

Peter Thiery, Jenniver Sehring und Wolfgang Muno

Wie misst man Recht? – Möglichkeiten und Grenzen der Messung von Rechtsstaatlichkeit

Zusammenfassung

Der vorliegende Beitrag vergleicht und bewertet die Messung von Rechtsstaatlichkeit anhand quantitativer Indizes. Dazu werden drei prominente Indizes (Freedom in the World, Worldwide Governance Indicators, Bertelsmann Transformation Index) hinsichtlich ihrer Konzeption, Operationalisierung und Ergebnisse verglichen. Die Analyse zeigt, dass alle drei Indizes methodische Schwächen aufweisen. Zudem erfassen sie zwar den Rechtsstaat bzw. Defizite des Rechtsstaates, nicht aber die gesamte Rechtswirklichkeit.

How to measure law? – Potentials and constraints of measuring the rule of law

Summary

This paper compares and assesses three approaches to measure rule of law using quantitative indices. It compares three indexes (Freedom in the World, the Worldwide Governance Indicators, and the Bertelsmann Transformation Index) concerning their conception, operationalisation, and results. It can be shown that all three indices suffer from methodical weaknesses. In addition, while they capture rule of law and deficiencies within the rule of law, they fail to capture the total legal reality.

1. Einleitung

Dieser Beitrag widmet sich mit dem Rechtsstaat einem besonderen Aspekt der Rechtswirklichkeit. Dabei geht es darum, eine besondere Perspektive zur Erfassung dieses Gegenstandes zu skizzieren und ihre Vor- und Nachteile zu analysieren, nämlich die quantitative Messung von Rechtsstaatlichkeit. Wie zu zeigen sein wird, besteht trotz aller Vorbehalte gegenüber den verschiedenen Methoden der Messung von Rechtsstaatlichkeit durchaus Grund zu der Annahme, dass ihr einheitlicher Befund richtig ist: Dass der Rechtsstaat – in einem globalen Maßstab gesehen – nicht so wirkt wie dies formal vorgesehen ist.

In der Demokratieforschung, wie in der Politikwissenschaft allgemein, zeigte sich in den vergangenen Jahren der Trend, mit großen Fallzahlen zu arbeiten, was insbesondere durch die politischen Entwicklungen im Zuge der so genannten „Dritten Welle“

der Demokratisierung seit den 1970er Jahren (Huntington 1991; Zimmerling 2003) möglich wurde. Rechtsstaatsmessungen versprechen dabei, über viele Fälle hinweg eine vergleichbare Erfassung des Gegenstandes zu ermöglichen und ihn so – als abhängige wie als unabhängige Variable – für die statistische Analyse von Ursache-Wirkungs-Zusammenhängen zugänglich zu machen. Eine kritische Analyse von Möglichkeiten und Grenzen der Messung von Rechtsstaatlichkeit erscheint daher sinnvoll. Aus der einfachen Frage: *Wie misst man Recht?* ergibt sich unser Leitfaden, der in folgenden Fragen besteht:

1. Was wird gemessen: Welches Konzept von Rechtsstaatlichkeit liegt der Messung zugrunde, und welche Indikatoren werden verwendet?
2. Wie wird gemessen: Mit welchen Methoden wird operiert und wie werden sie mit welchem Ziel angewendet? Wie sauber und zuverlässig wird gemessen (Validität und Reliabilität)?
3. Welche Ergebnisse werden produziert und wie ist deren Qualität einzuschätzen? Werden mit verschiedenen Methoden etwa ähnliche Ergebnisse erzielt oder gibt es signifikante Diskrepanzen?

Während es gegenwärtig noch keinen expliziten Index zur Messung von Rechtsstaat bzw. *Rule of law* gibt, wird Rechtsstaatlichkeit als Komponente in verschiedenen Indizes zur Messung von Demokratie oder *Good Governance* erfasst. Von diesen Indizes haben wir jene drei für den Vergleich ausgewählt, die nicht nur als die profiliertesten gelten, sondern vor allem Rechtsstaatlichkeit als substantielle Kategorie ihrer Untersuchungen erfassen: Freedom in the World von Freedom House, die Worldwide Governance Indicators der Weltbank sowie den Transformationsindex der Bertelsmann-Stiftung.¹ Diese drei Messungen werden anhand von Analysekatégorien, die sich mittlerweile in der Beurteilung der Demokratiemessung etabliert haben (Munck/Verkuilen 2002; Müller/Pickel 2007), vergleichend analysiert. Neben der Einordnung der Ansätze geht es um die Frage der Qualität der Messung (Konzeptualisierung, Operationalisierung, Messung, Aggregation) sowie schließlich um die Aussagekraft solcher Messungen von Rechtsstaatlichkeit.

Dabei wollen wir aufzeigen, dass alle drei Indizes zwar zentrale Aspekte des Rechtsstaates erfassen und mit einiger Zuverlässigkeit auch messen, aber aufgrund der je unterschiedlichen Rationalität ihrer Untersuchungen kein hinreichend differenziertes Bild der Rechtsstaatlichkeit ermöglichen und so nur mit Vorbehalt für weiterführende Analysen des Rechtsstaats nutzbar sind. Abschließend werden weiterführende Fragen formuliert, insbesondere dahingehend, ob die etablierten Messungen nicht generell unter einem eingeschränkten Blickwinkel leiden.

¹ Daneben wird Rechtsstaatlichkeit partiell auch in anderen Indizes erfasst, wie dem *Failed States Index* des *Fund for Peace*. Die *American Bar Association* konzipiert in ihrem *World Justice Project* einen eigenen *Rule of Law*-Index, der sich allerdings noch in der Anfangsphase befindet. Wir haben uns hier auf diejenigen beschränkt, die in der Literatur am häufigsten referiert und in der Entwicklungszusammenarbeit auch praktisch genutzt werden.

2. Der „Kontext“ der Rechtsstaatsmessung

Indizes verdichten relevante, aber im Einzelfall komplexe Informationen anhand möglichst präzise konstruierter Konzepte, Indikatoren und Verfahren zu einer Art Landkarte, die sowohl der Wissenschaft wie auch der Politik als Grundlage für ihre weiteren Analysen bzw. Entscheidungen dient. Keiner der hier analysierten Indizes beansprucht, ausschließlich einen Rechtsstaatsindex zu produzieren. Vielmehr sind die Rechtsstaatsdaten eingebunden in übergreifende Messziele, lediglich die WGI beinhalten einen eigenen Sub-Index *Rule of Law*. Dieser jeweilige Kontext beeinflusst auch die Konstruktion und Auswahl der Kriterien und Indikatoren sowie partiell die Methodik der Rechtsstaatsmessung.

Freedom in the World (FIW) – seit 1972 von der US-amerikanischen Nichtregierungsorganisation (NGO) Freedom House erstellt – misst den Stand der politischen und bürgerlichen Freiheiten in den Staaten der Erde. Damit stellt er keine Demokratie-messung im eigentlichen Sinne dar, weist aber viele Überschneidungen mit Demokratieindizes auf und wird deswegen (auch) als Demokratie-Index verstanden und verwendet (Lauth 2004: 268). Er ist einerseits einer der meistzitierten Demokratie-Indizes, der seine Bedeutung u.a. durch den großen Umfang der erfassten Länder, den langen Zeitraum der Messung, die jährliche Erstellung und durch die Abrufbarkeit der Daten im Internet Aktualität erlangte. Andererseits ist er von wissenschaftlicher Seite wegen der Art der Operationalisierung und der mangelnden Transparenz methodisch umstritten.

