102, 198-200; Hall 1987: 197-198).

Eine fehlende Einigung über Vergleichsstandards hatte häufig einen negativen Effekt auf den Ausgang der Rüstungskontrollverhandlungen (zum Beispiel während der Haager Konferenzen 1899 oder der Abrüstungskonferenz des Völkerbunds), auf denen die Teilnehmer sich nicht darüber einig wurden, wie und was verglichen werden sollte (vgl. Hall 1987: 116-142; Keefer 2016: 96-127). Fast ausnahmslos beruhen erfolgreiche Rüstungskontrollverhandlungen auf einer Bereitschaft der beteiligten Staaten, nur über einfache Vergleiche zu verhandeln (im Bereich der nuklearen Rüstungskontrolle also etwa ein Vergleich der Anzahl von Abschussrampen und Gefechtsköpfen, aber nicht von Sprengkraft und Zielgenauigkeit) – obwohl sich alle Beteiligten ob der Ungenauigkeit solcher Vergleiche im Klaren waren.²³

Von dieser Tendenz zu einfachen Vergleichen, die sich von den *Washington* und *London Naval Treaties* (1922 und 1930) bis zu den nuklearen Rüstungskontrollverhandlungen im Kalten Krieg beobachten lässt, wurde allerdings auch abgewichen, wie beispielsweise in den in der Mitte der 1970er Jahre initiierten Verhandlungen um konventionelle Rüstungskontrolle in Europa (MBFR). Der dortige Fokus auf Methoden und Vergleichsstandards sorgte dafür, dass in sechzehn Verhandlungsjahren keine Einigung erzielt werden konnte (Bluth 2012). Grundsätzlich verweisen solche Vergleichshindernisse jedoch nicht unbedingt auf ein „technisches“ Problem, sondern stellen ebenfalls willkommene öffentliche Begründungen für den Misserfolg von Rüstungskontrollverhandlungen im Falle mangelnden politischen Willens, ein Abkommen zu erzielen, dar.²⁴

23 Das einfache Ratio-System von Schlachtschiffen, das auf der *Washington Naval Conference* festgelegt wurde, basierte auf „a practical principal of easy application“ (Hall 1987: 198-199; auch Keefer 2016: 8).

24 Frankreich stimmte etwa nur aufgrund der öffentlichen Meinung zu, das Thema der Rüstungsbeschränkung auf der zweiten Haager Konferenz 1907 zu diskutieren (vgl. Keefer 2016: 187). Das Deutsche Reich (und die meisten anderen Großmächte), war auf der ersten Haager Konferenz 1899 von Anfang an nicht gewillt, potenziellen Abrüstungsvereinbarungen zuzustimmen. Der Verweis auf das Scheitern der Haager Konferenzen aufgrund der fehlenden Einigung auf eine Formel kann deshalb als eine Strategie in diesem Sinne interpretiert werden (vgl. Keefer 2016: 96, 120-127; Steller 2011: 248).

Im Anschluss an diese Beobachtung von Auseinandersetzungen bzw. Debatten um das Vergleichen sowie in Folge der vorhergehenden historischen Rekonstruktion können im Kern vier wesentliche Vergleichszwecke identifiziert werden:²⁵

Erstens wurde und wird für Zwecke der militärischen Planung im weiteren Sinne verglichen. Hierbei handelt es sich um den scheinbar *offensichtlichsten* Grund, Streitkräftevergleiche anzustellen, um etwa die Stärke und Ausrüstung der eigenen Streitkräfte an den entsprechenden Fähigkeiten anderer Streitkräfte auszurichten. Es geht hier um strategische und operative Planung, das Feld der vergleichenden Akteure war hier zunächst vergleichsweise überschaubar, wurde im Prozess fortwährender Institutionalisierung und Bürokratisierung jedoch zunehmend komplexer.

Während sich der vorgenannte Vergleichszweck primär auf ein militärisches bzw. sehr militärnahes organisatorisches Feld bezieht, wurde und wird in einem breiteren Kontext des politischen Systems *zweitens* zu Zwecken der Legitimation von bzw. der Argumentation über auf Streitkräfte bezogenen Maßnahmen verglichen. Sei es, um die Beschaffung neuer Waffen zu begründen, sei es, um in Abrüstungsverhandlungen einzutreten, sei es, um eine allgemeine außenpolitische bzw. strategische Ausrichtung zu legitimieren. Es wird hier weitaus mehr für bzw. vor den Augen der Öffentlichkeit verglichen als beim ersten Vergleichszweck (bei dem es zumeist allenfalls um Fachpublika geht; vgl. Förster 2016). Vor allem im Kontext der Legitimierung von Ressourcenallokationen stellt die *Veröffentlichung* von Streitkräftevergleichen einen wichtigen Teil der Vergleichspraktiken dar. Beginnend im späten 19. Jahrhundert, wurden die zunehmend institutionalisierten und häufigeren Streitkräftevergleiche im staatlichen Rahmen (bspw. durch Parlamente, General- und Admiralsstäbe, Nachrichtendienste)²⁶ durch die Berichterstattung über Streitkräftevergleiche durch PublizistInnen und JournalistInnen ergänzt.²⁷

Während sich der vorgenannte Vergleichszweck primär auf nationale Kommunikationsräume und politische Öffentlichkeiten bezieht, dient *drittens* das Vergleichen dem Zweck, einen entsprechenden internationalen Kommunikationsraum herzustellen und hierin Aushandlungsprozesse zu ermöglichen. Insbesondere bei Rüstungskontroll- bzw. Abrüstungsverhandlungen wird hier die gleichzeitig dar- wie herstel-

25 Vgl. Lunn (1989: 50) mit einem in Bezug auf die nachgenannten ersten drei Zwecke ähnlichen Vorschlag für die Kategorisierung von Streitkräftevergleichen spezifisch im Kalten Krieg.

26 Beispielsweise informierte das *Reichsmarineamt* regelmäßig den Reichstag im Rahmen der Debatten über den Marinehaushalt über die relativen Stärken der wichtigsten Marinen. Die entsprechenden Memoranden wurden der Öffentlichkeit dabei nicht nur durch die Veröffentlichung der Reichstagsdebatten zugänglich gemacht, sondern wurden in Form graphischer Darstellungen ebenfalls in Zeitschriften veröffentlicht (vgl. Neuhaus-Wilmersdorf 1906).

27 *Weyers* (seit 1944 als *Weyers Flottentaschenbuch*) ist ein Handbuch über Flotten, das seit 1900 erscheint. Es wurde zuerst vom ehemaligen Marineoffizier Bruno Weyer zusammengestellt und veröffentlicht (Weyer 1904); bei *Jane's Fighting Ships* handelt es sich um ein zuerst 1898 von Fred Jane herausgegebenes Handbuch zu Kriegsschiffen, zunächst zum Zwecke eines Nachschlagewerks für die Freizeitbeschäftigung der Kriegsspieler in Herrenclubs (vgl. Ireland/Grove 1997).

lende Funktion von Streitkräftevergleichen sichtbar. Einerseits ruht jegliche politische Motivation von Verhandlungen auf Vergleichsbeobachtungen, andererseits sind langwierige Auseinandersetzungen über das, was verglichen werden soll und darüber, wie verglichen werden soll, zentraler Kernbereich jeglicher Rüstungskontroll- bzw. Abrüstungsverhandlungen: das Vergleichen stellt hier gleichsam erst den relevanten Bezugsraum her.

In der *abstraktesten* Form dienen die Streitkräftevergleiche *viertens* der relativen Macht- bzw. der Gleichgewichtsbeobachtung. Als Vergleichspraxis stellen sie dabei Macht- und Statusverteilung nicht nur fest, sondern wenigstens in Teilen auch her (vgl. Jaschob 2018: 68). Spätestens seit Ende des 19. Jahrhunderts wurde die Beurteilung der systemweiten Verteilung von Macht und Status innerhalb der Gruppe der Großmächte zu einem wichtigen Zweck von Streitkräftevergleichen. Diese Funktion der Vergleiche, d.h. das Sichtbarmachen einer zuvor eher als „Sense of Place“ (Pouliot 2017: 124-129) wahrgenommenen Hierarchie, konnte mit der signifikanten Zunahme numerischer Vergleiche Ende des 19. Jahrhunderts wiederum Einfluss auf die Machtverteilung und Statushierarchie innerhalb der Gruppe der Großmächte ausüben, da Staaten durch Aufrüsten versuchten, den Platz innerhalb der Gruppenhierarchie zu verändern (vgl. Bönker 2012: 73-74; vgl. auch Renshon 2017: Kapitel 6; Jaschob 2018: Kapitel 4).