Für jedes der mittlerweile 193 Länder wird ein Bericht erstellt, wobei sowohl die formalen Regeln als auch die tatsächlichen Praktiken analysiert, letztere aber stärker gewichtet werden. Zur Messung von Freiheit werden zwei Kategorien gebildet: *political rights* und *civil liberties*, die getrennt bewertet werden und zusammen den Gesamtindex bilden. Die Bewertung erfolgt anhand eines numerischen Ratings von 1-7, wobei 1 der beste zu erreichende Wert ist. Die Wertungen werden zunächst für beide Kategorien getrennt vergeben, anschließend wird über das arithmetische Mittel der Gesamtwert gebildet. Dieser Gesamtindex ist die Basis für die typologische Zuordnung des jeweiligen Landes in „frei“, „teilweise frei“ und „nicht frei“. Die beiden Kategorien *political rights* und *civil liberties* setzen sich jeweils aus mehreren Subkategorien zusammen, die wiederum mehrere Leitfragen umfassen. Aktuell umfasst die Liste zehn Leitfragen bei *political rights* und 15 bei *civil liberties*. Rechtsstaatlichkeit (*Rule of law*) ist eine Subkategorie von *civil liberties*.

Die *Worldwide Governance Indicators* (WGI) – Ende der neunziger Jahre am World Bank Institute entwickelt – sind ein Meta-Indikator zur Messung von *Governance* bzw. *Good Governance*, bei dem Rechtsstaatlichkeit einen der Teilindizes bildet. Ausgangspunkt war die in der entwicklungspolitischen Diskussion rapide zunehmende Diskussion um „Good Governance“, d.h. um politische, institutionelle und rechtliche Rahmenbedingungen für die Entwicklung seit Ende der achtziger Jahre. Der Begriff „Good Governance“ war (und ist bis heute) nicht klar, da es keine Einigung

gibt über den Inhalt und die Definition (Landman 2003: 1). Die Autoren definieren *Governance* als „the traditions and institutions by which authority in a country is exercised“ (Kaufman/Kraay/Zoido-Lobaton 1999: 1). Drei Faktoren sind damit gemeint, die in jeweils zwei Clustern gemessen werden: (1) Der Prozess, durch den Regierungen gewählt, überwacht und ersetzt werden, wird erfasst durch *Voice and Accountability* sowie durch *Political Instability and Violence*. (2) Die Fähigkeit von Staaten, ihre Politiken durchzusetzen, werden durch *Government Effectiveness* und *Regulatory Burden* erfasst. (3) Der Respekt, den Bürger und Staaten für die Institutionen zeigen, die soziale Interaktionen regulieren, werden erfasst durch *Rule of Law* und *Control of Corruption* (Kaufmann/Kraay/Mastruzzi 2008: 7ff.).

Die WGI wurden zwischen 1999 und 2004 zweijährlich, seit 2004 jährlich weiterentwickelt. Als Meta-Index werden für die WGI keine eigenen Primärdaten erhoben, sondern Fremdquellen benutzt, die ihrerseits auf subjektiven Einschätzungen und nicht auf „harten“, „objektiven“ Daten beruhen. In der Version von 2008 wurden für die WGI 2007 340 Einzelindikatoren aus 35 verschiedenen Quellen herangezogen, die von 32 verschiedenen Organisationen produziert wurden und insgesamt 212 Länder erfassen (Kaufman/Kraay/Mastruzzi 2008: 8). Dabei handelt es sich einerseits um internationale Umfragen (Surveys), etwa von Transparency International, Gallup oder dem Latinobarómetro, andererseits um Experteneinschätzungen (jeweils neun von NGOs wie Freedom House oder dem Bertelsmann Transformation Index, neun von öffentlichen Stellen wie der Weltbank selbst oder dem US State Department, acht von kommerziellen Wirtschaftsinformationsdiensten wie dem Economist Intelligence Unit, Kaufman/Kraay/Mastruzzi 2008: 29). Dabei erfasst nicht jede Quelle 212 Länder, die Spannweite reicht von zwölf bis 192 Staaten.

Der Bertelsmann Transformation Index ist der jüngste der drei Indizes, er erscheint seit 2004 alle zwei Jahre. Als Transformationsindex ist er relativ komplex angelegt, da er nicht nur den Status quo der Transformationsziele Demokratie und Marktwirtschaft erfassen will („Status-Index“), sondern auch die Steuerungsleistungen der politischen Eliten auf dem Weg dahin („Management-Index“) (Bertelsmann Stiftung 2008). Letzteres ähnelt in der Intention den *Governance Indicators* der Weltbank, doch versucht der BTI weniger die strukturellen Gegebenheiten von *Good Governance* zu erfassen, sondern zielt dezidiert auf die handelnden politischen Akteure. Der BTI untersucht nur Transformations- und Entwicklungsländer; für die OECD-Länder wurde ein eigener „Reform-Index“ entwickelt (Bertelsmann Stiftung 2009).

Da rechtsstaatliche Demokratie und sozial verantwortliche Marktwirtschaft die normativen Bezugspunkte des Rankings darstellen, wird die Untersuchung nicht als relativer Vergleich von Staaten untereinander konzipiert, sondern als Bewertung (Rating) des Entwicklungsstandes und der Managementleistung im Hinblick auf die Verwirklichung dieser normativen Ziele. Diese drei Dimensionen sind in jeweils fünf bis sieben Einzelkriterien aufgeschlüsselt, jedes Kriterium wiederum ist in einzelne Fragen untergliedert, die als Indikatoren der Messung fungieren. Die insgesamt 17

Kriterien und 47 Indikatoren bilden die Grundlage der Teilindizes. Anhand dieser Indikatoren werden zum einen ausführliche Länderberichte erstellt, zum anderen dienen sie auch der numerischen Bewertung. Im Unterschied zu weniger voraussetzungsvollen Demokratiebegriffen beinhaltet hier rechtsstaatliche Demokratie die Gesetzesbindung staatlichen Handelns sowie die Teilung und wechselseitige Kontrolle von legislativer, exekutiver und richterlicher Gewalt.

3. Die Konzeptualisierung von Rechtsstaatlichkeit

Die Qualität von Messungen ist zunächst von der Konzeptvalidität abhängig, zu der u.a. die geeignete Mischung aus Sparsamkeit und Relevanz sowie logisch stringente Zuordnungen von Merkmalen und Dimensionen gehören (Müller/Pickel 2007: 514ff.). Im Falle von FIW ist zunächst hervorzuheben, dass Freedom House keinen expliziten Begriff der Rechtsstaatlichkeit entwickelt, in der Kriterienentwicklung aber einem *common sense*-Verständnis folgt, das wesentliche Aspekte des Rechtsstaates umfasst. Die Konzeptualisierung erfolgt in vier Dimensionen, für die Leitfragen formuliert werden. Diese Leitfragen wurden im Laufe der Jahre leicht geändert und umstrukturiert; im letzten Bericht von 2008 waren es folgende:

- Gibt es eine unabhängige Justiz?
- Ist Rechtsstaatlichkeit im Zivil- und Strafrecht gegeben? Steht die Polizei unter direkter ziviler Kontrolle?
- Gibt es Schutz vor politischem Terror, ungerechtfertigter Verhaftung, Exil, oder Folter, ob von systemunterstützenden oder gegnerischen Gruppen? Gibt es Freiheit von Krieg und Aufständen?
- Garantieren Gesetze, Politik und Praktiken gleiche Behandlung verschiedener Teile der Bevölkerung?

Diese vier Leitfragen werden jeweils in 5 bis 7 Einzelfragen aufgefächert (Freedom House 2008a). Über die Herleitung dieser Aspekte erfährt man jedoch nichts Genaueres. Laut Eigenangaben des Begründers von Freedom in the World wurden die Subkategorien nicht anhand von theoretischen Überlegungen ausgewählt, sondern stellen vielmehr eine „checklist“ aufgrund empirischer Erfahrungen dar, weil die für die Bewertung notwendigen Informationen in den meisten Ländern verfügbar sind. (Gastil 1990: 30).

Wie bereits erwähnt ist es das Anliegen von FIW, den Grad der Freiheit zu erfassen. Freiheit wird definiert als „the opportunity to act spontaneously in a variety of fields outside the control of the government and other centers of potential domination“ (Freedom House 2008a). Das Verständnis von Freiheit allgemein wird aus den Kriterien der Allgemeinen Erklärung der Menschenrechte abgeleitet. Bei einigen Prüffragen zur Rechtsstaatlichkeit lässt sich ein direkter Bezug zu dieser normativen Basis erkennen, so zu Art. 2 (keine Unterscheidung nach Geschlecht, Hautfarbe, Re-

ligion etc.), Art. 5 (Verbot der Folter und grausamer Behandlung), Art. 8 (Anspruch auf wirksamen Rechtsbehelf), Art. 9 (Verbot willkürlicher Festnahme oder Haft), Art. 10 (Anspruch auf gerechtes Verfahren vor unparteiischem Gericht), Art. 11 (Recht auf Unschuldsvermutung) und Art. 14 (Recht auf Asyl).

Am schwächsten ausgeprägt ist die Konzeptvalidität bei den Worldwide Governance Indicators der Weltbank. Rechtsstaatlichkeit wird von den Autoren weder theoretisch noch historisch umrissen, was insbesondere die Herleitung und Zuordnung der Merkmale und Subdimensionen wenig nachvollziehbar macht. Die Autoren der WGI verstehen Rechtsstaatlichkeit vor allem als das Vertrauen in Regeln und die Gesetzestreue, wobei es speziell um die Qualität von Vertragssicherheit, Eigentumsrechte, die Polizei und die Gerichte sowie um Verbrechen und Gewalt geht. Auffällig ist, dass Elemente, die gemeinhin als Teil von Rechtsstaatlichkeit angesehen werden, etwa der Schutz von Menschen- und Bürgerrechten oder Korruption, nicht als Bestandteil von Rechtsstaatlichkeit, sondern als eigene Cluster bzw. Teilindikatoren angesehen werden. Insgesamt überwiegen bei den WGI die Aspekte des formalen Rechtsstaates gegenüber denen des materiellen Rechtsstaates.

Der BTI nimmt die Messung von Rechtsstaatlichkeit im Rahmen seiner Demokratie-messung vor, wo sie ein Kriterium neben Staatlichkeit, politischer Partizipation sowie zwei Kriterien der demokratischen Konsolidierung bildet. Hierbei folgt der BTI einem explizierten Demokratiekonzept (Merkel/Puhle et al. 2003: 39ff.), das ausführlich die Notwendigkeit der rechtsstaatlichen Einhegung der (elektoralen) Demokratie herleitet und dabei auf Autoren wie Habermas (1992; 1994) oder Dworkin (1998) zurückgreift. Zentral für Rechtsstaatlichkeit sind dabei die „horizontale Verantwortlichkeit“ sowie die Gewährleistung bürgerlicher Freiheitsrechte. Sie werden in insgesamt vier Indikatoren aufgeschlüsselt, die jeweils mit entsprechenden Antwortvorgaben zur Bewertung versehen werden (Bertelsmann Stiftung 2008: 80f.):

- Inwieweit funktioniert die Teilung und wechselseitige Kontrolle der staatlichen Gewalten?
- Inwieweit existiert eine unabhängige Justiz?
- Inwieweit wird Amtsmissbrauch von Mandatsträgern rechtlich oder politisch geahndet?
- Inwieweit sind bürgerliche Freiheitsrechte gewahrt und geschützt und inwieweit können sie von den Bürgern eingeklagt werden?

Damit ist das Rechtsstaatskonzept hinreichend und sparsam konzipiert: Es umfasst sowohl den formalen wie den materiellen Rechtsstaat und schließt die Kontrolldimension der Gewaltenteilung mit ein. Wie oben angedeutet, fehlt dem BTI durch die komplexe Anlage als Transformationsindex allerdings die Tiefenschärfe, da die vier Indikatoren nicht weiter operationalisiert werden.

Den drei Indizes ist gemeinsam, dass sie im Rahmen ihrer unterschiedlichen Zielsetzungen Rechtsstaatlichkeit als Unterkriterium der Gesamtindizes messen und hierfür eigene Indikatoren konstruieren. Die Frage nach der Adäquatheit dieser Rechtsstaatsmessungen ist zunächst mit der Validität der von ihnen benutzten Konzepte verknüpft. Hier zeigt sich, dass die Herleitungen und Begründungen für die Konzepte wie für die benutzten Indikatoren unterschiedlich und teils recht mager ausfallen. Dies gilt insbesondere für das Rechtsstaatskonzept der Weltbank, während das BTI-Konzept eine stärkere demokratietheoretische Fundierung besitzt. Bei Freedom House wiederum stellen die Menschenrechte als normativer Bezugspunkt die Ausgangsbasis dar. Ausgehend von diesen unterschiedlichen Rechtsstaats- bzw. Demokratiekonzeptionen werden auch verschiedene Indikatoren herangezogen, die in Tabelle 1 nochmals vergleichend gegenübergestellt sind.

Tabelle 1: Indikatoren der Rechtsstaatsmessung im Vergleich

Messung	Indikatoren
FIW	Unabhängige Justiz
	Rechtsstaatlichkeit in Zivil- und Strafverfahren
	Schutz vor politischem Terror, ungerechtfertigter Verhaftung, Exil, Folter sowie Abwesenheit von Krieg und Aufständen
	Gleichbehandlung vor dem Gesetz, in Politik und Verfahren
WGI	Vertrauen in die gesellschaftlichen Regeln
	Befolgung von Regeln
	Hohes Niveau der Vertragsdurchsetzung
	Hohe Qualität der Polizei
	Hohe Qualität der Gerichte
	Geringe Wahrscheinlichkeit von Verbrechen
	Geringe Wahrscheinlichkeit von Gewalt
BTI	Funktionierende Gewaltenteilung
	Unabhängige Justiz
	Ahndung von Amtsmissbrauch
	Bürgerliche Freiheitsrechte

4. Methodik

Indexkonstruktionen weisen immer eine Reihe methodischer Probleme auf (Müller/Pickel 2007: 525ff.), die wir im Folgenden im Hinblick auf die Verfahren der Messung und Aggregation untersuchen. Bei FIW wird zur Erstellung der Rohdaten jede der vier Leitfragen für die Subkategorie *Rule of law* mit jeweils 0 (nicht vorhanden) bis 4 (voll vorhanden) Punkten bewertet. Insgesamt können also 16 Punkte für Rechtsstaatlichkeit erreicht werden, die zu den maximal 60 erreichbaren Punkten der Kategorie *civil liberties* gehören. Die letztlich quantitative Bewertung erfolgt nach einem mehrstufigen qualitativen Berichts- und Begutachtungsverfahren. Während

der letzten Jahrzehnte wurde die Methodik modifiziert, z.B. neue Unterfragen für einzelne Kategorien eingeführt. Deswegen wird kritisiert, dass dadurch die Ratings weniger konsistent seien (Lauth 2004: 269, Munck/Verkuilen 2002: 21). Heutige Datengrundlage für die Analyse sind in- und ausländische Medienberichte, wissenschaftliche Analysen, Berichte von NGOs, Think Tanks, Informationen durch individuelle professionelle Kontakte sowie Forschungsbesuche vor Ort. Im Internet ist eine Auswahl der Quellen veröffentlicht. Das Rating wird intern sowie auf Regional Konferenzen mit Regionalexperten überprüft, anschließend nochmals interregional, um die Vergleichbarkeit sicherzustellen.

Die Kodierung erfolgte bis zum Bericht von 1988/89 allein durch den Leiter der Surveys. Nach einer ersten Einordnung der Länder wurde in den Folgejahren geprüft, ob sie weiterhin in die Kategorie passten oder nicht (Gastil 1990: 26-28). Erst mit dem kontinuierlichen Ausbau des Mitarbeiterstabes sind mehrere Personen beteiligt. Die Kodierung enthält aber trotz der eingeführten Überprüfungen immer noch einen subjektiven Bias, der die Kategorisierung beeinflusst (Gaber 2000: 121). Hauptkritikpunkt ist allerdings die mangelnde Transparenz bezüglich der Kodierungsregeln, die die Kategorisierung nicht nachvollziehbar macht. Die mangelnde Transparenz betrifft auch die Informationsquellen, die nur sehr allgemein genannt werden. Ebenso wurde in den Jahresberichten sowie in den einzelnen Ländergutachten lediglich das Gesamtrating für die beiden Dimensionen *civil liberties* und *political rights* veröffentlicht und nicht die Einzelbewertungen (Munck/Verkuilen 2002: 19-21). Deswegen kommen Munck und Verkuilen (2002: 21) zu dem Schluss: „In the end, the aggregate data offered by Freedom House has to be accepted largely on faith“. Wohl als Reaktion auf diese Kritik veröffentlicht Freedom House seit 2006 auch die Bewertungen der Subkategorien, so auch die Werte für Rechtsstaatlichkeit.