Wichtig erscheint nun, dass sich die hier empirisch wie analytisch unterscheidbaren Vergleichszwecke selbstverständlich nicht auf gänzlich voneinander verschiedene Vergleichspraktiken beziehen. Es existieren vielfältige Übergänge und Verknüpfungen, die sich im Laufe der Zeit wandeln: Der zweite Vergleichszweck nahm erst im Zuge von Konstitutionalisierung und Parlamentarisierung, d.h. der Konsolidierung moderner nationalstaatlicher politischer Systeme, einen größeren Raum ein. Bis ins 19. Jahrhundert dienten Vergleichsbeobachtungen vor allem dem Zweck, die relative Stärke der eigenen Streitkräfte zum Zwecke der Kriegsvorbereitung und -durchführung abzuschätzen. Dennoch lässt sich auch hier bereits früh der Vergleichszweck der internen Legitimierung von Ressourcenallokationen feststellen (vgl. Kaunitz 1761; Joseph II 1761; Kluebing 1986: 177-179). Andererseits wirken etwa auch Streitkräftevergleiche im Rahmen von Haushaltsverhandlungen (zweiter Vergleichszweck) unmittelbar auf internationale Beobachtungszusammenhänge (dritter Vergleichszweck) ein (vgl. Seligmann 2016; Stein 2007: 44-45; Cholet 2011: 254-255, 298): Militärhaushalte sind ein Indikator für Aufrüstung und spielten eine wichtige Rolle in der gegenseitigen Beobachtung in Rüstungswettläufen bzw. in Bezug auf Rüstungskontrollbemühungen. Dies hat wiederum Einfluss auf die Macht- und Statusverteilung innerhalb des Systems (vierter Vergleichszweck), da sich diese durch Auf- oder auch Abrüsten verändern kann. Der verwendete Vergleich kann dabei ein und derselbe sein. Ein Vergleich der Zahl und/oder der Tonnage von Schlachtschiffen diente etwa sowohl zur Bestimmung der Verteilung der

relativen Seemacht als auch zur Verdeutlichung der Statushierarchie.²⁸ Schließlich dienten die insbesondere zum ersten Vergleichszweck angesammelten Daten auch schon vereinzelt im 18. und in der ersten Hälfte des 19. Jahrhunderts dazu, Aussagen über die relative Machtposition eines Staates im europäischen Staatensystem (bzw. in Beziehung zu den europäischen Mächten im Fall der USA) zu treffen (vgl. z.B. Horsley 1744; Kolb 1860; Coxe 1806).

5. Streitkräftevergleiche und die Moderne der Weltpolitik

Wie in der historischen Rekonstruktion deutlich wird, verweisen *Praktiken* des Streitkräftevergleichs weder nur auf isolierte Vergleichsbeobachtungen noch beziehen sie sich auf nur eine distinkte Ebene von Weltpolitik als sozialem System. Vielmehr wird gerade in ihrer historischen Rekonstruktion deutlich, dass und wie sich hier verschiedene Konstitutionsebenen in ihrer Entwicklung miteinander verknüpfen, d.h. vor allem die *Meso*- Ebene vergleichender Organisationen und Institutionen, sowie die *Makro*- Ebene eines sich wesentlich über das Programm des fortwährenden Machtvergleichens beschreibenden weltpolitischen Systems.

Streitkräftevergleiche stellen eine wichtige Praktik dar, durch welche Staaten sich und andere Akteure im Verhältnis zu anderen – und andere im Verhältnis zu sich selbst – beobachten. Sie unterfüttern damit gleichsam das *Machtgleichgewicht* als ein generelles Beobachtungsschema mit spezifischen Vergleichsbeobachtungen. Es kann dabei als offene und in weiteren, insbesondere historischen, Studien noch weiter zu klärende Frage angesehen werden, welchen Stellenwert Streitkräftevergleiche im Vergleich zu anderen, auf relative Machtverhältnisse bezogenen Vergleichspraktiken zu unterschiedlichen Zeitpunkten hatten. Dass Streitkräftevergleiche hier eine zentrale Rolle spielen, dürfte allerdings kaum zu bestreiten sein (und die intuitive Eingängigkeit des politischen Realismus mit seiner Betonung der Bedeutung sowohl von Machtgleichgewicht als auch von militärischer Macht spiegelt in dieser Hinsicht ja nicht zuletzt auch die Tatsache wider, dass er historische Praktiken aufgreift und als scheinbar a-historische Gesetzmäßigkeit objektiviert).

Allgemein reflektieren verschiedene Vergleichspraktiken die Verfügbarkeit und die Verarbeitung von immer mehr quantitativen Daten im Zusammenhang rational-bürokratischen Organisierens als wesentlichem Charakteristikum des modernen Territorialstaates (und tragen umgekehrt zu dessen Konsolidierung bei). Nichtsdestotrotz durchlaufen unterschiedliche Vergleichspraktiken auch jeweils spezifische Entwicklungsgeschichten und trotz vieler gemeinsamer Charakteristika, lässt es sich nicht etwa von der Geschichte von Streitkräftevergleichen auf die Geschichte etwa wirtschaftlicher oder demographischer Vergleiche schließen. Ein grundlegen-

28 Dies ist unter anderem ersichtlich an den zentralen Begründungen, die für ein Aufrüsten der Seestreitkräfte in den Rüstungsdebatten ab ca. 1890 genannt wurden: Sicherheit, Machterhalt- oder -zuwachs und Prestige waren darin zentral (etwa von Bülow 1899: 3294; Churchill 1912).

der Unterschied besteht dabei schon alleine darin, dass allen Daten über Streitkräfte der Vergleichsaspekt durch den Bezug auf Macht als relationalem Konzept gleichsam ursprünglich eingeschrieben ist.