Die Auseinandersetzung mit der Methodik von FIW erfolgte daher bislang zumeist nur mit dem Gesamtindex. Einige der in der bisherigen Debatte thematisierten Probleme lassen sich aber auch in der Subkategorie wiederfinden:

- Der Fragenkatalog hat sich über die Zeit geändert. Auch wenn Freedom House lieber von „checklist“ als von Indikatoren spricht, haben die Fragen letztendlich den Status von Indikatoren. Insofern bedeutet eine Änderung der Fragen eine Änderung der Indikatoren, was die zeitliche Konsistenz der Daten und damit ihre Vergleichbarkeit einschränkt. So wurde z.B. erst 2007 die Frage bezüglich der Rechte von Flüchtlingen aufgenommen, die angesichts der restriktiven Asylpolitik in vielen Ländern die Wertung ändern kann.
- Ein weiterer Punkt ist die Genauigkeit der Messung: Die unter den vier Leitfragen subsumierten Fragen umfassen teils verschiedene Aspekte, die unterschiedlich beantwortet werden können. Da die Kodierung und Aggregation der Daten unterhalb der Ebene der Subkategorien immer noch intransparent ist, kann dies nicht nachvollzogen werden. Warum die fünfgliedrige Skala (0-4) gewählt wurde

und wo die Schwellen zwischen den einzelnen Werten liegen, wird nicht dargestellt.

- Das Problem der Genauigkeit findet sich auch in der Trennschärfe der vier Dimensionen von Rechtsstaatlichkeit, die sich in bestimmten Aspekten überschneiden können. Beispielfhaft genannt sei hier die Frage „Sind Staatsanwälte unabhängig von politischer Kontrolle und Einfluss?“, die zur Dimension 2 (Rechtsstaatlichkeit in Zivil- und Strafrecht) gehört, aber genauso Dimension 1 (Unabhängigkeit der Justiz) messen könnte.
- Ähnliches gilt für die Problematik der Gewichtung und Aggregation. Gastil weist darauf, dass verschiedene Faktoren in Ländern unterschiedliche Bedeutung haben können und man deswegen keine generelle Gewichtung vornehmen kann (Gastil 1990: 29). So werden auch alle vier Komponenten bei Rechtsstaatlichkeit gleich gewichtet. Durch das Aggregieren per Addition statt Multiplikation können geringe Werte bei einer Komponente mit höheren Werten bei einer anderen ausgeglichen werden. Das heißt, dass ein Land nicht in allen Komponenten einen Mindestwert erreichen muss: Selbst wenn z.B. die Unabhängigkeit der Judikative mit 0 (nicht vorhanden) bewertet wird, kann die Gesamtbewertung der Subkategorie noch im Mittelmaß liegen. Dieses Problem setzt sich fort bei der weiteren Aggregation zu den beiden Kategorien und zum Gesamtindex. Wenn man sich die veröffentlichten drei Jahrgänge der *Rule of law*-Bewertungen von Freedom House ansieht, sind freilich alle Fälle, in denen *Rule of law* mit „0“ bewertet wurde, auch in der Kategorie „nicht frei“ angesiedelt.
- Da für die einzelnen Subkategorien keine Schwellenwerte festgelegt wurden, können auch Staaten, in denen die Rechtsstaatlichkeit nur im Mittelmaß liegt, als „frei“ eingestuft werden. Das ist z.B. in einigen lateinamerikanischen Staaten der Fall. Etliche als „teilweise frei“ bewertete Staaten besitzen bessere Rechtsstaatlichkeit-Werte als die „freien“. Auf der einen Seite zeigt das, dass die einzelnen Indikatoren nur in ihrer Wechselwirkung mit anderen zu bewerten sind und eine relativ hohe Rechtsstaatlichkeit allein noch keine Freiheit garantiert. Auf der anderen Seite ist es problematisch, dass ein Land, welches nicht einmal die Hälfte des geforderten Maßes an Rechtsstaatlichkeit erreicht, als frei eingestuft werden kann (Freedom House 2008b, 2007, 2006).

Bei den WGI werden für die Messung von Rechtsstaatlichkeit insgesamt 77 Indikatoren aus 26 Quellen herangezogen. Zwölf der Quellen werden als repräsentativ, 14 als nichtrepräsentativ eingestuft. Hierbei handelt es sich um regionale Surveys wie das Latinobarómetro oder um regionale Organisationen wie die Afrikanische Entwicklungsbank. Die Einstufung als nichtrepräsentativ ist für die spätere Gewichtung von Bedeutung (Kaufmann/Kraay/Mastruzzi 2008: 76f.). Die einzelnen Informationen werden in einem komplexen statistischen Verfahren, dem „Unobserved Components Model“, zu einem Teilindex für Rechtsstaatlichkeit umgerechnet (zu den technischen Details Kaufmann/Kraay/Mastruzzi 2008: 97ff, Arndt/Oman 2006:

103ff.). Das funktioniert, etwas vereinfacht, in fünf Schritten (Arndt/Oman 2006: 49ff):

- 1) Alle Indikatoren einer Quelle zu einem Land werden zusammengerechnet und ein einfacher, ungewichteter Durchschnitt genommen. Letztlich liegen nach diesem Verfahren noch insgesamt 24 Indikatoren für den Teilindex vor.
- 2) Jeder Indikator wird reskaliert, die 24 Quellen werden nach einer bestimmten Formel geographischer und Einkommensverteilung der untersuchten Länder in repräsentativ oder nichtrepräsentativ eingeteilt.
- 3) Die repräsentativen Einzelindikatoren werden gewichtet und aggregiert, je höher die Korrelation untereinander, umso größer die Gewichtung.
- 4) Nun werden die nicht-repräsentativen Quellen per Regressionsanalyse mit dem vorläufigen Index kombiniert, um die Standardabweichung zu errechnen, und wie oben erneut gewichtet. Dieses Verfahren führt im Endeffekt dazu, dass die kommerziellen Business Information Provider 67 % Anteile am Endindikator bekommen, Surveys nur 10 %, NGOs 10 % und öffentliche Stellen 12 % (Kaufmann/Kraay/Mastruzzi 2008: 30).
- 5) Schließlich werden die Einzelindikatoren erneut aggregiert und reskaliert, so dass der Durchschnitt der Werte bei Null liegt und jedes Land einen Rechtsstaatswert zwischen -2,5 und +2,5 erhält.

Die WGI sind zweifellos einer der umfassendsten Versuche, Rechtsstaatlichkeit zu messen. Drei Aspekte stechen positiv hervor: Zum einen die äußerst umfangreiche Quellenlage, die eine Vielzahl anderer Versuche, Rechtsstaatlichkeit zu messen, erfasst und berücksichtigt; zum zweiten die mit 212 erfassten Ländern überaus große Reichweite; und schließlich die Transparenz. Sämtliche technisch-mathematischen Details zur Berechnung der Indizes sind dokumentiert und alle Quellen detailliert aufgelistet. Kritikpunkte werden ebenfalls offen diskutiert, zum Teil sogar auf der Homepage der WGI selbst.