Praktiken des Streitkräftevergleichs erstellen und reproduzieren auf der Ebene von Praktiken und Institutionen Differenzen und Merkmale, die gemeinhin der Ebene des *internationalen Systems* zugeschrieben werden, sprich Machtverteilung und Machtgleichgewicht. Ihre Untersuchung bzw. historische Rekonstruktion leistet dabei einen Beitrag zum besseren Verständnis, wie Systemeigenschaften durch Prozesse und Praktiken konstituiert werden, die sich zwischen dem System als relativer sozialer *Makro-* Ebene und dem Handeln einzelner Akteure als relativer *Mikro-* Ebene auf einer *Meso-* Ebene sozialer Konstitution abspielen. In diesem Sinne wird vor allem auch Gleichgewicht als *Beobachtungsschema* im weltpolitischen System vor allem über einfache Zahlenverhältnisse dargestellt. Es wäre jedoch schlichtweg falsch daraus zu folgern, dass diese einfachen Zahlenverhältnisse das Ergebnis von ebenfalls *einfachen* Beobachtungen und Vergleichen – oder gar dem Beobachten oder Vergleichen Einzelner – wären. Wie wir rekonstruiert haben, liegen ihnen langwierige Auseinandersetzungen darüber zugrunde, was zu vergleichen ist und wie dies zu geschehen hat.²⁹ (Militärische) Macht und Machtgleichgewichte harren nicht einfach ihrer Messung wie etwa die Höhe eines Berggipfels. Vielmehr handelt es sich bei Vergleichsbeobachtungen um mit wissenschaftlichen und politischen Auseinandersetzungen durchsetzte, dynamische soziale Praktiken, die in jeweils historisch spezifische soziale Kontexte eingebettet sind. Wir hoffen mit vorliegendem Aufsatz einen Beitrag in zweierlei Hinsicht geleistet zu haben: *Erstens* eine historische Rekonstruktion von Praktiken des Streitkräftevergleichs, welche als Praktiken, die wechselseitige Beobachtungsverhältnisse von Macht widerspiegeln und mitkonstituieren, bislang kaum Aufmerksamkeit gefunden haben (trotz einer langen Beschäftigung mit den Techniken des Vergleichens selbst). *Zweitens* haben wir damit versucht aufzuzeigen, wie das weltpolitische System durch Vergleichspraktiken charakterisiert ist, die spezifisch für dieses System sind, es aber gerade auch an eine soziale Umwelt binden, die durch eine scheinbar nicht enden wollende Zunahme des Vergleichens (durch Zahlen, Rankings, qualitative Vergleiche etc.) charakterisiert ist. Schließlich verstehen wir diesen Aufsatz nicht nur inhaltlich als Beitrag zur historisch-soziologischen IB-Forschung zur Entwicklung des modernen weltpolitischen Systems, sondern im forschungspragmatischen und disziplinpolitischen Sinne auch als Demonstration einer nicht nur rein aus der IB (und von IBlern/-innen) stammenden, sondern einer in Kooperation zwischen einer soziologisch informierten IB und Geschichtswissenschaft gewonnenen, Verwendung historischen Materials.

29 Vgl. Brooks und Wohlforth (2016) als gutes jüngeres Beispiel.

Literatur

- Admiralty* 1919-1921: Capital Ships: Future Policy, TNA ADM 116/3610.
- Admiralty* 1933: Return of the Fleets – History, TNA ADM 116/3745.
- Albert, Mathias* 2016: A Theory of World Politics, Cambridge.
- Albert, Mathias/Müller, Thomas* 2019: Apples and Oranges and Warheads, in: Security Times (Special Edition of The German Times for the 55th Munich Security Conference), 31.
- Auswärtiges Amt* 1917: Denkschrift über die Frage einer vertragsmäßigen Beschränkung der militärischen Rüstungen, PA AA, R 180, Bd. 10448.
- Auswärtiges Amt* 1979: Bewertung des International Institute for Strategic Studies vom September 1979, PA AA, B14, Bd. 125597.
- Awdry, R.D.* 1896: Return to an Order of the Honourable The House of Commons, Navy (Fleets of Great Britain and Foreign Countries), London.
- Bethmann Hollweg, Theobald von* 1911: 159. Sitzung, Donnerstag den 30. März 1911, in: Verhandlungen des Reichstags. XII. Legislaturperiode. II. Session. Band 266, Berlin, 6001-6003.
- Biefang, Andreas* 2009: Die andere Seite der Macht: Reichstag und Öffentlichkeit im „System Bismarck“ 1871–1890, Düsseldorf.
- Bluth, Christoph* 2012: Arms Control as a Part of Strategy. The Warsaw Pact in MBFR Negotiations, in: Cold War History 12: 2, 245-268.
- Bönker, Dirk* 2012: Militarism in a global Age. Naval Ambitions in Germany and the United States before World War I, Ithaca and London.
- Bridge, Roy/Bullen, Roger* 2005: The Great Powers and the European States System 1814-1914, 2. Aufl., Harlow.
- Brooks, Stephen/Wohlforth, William C.* 2016: The Rise and Fall of the Great Powers in the Twenty-First Century, in: International Security 40: 3, 7-53.
- Broome, André/Quirk Joel* (Hrsg.) 2015: The Politics of Numbers (Review of International Studies 41 Special Issue), Cambridge.
- Bülow, Bernhard von* 1899 [1900]: 119. Sitzung, Montag den 11. Dezember 1899, in: Verhandlungen des Reichstags. X. Legislaturperiode. I. Session. Band 4, Berlin, 3292-3295.
- Bundesministerium der Verteidigung* 1979: Zur Military Balance 79/80. Analyse und Bewertung der Aussagen zum nuklearen Kräfteverhältnis in Europa im Vergleich zum Weißbuch der Bundesregierung, PA AA, B14, Bd 125597.
- Buzan, Barry/Lawson, George* 2013: The Global Transformation: The Nineteenth Century and the Making of Modern International Relations, in: International Studies Quarterly 57: 3, 620-634.
- Buzan, Barry/Little, Richard* 2000: International Systems in World History, Oxford.
- Caprivi, Georg Leo von* 1891 [1892]: 128. Sitzung, Freitag den 27. November 1891, in: Verhandlungen des Reichstags. VIII. Legislaturperiode. I. Session. Band 5, Berlin, 3107-3114.
- Cholet, Julia* 2011: Der Etat des Deutschen Reiches in der Bismarckzeit, Berlin.
- Churchill, Winston* 1912: Commons Sitting of Monday 18th March, in: The Parliamentary Debates (Official Report). Fifth Series. Vol. 35, London.
- Collins, John M./Cordesman, Anthony H.* 1978: Imbalance of Power. An Analysis of Shifting U.S.-Soviet Military Strengths and Net Assessment Appraisal, San Rafael.
- Conrad Chef des Generalstabes* 1893–1914: Denkschriften des Chef des Generalstabes, AT-OeStA/KA AhOB Gst Op 740-742.
- Corry, Henry* 1858: Memorandum on the Relative State of the English and French Naval Forces, TNA ADM 116/1.
- Coxe, Tench* 1806: Thoughts on the Subject of Naval Power in the United States of America; and on certain means of encouraging and protecting their commerce and manufactures, Philidelphia.