Diese positiven Aspekte dürfen aber nicht über einige wesentliche Schwächen hinwegtäuschen (zur kritischen Diskussion Arndt/Oman 2006, Iqbal/Shah 2008, Langbein/Knack 2008 sowie die Debatte in *The Journal of Politics* mit Kurtz/Schrank 2007a, Kaufmann/Kraay/Mastruzzi 2007a, Kurtz/Schrank 2007b, Kaufmann/Kraay/Mastruzzi 2007b):

- Mangelnde Transparenz: Zwar wurde zu Recht die Transparenz des Meta-Indikators gelobt, doch gilt dies nicht für alle Einzelquellen wie etwa Freedom House (Methodik) oder die Weltbankberichte CPIA, die nur teilweise öffentlich sind. Wiederum andere, vor allem die kommerziellen, sind nur gegen Bezahlung zugänglich. Business Risk Service etwa berechnet 575 US-Dollar für einen Länderbericht, Global Insight soll für die Global Risk-Daten 12700 US-Dollar verlangen (Arndt/Oman 2006: 72). Letztlich ist die Transparenz des Meta-Indikators dann nicht mehr überzeugend.

- Fehlerkumulation: Sie ergibt sich aus der Gestaltung als Meta-Indikator, d.h. methodische Fehler einzelner Quellen summieren sich. Es kann hier nicht auf alle Quellen eingegangen werden, aber ungenaue Fragen, nicht-transparente Auswahl der Befragten bzw. der Experten, was zu *selection-bias* führt, werden vielfach kumuliert. Teilweise sind Indikatoren doppelt oder dreifach vorhanden. So taucht etwa im Teilindikator *Rule of law* der WGI dreimal Rechtsstaatlichkeit als Einzelindikator auf. Wie erwähnt, tauchen dafür andere Aspekte, die möglicherweise Bestandteil von Rechtsstaatlichkeit sind, wie die Beachtung von Menschenrechten, in diesem Teilindikator gar nicht auf.
- Gewichtungen: Sie führen letztlich zu einer Überbewertung der kommerziellen Informationsanbieter, in den Worten von Iqbal/Shah zu „Western business perspectives“ (Iqbal/Shah 2008: 24). Dies beinhaltet einen ideologischen *bias* zugunsten von freier Marktwirtschaft, Eigentumsrechten, eine Ablehnung von Staatseingriffen etc.
- Varianz der Quellen: Die Anzahl der Länder, die von den WGI erfasst werden, variiert mit jeder Ausgabe, ebenso die Anzahl der Quellen wie auch die Art der Quellen (kommerzielle Anbieter, NGOs, Surveys oder staatliche Stellen). Zudem werden die Ergebnisse jedes Mal so reskaliert, dass der Durchschnitt bei Null liegt. Dies bedeutet, dass methodisch gesehen ein Vergleich der Länder ebenso unzulässig ist wie ein Vergleich eines Landes über die Zeit, da die jeweiligen Samples nicht übereinstimmen. Zwar wird das von den Autoren, gewissermaßen im Kleingedruckten der Methodik, eingeräumt, von den Anwendern in der Regel aber nicht beachtet. Nicht nur die Presse, auch Entwicklungsorganisationen sowie staatliche Stellen ziehen die Daten zu Vergleichszwecken heran. Im Millennium Challenge Account beispielsweise, dem Grundsatzprogramm der US-amerikanischen Entwicklungspolitik, wird explizit auf die WGI verwiesen, um Entwicklungen im Bereich der Good Governance (und damit auch in der Rechtsstaatlichkeit) zu beobachten, was direkt mit Mittelzuweisungen verbunden ist.
- Subjektivität der Indikatoren: Es handelt sich, wie erwähnt, ausschließlich um Umfragen oder Expertenmeinungen. Diese Kritik kontern die Autoren mit „perceptions matter because agents base their actions on their perceptions, impressions, and views“ (Kaufmann/Kraay/Mastruzzi 2008: 3). Hier geht es um grundlegende konzeptionelle Fragen zur Messung von sozialen Phänomenen überhaupt, die nicht vorschnell entschieden werden können, aber die Überlegung erscheint lohnenswert, ob nicht auch „objektive“ Daten, etwa Kriminalitätsraten oder die Anzahl von Richtern, in Kombination mit subjektiven Daten aussagekräftig wären. Die Autoren der WGI weisen diese Kritik zurück, sie sehen, trotz aller Defizite, keine sinnvolle Alternative zu ihrer Vorgehensweise (Kaufmann/Kraay 2008).

Der kurze Überblick über die Kritikpunkte zeigt, dass die WGI nicht unproblematisch sind; dennoch sollten die positiven Aspekte nicht übersehen werden. Die Viel-

falt der Quellen ist unübertroffen, was zwar einerseits Fehler und *biases* kumulieren, andererseits aber auch wieder relativieren könnte.

Die methodische Anlage des BTI ähnelt mehr dem Vorgehen von Freedom House als dem der Weltbank, doch ist der BTI konzeptionell und methodisch präziser konzipiert als der Freedom House Index. Grundlage der Bewertungen sind relativ ausführliche Ländergutachten. Sie basieren auf einem detaillierten Codebuch, das die verschiedenen Dimensionen, Kriterien und Indikatoren aufschlüsselt und erläutert. Für die numerische Bewertung werden pro Indikator vier Antwortvorgaben formuliert, die den Länderexperten die Orientierung vorgeben, inwieweit die Kriterien erfüllt sind. Diese vier Antwortniveaus sind auf einer Skala von 1 (nicht erfüllt) bis 10 (voll erfüllt) gespreizt und sie sind die Grundlage für die Rohdaten des gesamten Index. Ein anschließendes Reviewverfahren mit je einem Zweitgutachter pro Land, intra- und interregionalem Abstimmungsprozess sowie einem weiteren globalen Abgleich sollen die Verlässlichkeit der Daten sicherstellen. Die anschließende Indexbildung erfolgt durch die Bildung der arithmetischen Mittelwerte, die zunächst auf Ebene der Kriterien (Beispiel: Rechtsstaat) und anschließend auf Ebene der Dimensionen (Beispiel: Demokratie) erfolgt.

Aufgrund seiner methodischen Anlage kann der BTI als verlässlicheres Messinstrument der Demokratiemessung als etwa Freedom House gelten. Dennoch sind auch hier Kritikpunkte zu nennen (Müller/Pickel 2007):

- Zunächst bringt es die relativ komplexe Anlage als Transformationsindex mit sich, dass das Potential der zugrunde liegenden Konzepte nicht hinreichend ausgeschöpft wird. So hätte eine Anzahl von Indikatoren, die mehrdimensional konstruiert wurden, weiter operationalisiert werden können. Zum Beispiel weist etwa der Indikator „bürgerliche Freiheitsrechte“ in der Erläuterung des Codebuchs explizit eine Vielzahl von Aspekten auf, für die insgesamt aber eine gemeinsame Bewertung abgegeben wird (Mihir 2008: 149ff).
- Desweiteren fällt auf, dass der BTI einen ausgeprägten Schwerpunkt auf Gewaltenteilung legt. Dabei fokussiert das Kriterium „Unabhängigkeit der Justiz“ sowohl auf Gewaltenteilung als auch auf die Sicherung der bürgerlichen Freiheitsrechte. Andere Aspekte sind in dem Kriterium „Rechtsstaat“ teils nur implizit enthalten.
- Ein weiterer Kritikpunkt betrifft – ähnlich wie bei Freedom House – die Aggregation der Rohdaten und damit die Gewichtung der Indikatoren. Abgesehen von der Bildung des Gesamtindex, bei dem u.a. die Zusammenführung des Status-Index aus Demokratie und Marktwirtschaft kritisiert wurde, betrifft dies auch die Einzelkriterien. Zwar wird die Vorgehensweise der Aggregation – wie insgesamt die Methodik (Bertelsmann Stiftung 2008: 74ff) – ausführlich erläutert, doch fällt die theoretische Begründung des Verfahrens knapp aus.

- Da der BTI – aus nachvollziehbaren Gründen – ein eigenes Kriterium Staatlichkeit für die Messung der politischen Transformation insgesamt enthält, werden Beeinträchtigungen durch Krieg und Gewalt zumindest nicht explizit in die Rechtsstaatsmessung einbezogen, wie dies die beiden anderen Indizes tun. Umgekehrt ist festzuhalten, dass mit dem Indikator „Ahndung von Amtsmissbrauch“ ein Element aufgenommen wurde, das auf die Existenz informeller Regeln hinweist, die das Funktionieren des Rechtsstaates beeinträchtigen.

Im Hinblick auf die methodische Vorgehensweise vor allem bei der Messung und Aggregation unterscheidet sich der Weltbank-Index deutlich von FIW und BTI. Als „Meta-Index“ profitieren die *Governance Indicators* zwar von einer Fülle an Informationen zu den jeweils zu messenden Aspekten sowie von der insgesamt transparenten Vorgehensweise. Die methodische Raffinesse bringt jedoch auch Probleme mit sich, die vor allem in der Qualität der Rohdaten selbst (Transparenz, Subjektivität) und der faktischen Gewichtung liegen. Demgegenüber sind FIW und der BTI grundlegend als Ratings konzipiert, die unter den dafür typischen Problemen leiden. Dazu gehören neben den Gewichtungs- und Aggregationsregeln in erster Linie die Validität und Reliabilität der Messungen, die beide in unterschiedlicher Weise zu minimieren versuchen. Bei Freedom House kommt zudem die mangelnde Transparenz bei der Erstellung und Aggregation der Primärdaten hinzu.

Tabelle 2: Methodik der Rechtsstaatsmessung im Vergleich

	FIW	WGI	BTI
Verfahren	Rating / Ranking	Meta-Index	Rating / Ranking
Indikatorenauswahl	unklare Begründung	unklare Begründung	theoretische Begründung
Trennschärfe	teils überlappend	teils überlappend	teils überlappend
Datenbasis	„subjektiv“	„objektiv“*	„subjektiv“
Aggregation	Standard (aber ohne Begründung)	elaboriert	Standard
Reliabilität	eingeschränkt	hoch*	eher hoch
Transparenz	unzureichend	hoch*	hoch

* die Bewertung bezieht sich auf die Governance Indicators als solche, nicht auf die Rohdaten (zu dieser Problematik siehe die Ausführungen in Kap. 4).

5. Rechtsstaatsmessung und Rechtswirklichkeit

Die aufgezeigten Unterschiede in Anlage, Konzeption und Messverfahren sowie deren Qualität legen die Frage nahe, welches Bild der Wirklichkeit des Rechtsstaates gezeichnet wird und ob die Messungen relativ übereinstimmen. Hier zeigt die statistische Analyse, dass sich die Messungen auf unterschiedliche Aspekte der Rechtsstaatlichkeit beziehen und somit die Rechtswirklichkeit nicht identisch abbilden. Ein

Hinweis auf die Validität von Messungen wird gemeinhin in einer hohen statistischen Korrelation mit anderen Messungen desselben Phänomens gesehen. Dabei ist freilich zu berücksichtigen, dass stark korrelierende Messungen möglicherweise einen *bias* beziehungsweise Irrtümer enthalten. Dennoch ist eine hohe Korrelation zumindest ein Hinweis darauf, dass ähnliche Aspekte der Realität gemessen werden. Die Ergebnisse zeigen, dass dies nicht durchweg der Fall ist. So existiert zwar zwischen Freedom House und dem Bertelsmann Transformationsindex eine sehr hohe Korrelation (0,938) doch korrelieren beide deutlich geringer mit der Messung der Weltbank (0.664 bzw. 0.663).² Hierbei ist zusätzlich zu berücksichtigen, dass die Weltbank-Messung die beiden anderen Indizes bei ihrer Daten-Aggregation berücksichtigt. Dies muss als Hinweis darauf gewertet werden, dass in die Messung der Weltbank deutlich andere Aspekte der Rechtswirklichkeit einfließen. Wie der Blick auf die Konzepte der Rechtsstaatlichkeit zeigte, ist diese Abweichung zu erklären mit der unterschiedlichen Konzeption des Rechtsstaates und damit der unterschiedlichen Anlage und Gewichtung der Indikatoren, insbesondere die deutlich stärkere Gewichtung des materiellen Rechtsstaates bei FIW und BTI.

Die Diskrepanz zwischen FIW und BTI einerseits sowie WGI andererseits bestätigt sich auch, wenn man den Blick auf die Streuung der Messergebnisse richtet und somit auch die deutlichen Abweichungen in Einzelfällen erkennbar werden. Das Beispiel Singapur verdeutlicht dies am besten: Bei FIW und BTI, deren Bewertung der Rechtsstaatlichkeit deutlich durch eine stärkere Gewichtung demokratischer Verhältnisse geprägt ist, landet Singapur nur auf einem mittleren Rang, während die Weltbank die Rechtsstaats-Qualität des autoritär regierten Singapur (wie auch anderer Autokratien wie etwa im Nahen Osten: Saudi-Arabien, Bahrain, Oman) deutlich höher bewertet – Singapur erzielt gar den höchsten Wert für die Nicht-Industrieländer. Schon dies deutet darauf hin, dass solche Messungen nicht umstandslos für politische Beratung und Gestaltung eingesetzt werden sollten.

Mit den genannten Vorbehalten weisen die Ergebnisse der drei Indizes darauf hin, dass der Rechtsstaat in vielen Staaten nur unzureichend funktioniert und damit ein Hindernis für die Entfaltung der Demokratie und weiterer gesellschaftlicher Funktionsbereiche darstellt. So liegen die Mittelwerte von BTI bzw. FIW für das BTI-Sample der 125 Transformationsländer jeweils unter dem Mittelpunkt der Skala, d.h. vereinfacht ausgedrückt: Sie erreichen zusammen nicht einmal ein „mittleres Niveau“ an Rechtsstaatlichkeit. Dieses Bild bessert sich zwar, wenn nur die 75 Demokratien des Samples betrachtet werden, doch bleiben auch sie weit unter dem Standard eines leidlich funktionierenden Rechtsstaates. Ein ähnliches Bild zeigen die *Governance Indicators*: Hier lässt sich aus methodischen Gründen zwar kein aussagekräftiger Mittelwert errechnen, doch weist die relativ solide Korrelation mit den beiden anderen Indizes auf einen ähnlichen Befund hin. Das lässt darauf schließen, dass die drei

² Berechnung der Rangkorrelationen (Spearman-Rho) für das Sample der 125 Länder des BTI 2008.

Indizes zumindest relevante Ausschnitte der Qualität der Rechtsstaatlichkeit zuverlässig abbilden.

In den Konzepten und Indikatoren der drei Messungen spiegeln sich trotz der vorgebrachten Einwände grundlegende Aspekte der Rechtsstaatlichkeit wider. Es ist somit davon auszugehen, dass sie kein völlig verzerrtes Abbild dieser Realität liefern, sondern Auskunft über den Zustand des Rechtsstaates weltweit geben. Dabei geben die Messungen zumindest teilweise – je nach Prägnanz und Aufschlüsselung der Indikatoren – auch Auskunft darüber, in welchen Funktionsbereichen zentrale Defizite liegen und wie diese über das jeweilige Ländersample hinweg variieren.

Allerdings lassen die Messungen die Frage offen, welche Rechtsverhältnisse genau sich ergeben, wenn der Rechtsstaat nur unzureichend funktioniert. Geht man davon aus, dass in solchen Fällen kein Normenvakuum vorherrscht, stellt sich des Weiteren die Frage, welche Normen oder Normensysteme entweder an die Stelle des Rechtsstaates treten oder mit ihm koexistieren. Mit anderen Worten mangelt es den etablierten Indizes an einem vollständigen Bild der Rechtswirklichkeit und damit auch an einer Einschätzung der tatsächlichen Gefährdung der Rechtsstaatsentwicklung. Dies geht einher mit der Tatsache, dass im Unterschied zur Demokratie die rechtsstaatliche Entwicklung bislang in der Forschung weitaus weniger differenziert betrachtet wurde. So spricht Carothers, der als Berater jahrelang Rechtsstaatsreformen aus wissenschaftlicher wie aus entwicklungspraktischer Sicht begleitet hat, von einem „problem of knowledge“ (Carothers 2006), das sich sowohl auf systematische und verallgemeinerbare Erkenntnisse als auch auf das Verständnis spezieller Dynamiken erstreckt und insbesondere die Wirkungsmechanismen von Reformen einschließt.