- Duchhardt, Heinz 1997: Balance of Power und Pentarchie. Handbuch der Geschichte der internationalen Beziehungen, Band 4, Paderborn.
- Dülffer, Jost 1981: Regeln gegen den Krieg: Die Haager Friedenskonferenzen von 1899 und 1908 in der internationalen Politik, Frankfurt am Main.
- Dülffer, Jost 2003: Vom europäischen Mächtesystem zum Weltstaatsystem um die Jahrhundertwende, in: Kröger, Martin u.a. (Hrsg.): Jost Dülffer. Im Zeichen der Gewalt. Frieden und Krieg im 19. Und 20. Jahrhundert, Köln, 49-65.
- Emrich, Eike/Gassmann, Freya/Haut, Jan/Pierdzioch, Christian/Prohl, Robert 2015: Medaillen für die nationale Repräsentanz? Zur Bedeutung von Medaillenerfolgen bei Olympischen Spielen, in: Sport und Gesellschaft 12: 1, 39-68.
- Epple, Angelika/Erhart, Walter (Hrsg.) 2015: Die Welt beobachten. Praktiken des Vergleichens, Frankfurt am Main.
- Förster, Stig (Hrsg.) 2016: Vor dem Sprung ins Dunkle. Die militärische Debatte über den Krieg der Zukunft 1880–1914, Paderborn.
- Freedman, Lawrence 2003: The Evolution of Nuclear Strategy, Basingstoke.
- Friedberg, Aaron 1988: The Weary Titan. Britain and the Experience of relative Decline, 1895-1905, Princeton.
- Friedrich II 1775 [1913]: Betrachtungen über die Feldzugspläne, in: Volz Gustav Bertold (Hrsg.): Die Werke Friedrich des Großen, Band 6, Berlin, 201-221.
- Garthoff, Raymond L. 1990: Estimating Soviet Military Force Levels. Some Light from the Past, in: International Security 14: 4, 93-116.
- Goldblat, Jozef 2002: Arms Control. The New Guide to Negotiations and Agreements, London et. al.
- Goldhamer, Herbert 1979: Reality and Belief in Military Affairs: A First Draft (June 1977). A Report prepared for Director of Net Assessment, Office of the Secretary of Defense. Hrsg. Von Goldhamer, Joan, R-2448-NA, Santa Monica, CA.
- Grawe, Lukas 2017: Deutsche Feindaufklärung vor dem Ersten Weltkrieg. Informationen und Einschätzungen des deutschen Generalstabs zu den Armeen Frankreichs und Russlands 1904–1914, Paderborn.
- Guzzini, Stefano 2013: Power, Realism and Constructivism, London.
- Hall, Christopher 1987: Britain, America and Arms Control, 1921-1937, Basingstoke.
- Hayard, Léon. (Hrsg.) 1892: Tableau Comparatif des Forces Militaires de la France et de la Russie en Face de Celles de la Triple-Alliance, Paris.
- Heintz, Bettina 2016: „Wir leben im Zeitalter der Vergleichung“. Perspektiven einer Soziologie des Vergleichs, in: Zeitschrift für Soziologie 45: 5, 305-323.
- Heintz, Bettina/Werron Tobias 2011: Wie ist Globalisierung möglich? Zur Entstehung globaler Vergleichshorizonte am Beispiel von Wissenschaft und Sport, in: Kölner Zeitschrift für Soziologie und Sozialpsychologie 63: 3, 359-94.
- Herrera, Geoffrey L./Mahnken, Thomas Gilbert 2003: Military Diffusion in Nineteenth-Century Europe. The Napoleonic and Prussian Military System, in: Goldman Emily O./Eliason, Leslie C. (Hrsg.): The Diffusion of Military Technology and Ideas, Stanford, Calif., 205-242.
- Herzog, Richard J./Wildgen, John K. 1986: Tactics in Military Propaganda Documents: A Content Analysis of Illustrations, in: Defense Analysis 2: 1, 35-46.
- Hobson, Rolf 2004: Maritimer Imperialismus. Seemachtideologie, seestrategisches Denken und der Tirpitzplan 1875 bis 1914, München.
- Horowitz, Michael C. 2010: The Diffusion of Military Power. Causes and Consequences for International Politics, Princeton.
- Horsley, William 1744: A Treatise on Maritime Affairs: Or a Comparison between the Commerce and Naval Power of England and France with a View to some Paradoxes Advanced by M. Deslandes: and More Particularly, to Several Popular Prejudices Entertained of Late Concerning the Formidable Power of France, London.
- International Institute for Strategic Studies 1976: The Military Balance 1976-1977, Oxford.