Neuere Ansätze versuchen deshalb, die Rechtsstaatsentwicklung differenzierter zu betrachten. Ein wesentlicher Schritt besteht darin, zur Systematisierung verschiedener Rechtsstaatstypen zwischen defizitärem Rechtsstaat und hybridem Rechtssystem zu unterscheiden (Lauth 2001, Lauth/Sehring 2009). Ein defizitärer Rechtsstaat erfüllt zwar nicht mehr alle Kriterien des Rechtsstaats, dominiert aber noch über alternative Regelsysteme. In einem hybriden Rechtssystem dagegen ist der Rechtsstaat eines von zwei oder mehreren Rechtssystemen, ohne mehr Durchsetzungskraft als diese zu besitzen. Dabei wurden auch grundlegende Zusammenhänge analysiert, die zu solchen Rechtsformen führen und nicht in einen funktionierenden Rechtsstaat münden. Verschiedene Ursachenbündel wurden betrachtet: a) die unzureichende administrative und finanzielle Ausstattung, b) die vorhandenen Macht- und Interessenkonstellationen und c) die Existenz informeller Normensysteme. Dieser letzte Punkt ist als zentraler Faktor für die mangelnde Rechtsstaatlichkeitsausprägung anzusehen. Dies bedeutet nicht, dass mangelnde Ressourcen und entgegenlaufende Interessen keine Bedeutung besitzen. Vielmehr ist es gerade die Einbettung der Akteure in den Kontext informeller Normensysteme, die deren Interessen und Ressourcen modelliert. Je nach Stärke und Ausprägung informeller Normensystemen ergeben sich unterschiedliche Typen des Rechtsstaats.

Informelle Normensysteme fungieren dabei als informelle Institutionen, die definiert werden als “socially shared rules, usually unwritten, that are created, communicated, and enforced outside of officially sanctioned channels” (Helmke/Levitsky 2004: 727). Während formal verfasste Institutionen also im offiziellen Code des politischen Systems verankert und somit als Ausdruck rationaler Herrschaft – in der Festlegung durch Verfassung und Gesetze – schriftlich kodifiziert sind, wirken informelle Institutionen außerhalb dieser Festlegung. Sie sind nicht durch formale Vereinbarungen fixiert und gewinnen ihre Bedeutung allein durch ihre faktische verhaltensprägende und verhaltensstrukturierende Existenz, mit der sie Einfluss auf die Struktur und Funktion des politischen Prozesses ausüben (Lauth 2004). Während formale Institutionen von Akteuren mit Regelungskompetenz sinnhaft gestaltet und verändert werden können (Mayntz/Scharpf 1995: 45), gilt dies bei informellen Institutionen nicht, da sie quasi naturwüchsig entstehen. Sie besitzen kein steuerndes und sie koordinierendes Zentrum (Lauth 2004), weshalb sie bei Vertretern des institutionenökonomischen Ansatzes als kaum beeinflussbare Konstanten gelten (North 1992: 7).

Eine angemessene Einschätzung der Qualität des Rechtsstaates kann somit nur unter Einbeziehung von informellen Institutionen geschehen (O’Donnell 1999). Dies gilt umso mehr, wenn man bedenkt, dass diesen unter Umständen der Rang einer zweiten Verfassung zukommen kann (Waldmann 2001). Dabei ist es durchaus möglich, dass informelle Institutionen zu einer positiven Entwicklung des Rechtsstaates beitragen, aber es ist gleichfalls möglich, dass sie diese behindern, indem sie die Funktionsweise formaler Institutionen unterminieren. Relevant sind hier vor allem konkurrierende Regelsysteme in Form von Korruption und Klientelismus.

Vor diesem Hintergrund wird deutlich, dass die hier analysierten Indizes zwar relativ gut das Nicht-Funktionieren des Rechtsstaates abbilden, nicht aber das Funktionieren des daraus resultierenden Normensystems. Mit anderen Worten messen sie zwar die Verwirklichung des Rechtsstaates, nicht aber die gesamte Rechtswirklichkeit eines Landes. So kann ein nicht voll funktionierender Rechtsstaat durch ein *responsive customary law* ergänzt werden. Er kann aber auch defizitär sein, weil in bestimmten funktionalen oder territorialen Bereichen mächtige Akteure (Guerilla, Oligarchen, etc.) ein eigenes Regelsystem aufgebaut haben, das sich staatlicher Kontrolle entzogen hat oder toleriert wird. In beiden Beispielen könnte die Messung denselben Wert für Rechtsstaatlichkeit ergeben, die Rechtswirklichkeit sähe jedoch sehr unterschiedlich aus. Dies ist insbesondere dann von Bedeutung, wenn Aussagen über die Dynamik und Stabilität des Rechtsstaates getroffen oder gar praktische Schlüsse – etwa im Hinblick auf Rechtsstaatsreformen – gezogen werden sollen.

6. Fazit

Die Messung der Rechtsstaatlichkeit hat in den letzten Jahren zweifelsohne Fortschritte erzielt. Anhand der Indizes von Freedom House, Weltbank und Bertelsmann-Stiftung konnte gezeigt werden, dass diese trotz einiger konzeptioneller und metho-

discher Probleme zentrale Aspekte der Wirklichkeit des Rechtsstaates abbilden und sich für weitere Analysen verwenden lassen. Dennoch zeigen sich auch die Grenzen solcher Ansätze für ein komplexes Phänomen wie Rechtsstaatlichkeit. Der quantitativen Ansätzen inhärenten großen Reichweite und Präzision stehen Rigidität und Reduktion von Komplexität gegenüber. Neben der weiteren Eliminierung von genuinen Schwachstellen in Konzeptualisierung und Methodik wird in Zukunft allerdings auch dem Aspekt der informellen Institutionen größere Aufmerksamkeit zu widmen sein, um die gesamte Rechtswirklichkeit gerade in defizitären Rechtsstaaten adäquater erfassen zu können. Dies lässt sich auf der Grundlage der etablierten Indizes allerdings kaum bewerkstelligen, ohne deren Grunddesign grundlegend zu ändern und damit die Vergleichbarkeit über die Zeit stark zu beeinträchtigen. Zudem kann ohne qualitative Einzelfallstudien oder vergleichende qualitative Analysen das Phänomen der informellen Institutionen (noch) nicht adäquat verstanden werden. Das Dilemma qualitativer Studien liegt allerdings wiederum darin, dass aufgrund der notwendigen multifaktoriellen Untersuchungsanlagen Aussagen über kausale Zusammenhänge immer begrenzt sind und die Vergleichbarkeit wegen des starken Kontextbezuges zumeist nicht gegeben ist. Dies bedeutet letztlich, dass weder quantitativ-statistische Untersuchungen noch qualitative Einzelfallstudien oder vergleichende qualitative Analysen allein ausreichend sind, um Zustand und Funktionsweise von Rechtsstaatlichkeit beziehungsweise seine Ursachen und Wirkungen adäquat zu analysieren.

Literatur

Arndt, Christiane/ Oman, Charles (2006) *Uses and Abuses of Governance Indicators*, Paris.

Bertelsmann Stiftung (Hrsg.) (2008) *Bertelsmann Transformation Index 2008. Politische Gestaltung im internationalen Vergleich*, Gütersloh.

Bertelsmann Stiftung (Hrsg.) (2009) *Sustainable Governance Indicators 2009*, Gütersloh.

Carothers, Thomas (2006): *The Problem of Knowledge*, in: Carothers, Thomas (Hrsg.): *Promoting the Rule of Law Abroad: In Search of Knowledge*, Washington, DC: 15ff.

Dworkin, Ronald (1998) *Freiheit, Selbstregierung und der Wille des Volkes. Ist Demokratie heute noch möglich?*, in: Gosepath, Stefan/ Lohmann, Georg (Hrsg.): *Philosophie der Menschenrechte*, Frankfurt a.M.: 292ff.