- International Institute for Strategic Studies* 1979: *The Military Balance 1979-1980*, Oxford.
- International Institute for Strategic Studies* 1981: *The Military Balance 1981-1982*, Oxford.
- International Institute for Strategic Studies* 1983: *The Military Balance 1983-1984*, Oxford.
- Ireland, Bernard/Grove, Eric* 1997: *Jane's War at Sea, 1897–1997: 100 Years of Jane's Fighting Ships*, London.
- Jackson, Peter J./Siegel, Jennifer* (Hrsg.) 2005: *Intelligence and Statecraft. The Use and Limits of Intelligence in International Society*, Westport, Conn.
- Janjoch, Lena* 2018: *Status im internationalen System. Das Deutsche Reich und sein Statusstreben in Europa 1890-1914*, Baden-Baden.
- Joseph II* 1761: *Memoria*, AT-OeStA/HHStA StK Vorträge 88, IV-VI, fl. 1-9.
- Kaunitz, Anton Wenzel Graf von* 1761: *Ohnmaaßgebliche Gedancken und Anmerkungen über den gegenwärtigen Stand der Kriegs-Operationen*, AT-OeStA/HHStA StK Vorträge 88, VII-IX, fl. 179-194.
- Keefe, Scott Andrew* 2016: *The Law of Nations and Britain's Quest for Naval Security: International Law and Arms Control, 1898–1914*, Cham.
- Kloosterhuis, Jürgen/Bahl, Peter* (Hrsg.) 2015: *Militär und Gesellschaft in Preußen. Quellen zur Militärsozialisation 1713–1806*, Archivalien in Berlin, Dessau und Leipzig. Geheimes Staatsarchiv, Berlin.
- Kluebing, Harm* 1986: *Die Lehre von der Macht der Staaten. Das aussenpolitische Machtproblem in der „politischen Wissenschaft“ und in der praktischen Politik im 18. Jh.*, Berlin.
- Kolb, Georg Friedrich* 1860: *Handbuch der vergleichenden Statistik der Völkerzustands und Staatenkunde*, 2. Auf., Leipzig.
- Krakau, Anton/Diehl, Ole* 1989: *„Annähernde Parität“ der Streitkräfte in Europa? Eine kritische Analyse des östlichen Streitkräftevergleichs*, Köln.
- Krauseneck, Johann Wilhelm von* 1831: *Krauseneck an Bernstorff*, GStA PK, III. HA MdA I, Nr. 10493.
- League of Nations* 1922: *Statistical Enquiry on Armaments. First Series of Data*, A.21.1922.IX.
- League of Nations* 1923: *Statistical Enquiry into National Peace-time Armaments, Part I*, A.20.1923.IX.
- League of Nations* 1924: *Armaments Year-Book. General and Statistical Information*. 2. Auflage, Geneva.
- League of Nations* 1939/1940: *Armaments Year-Book. General and Statistical Information*, Geneva.
- Lunn, Simon* 1989: *The East-West Military Balance: Assessing Chance*, in: Kaiser, Karl (Hrsg.): *The Changing Strategic Landscape*, London, 49-71.
- Lutz, Dieter S.* 1986: *Towards a Methodology of Military Force Comparison. An Introduction to Some Aspects of an Instrument of Security Policy with Special Emphasis on Nuclear Forces at the Beginning of the Eighties*, Baden-Baden.
- Macnamara, Patrick* 1919: *Limitation of Naval Armaments*, in: TNA ADM 116/1863.
- Murray, Oswyn*. 1921: *Return to an Order of the Honourable The House of Commons, Navy (Fleets of Great Britain and Foreign Countries)*, London.
- NATO* 1982: *NATO and the Warsaw Pact: Force Comparisons*, Brussels.
- NATO* 1984: *NATO and the Warsaw Pact: Force Comparisons*, Brussels.
- NATO* 1988: *Conventional Forces in Europe: The Facts*, Brussels.
- NATO Information Service* 1972: *Allied Defence in the Seventies*, Brussels.
- Nauticus* (Hrsg.) 1908: *Jahrbuch für Deutschlands Seeinteressen*, Zehnter Jahrgang, Berlin.
- Navy Board and Admiralty* 1919: *Navy Estimates 1919–1920. Admiralty Memorandum for the War Cabinet*, TNA CAB/24/82/100.
- Neuhaus-Wilmersdorf, Erich* 1906: *Illustrierte Darstellung der See-Interessen des Deutschen Reichs*, BArch RM 2/112.
- Office of Naval Intelligence* 1909: *Information Concerning some of the Principal Navies of the World: A Series of Tables Compiled to Answer Popular Inquiry*, Washington.