Freedom House (2006): *Tables and Charts*. <http://www.freedomhouse.org/template.cfm?page=25&year=2006> (Zugriff 20.08.2008).

Freedom House (2007) *Tables and Charts*. <http://www.freedomhouse.org/template.cfm?page=25&year=2007> (Zugriff 20.08.2008).

- Freedom House (2008a) Methodology. <http://www.freedomhouse.org/template.cfm?page=351&year=2006> (Zugriff 20.08.2008).
- Freedom House (2008b) Tables and Charts. <http://www.freedomhouse.org/template.cfm?page=25&year=2008> (Zugriff 20.08.2008).
- Gaber, Rusanna (2000) Demokratie in quantitativen Indizes: Ein mehr- oder eindimensionales Phänomen? In: Lauth, Hans-Joachim/ Pickel, Gert/ Welzel, Christian (Hg.), *Demokratiemessung. Konzepte und Befunde im internationalen Vergleich*. Wiesbaden: 112ff.
- Gastil, Raymond D. (1990) The Comparative Survey of Freedom: Experiences and Suggestions, *Studies in Comparative International Development* 25 (1): 25ff.
- Habermas, Jürgen (1992) *Faktizität und Geltung. Beiträge zur Diskurstheorie des Rechts und des demokratischen Rechtsstaats*, Frankfurt a.M.
- Habermas, Jürgen (1994) Über den internen Zusammenhang von Rechtsstaat und Demokratie, in: Preuß, Ulrich K. (Hrsg.): *Zum Begriff der Verfassung. Die Ordnung des Politischen*, Frankfurt a.M.: 83ff.
- Helmke, Gretchen/ Levitsky, Steven (2004) Informal Institutions and Comparative Politics: A research Agenda, *Perspective on Politics* 2 (4): 725ff.
- Huntington, Samuel P. (1991) *The Third Wave. Democratization in the Late Twentieth Century*, Norman/London.
- Iqbal, Kazi/ Shah, Anwar (2008) How do worldwide governance indicators measure up (<http://siteresources.worldbank.org/PSGLP/Resources/Howdoworldwidegovernanceindicatorsmeasureup.pdf>) (Zugriff 30.9.2008).
- Kaufmann, Daniel/ Kraay, Art (2008) Governance Indicators: Where are we, where should we be going?, *The World Bank Research Observer* 23 (1): 1ff.
- Kaufmann, Daniel/ Kraay, Aart/ Mastruzzi, Massimo (2007a) Growth and Governance: A Reply, *The Journal of Politics* 69 (2): 555ff.
- Kaufmann, Daniel/ Kraay, Art/ Mastruzzi, Massimo (2007b) Growth and Governance: A Rejoinder, *The Journal of Politics* 69 (2): 570ff.
- Kaufmann, Daniel/ Kraay, Art/ Mastruzzi, Massimo (2008) *Governance Matters VII: Aggregate and Individual Governance Indicators 1996-2007*, Policy Research Working Paper 4654, Worldbank, Washington.
- Kaufmann, Daniel/ Kraay, Art/ Zoido-Lobaton (1999) *Governance Matters*, Policy Research Working Paper 2196, Worldbank, Washington.
- Kurtz, Marcus/ Schrank, Andrew (2007a) Growth and Governance: Models, Measures, and Mechanisms, in: *The Journal of Politics* 69 (2): 538ff.

Kurtz, Marcus/ Schrank, Andrew (2007b) Growth and Governance: A Defense, *The Journal of Politics* 69 (2): 563ff.

Landman, Todd (2003) Map-Making and Analysis of the Main International Initiatives on Developing Indicators on Democracy and Good Governance, Eurostat Contract No. 200221200005, Final Report, Human Rights Centre, Essex.

Langbein, Laura/ Knack, Stephen (2008) The Worldwide Governance Indicators and Tautology: Causally related separable concepts, indicators of a common cause, or both? Policy Research Working Paper 4669, Worldbank, Washington.

Lauth, Hans-Joachim (2004) Demokratie und Demokratiemessung, Wiesbaden.

Lauth, Hans-Joachim (2001) Rechtsstaat, Rechtssysteme und Demokratie, in: Becker, Michael/ Lauth, Hans-Joachim/ Pickel, Gert (Hg.), *Rechtsstaat und Demokratie*, Wiesbaden: 21ff.

Lauth, Hans-Joachim/ Sehring, Jenniver (2009) Putting the deficient Rechtsstaat on the research agenda. Reflections on diminished subtypes, *Comparative Sociology* 8 (2), 2009: 165ff.

Mayntz, Renate/ Scharpf, Fritz W. (1995) Der Ansatz des akteurszentrierten Institutionalismus, in: dies. (Hrsg.): *Gesellschaftliche Selbstregulierung und politische Steuerung*, Frankfurt a.M.: 39ff.

Merkel, Wolfgang/ Puhle, Hans-Jürgen/ Croissant, Aurel/ Eicher, Claudia/ Thiery, Peter (2003) *Defekte Demokratie*. Band 1: Theorie, Opladen.

Mihr, Anja (2008) Demokratie und Menschenrechte. Der dritte Bertelsmann Transformation Index 2008, *Zeitschrift für Menschenrechte* Nr. 1: 143ff.

Müller, Thomas/ Pickel, Susanne (2007) Wie lässt sich Demokratie am besten messen? Zur Konzeptqualität von Demokratie-Indizes, *Politische Vierteljahresschrift* 48 (3): 511ff.

Munck, Geraldo L./ Verkuilen, Jan (2002) Conceptualizing and measuring democracy. Evaluating Alternative Indices, *Comparative Political Studies* 35 (1): 5ff.

North, Douglass (1992) Institutionen, institutioneller Wandel und Wirtschaftsleistung, Tübingen.

O'Donnell, Guillermo (1999) Illusions about Consolidation, in: ders., *Counterpoints. Selected Essays on Authoritarianism and Democratization*, Notre Dame: 175ff (reprint aus *Journal of Democracy* 7, 2 (April 1996)).

Waldmann, Peter (2001) Rechtsstaat und alternative Normsysteme in Lateinamerika, in: Becker, Michael/ Lauth, Hans-Joachim/ Pickel, Gert (Hrsg.), *Rechtsstaat und Demokratie*, Wiesbaden: 112ff.

Zimmerling, Ruth (2003) Samuel Huntingtons demokratische Wellen – viel Lärm um Gischt? *Politische Vierteljahresschrift* 44 (2): 199ff.

Autoren

Dr. Peter Thiery (1959) Centrum für angewandte Politikforschung (C.A.P.), Ludwig-Maximilians-Universität München; Institut für Politikwissenschaft und Sozialforschung, Universität Würzburg. Zuvor Mitarbeiter an den Universitäten Heidelberg und Mainz. Studium der Politikwissenschaft, Germanistik, Publizistik und Philosophie in Mainz. Forschungsschwerpunkte: Demokratie, Rechtsstaat, Lateinamerika.

E-Mail: peter.thiery@lrz.uni-muenchen.de

Dr. Jenniver Sehring (1974) Institut für Politikwissenschaft und Sozialforschung, Universität Würzburg. Zuvor wissenschaftliche Mitarbeiterin an der Fernuniversität in Hagen und der Universität Gießen. Studium der Politikwissenschaft und Ethnologie in Mainz. Forschungsschwerpunkte: Autoritäre Regime, Wasserpolitik, Zentralasien.

E-Mail: jenniver.sehring@uni-wuerzburg.de

Dr. Wolfgang Muno M.A. (1968) Institut für Politikwissenschaft, Universität Mainz, Institut für Politikwissenschaft und Sozialforschung, Universität Würzburg. Studium der Politikwissenschaft, Jura und Ethnologie in Mainz und Caracas, Venezuela. Forschungsschwerpunkte: Lateinamerika (vergleichende Politik und Ökonomie), Internationale Beziehungen (Entwicklungstheorien, Nord-Süd-Beziehungen).

E-Mail: muno@uni-mainz.de