- Ostrich, John T./Green William C. 1981: Methodological Problems Associated with the IISS Military Balance, in: Comparative Strategy 3: 2, 151-171.
- Parkinson, Roger 2008: Late Victorian Navy: The Pre-dreadnought Era and the Origins of the First World War, Suffolk.
- Persson, Gudrun 2010: Learning from Foreign Wars. Russian Military Thinking, 1859–1873, Solihull.
- Pouliot, Vincent 2017: Against Authority. The Heavy Weight of International Hierarchy, in: Zarakol, Ayşe (Hrsg.): Hierarchies in World Politics, Cambridge, 113-134.
- Prados, John 1982: The Soviet Estimate. U.S. Intelligence Analysis [and] Russian Military Strength, New York.
- Reichsmarineamt 1897: Die Seeinteressen des Deutschen Reichs zusammengestellt auf Veranlassung des Reichsmarineamt, in: Stenographische Berichte über die Verhandlungen des Reichstags. IX. Legislaturperiode. V. Session. Erster Anlagenband, Berlin, 11-90.
- Renshon, Jonathan 2017: Fighting for Status. Hierarchy and Conflict in World Politics, Princeton, New Jersey.
- Roberts, Peter/Kaushal, Sidhardt 2018: Strategic Net Assessment: Opportunities and Pitfalls, in: The RUSI Journal 163: 6, 66-76.
- Rohn, Laurinda L. 1990: Conventional Forces in Europe. A New Approach to the Balance, Stability, and Arms Control, Santa Monica.
- Roon, Albrecht von 1862: Denkschrift zum Flottenplan, GStPK, III. HA MdA I, Nr. 10671, 3-20.
- Rose, Andreas 2013: Der politische Raum Londons und die öffentlichen Beziehungen zwischen England und Deutschland vor 1914, in: Bösch, Frank/Hoeres, Peter (Hrsg.): Außenpolitik im Medienzeitalter. Vom späten 19. Jahrhundert bis zur Gegenwart, Göttingen, 95-121.
- Seligmann, Matthew S. 2006: Spies in Uniform. British Military and Naval Intelligence on the Eve of the First World War, Oxford.
- Seligmann, Matthew S. 2016: The Anglo-German Naval Race, 1898–1914, in: Mahnken, Thomas Gilbert/Maiolo, Joseph A./Stevenson, David (Hrsg.): Arms Races in International Politics. From the Nineteenth to the Twenty-First Century, Oxford, 21-40.
- Sittart Neale, Henry J. van 1898: Return to an Order of the Honourable The House of Commons, Navy (Fleets of Great Britain and Foreign Countries), London.
- Soviet Union 1982-1987: Whence the Threat to Peace (4 Auflagen), Moskau.
- Speich, Daniel 2013: Die Erfindung des Bruttosozialprodukts: Globale Ungleichheit in der Wissensgeschichte der Ökonomie, Göttingen.
- Sprengel, Matthias Christian 1793: Grundris der Staatenkunde der vornehmsten europäischen Reiche, Teil 1 und 2, Halle.
- Stein, Oliver 2007: Die Deutsche Heeresrüstungspolitik 1890–1914: Das Militär und das Primat der Politik, Paderborn.
- Steller, Verena 2011: Diplomatie von Angesicht zu Angesicht. Diplomatische Handlungsformen in den deutsch-französischen Beziehungen 1870-1919, Paderborn.
- Stollberg-Rilinger, Barbara 2015: The Emperor's Old Clothes: Constitutional History and the Symbolic Language of the Holy Roman Empire, London.
- Storz, Dieter 2003: Modernes Infanteriegewehr und taktische Reform in Deutschland in der Mitte des 19. Jahrhunderts, in: Epkenhans, Michael/Gross, Gerhard Paul (Hrsg.): Das Militär und der Aufbruch in die Moderne, 1860 bis 1890. Armeen, Marinen und der Wandel von Politik, Gesellschaft und Wirtschaft in Europa, den USA sowie Japan, Oldenbourg, 209-230.
- Tachikawa, Kyoichi 2015: Japanese Pre-War Military Attaché System, in: NIDS Journal of Defense and Security 16: 147-185.
- Tooze, J. Adam 2014: The Deluge. The Great War, America and the Remaking of the Global Order, 1916–1931, New York.
- United Nations 1989: The United Nations Disarmament Yearbook, Band 13, New York.

- US War Department* 1911: Strength and Organization of the Armies of France, Germany, Austria, Russia, England, Italy, Mexico and Japan, Washington.
- Vagts, Alfred* 1967: The Military Attaché, Princeton NJ.
- Sass, Hartmut von* 2018: Comparisons: A Typology, in: https://www.academia.edu/37901487/von_Sass_Comparisons_A_Typology; 25.03.2020.
- Ward, Steven* 2017: Status and the Challenge of Rising Powers, Cambridge.
- Warner, Michael* 2014: The Rise and Fall of Intelligence: An International Security History, Washington.
- Webster, Andrew* 2005: The Transnational Dream: Politicians, Diplomats and Soldiers in the League of Nations Pursuit of International Disarmament, 1920–1938, in: Contemporary European History 14: 4, 493-518.
- Werron, Tobias* 2012: Wie ist globale Konkurrenz möglich? Zu rsozialen Konstruktion globaler Konkurrenz am Beispiel des Human Development Index, in: Soziale Systeme 18: 1+2, 168-203.
- Wettstein, Adrian E.* 2016: Die Grenzen militärischer Prognostik. Die Diskussion um den Krieg der Zukunft in den französischen Militärzeitschriften, in: Förster, Stig (Hrsg.): Vor dem Sprung ins Dunkle. Die militärische Debatte über den Krieg der Zukunft 1880–1914 Paderborn, 133-243.
- Weyer, Bruno* 1904: Taschenbuch der Kriegsflotten, V. Jahrgang, 3. Aufl., München.
- Widenmann, Wilhelm* 1909 [1927]: Der Marineattaché in London Korvettenkapitän Widenmann an den Staatssekretär des Reichsmarinemats von Tirpitz, in: Lepsius, Johannes/Mendelssohn Bartholdy, Albrecht/Thimme, Friedrich (Hrsg.): Große Politik der Europäischen Kabinette 1871–1914. Band 28, Berlin, Nr. 10291.
- Willett, Marcus* 2019: Assessing Cyber Power, in: Survival 61: 1, 85-90.