

# Möglichkeiten und Grenzen der Demokratiemessung

von Hans-Joachim Lauth

*Die Bedeutung der Demokratiemessung in der empirischen Demokratieforschung wächst kontinuierlich und trifft sich mit dem generellen Interesse der Rechtswissenschaften an der Normmessung. Ein wesentlicher Grund liegt in der gestiegenen Anzahl von Demokratien und der damit gewachsenen Varianz ihrer Ausprägungen, die nicht zuletzt in dem unterschiedlichen Grad ihrer Qualität besteht. Deren Bestimmung ist das maßgebliche Anliegen der noch jungen Subdisziplin der Demokratiemessung. Zentrale quantifizierende Ansätze dieses Bereichs werden vorgestellt und anhand gängiger Kriterien beurteilt. Dazu gehört die Konsistenz des Demokratiekonzeptes sowie die Validität und Reliabilität der Messungen. Auch wenn sich dabei markante Unterschiede zeigen, so sind einige Messanlagen durchaus in der Lage, ein differenziertes Bild der demokratischen Entwicklung aufzuzeigen.*

*The growing need to measure democracy empirically coincides with the general legal interest in the measurement of norms. This is mainly due to the noticeable rise in the number of democracies and the subsequently expanding variance of their different shapes and forms, not least in relation to the observable degree of democratic quality. Determining this quality is the core concern of the young sub-discipline that seeks to measure democracy. Central quantitative approaches will be introduced and judged on the basis of accepted criteria, not least the internal consistency of the respective conceptualisations of democracy and the validity and reliability of the measures used. Even though some important distinctions are apparent, several measures are sufficiently capable of drawing a differentiated picture of democratic development.*

## I. Einleitung

Demokratiemessung verfolgt mit der Regimeklassifizierung und der Bestimmung der Qualität der Demokratie zwei Ziele, die sich in der Vergangenheit maßgeblich auf Transformationskontexte konzentrierte. Inzwischen richtet sich das Interesse auch auf etablierte Demokratien, denen aus verschiedener Perspektive ein Verlust der Qualität der Demokratie konstatiert wird. Vielfältige Anfra-

*Anmerkung der Schriftleitung:* Ein ergänzender Beitrag des Autors, der sich auf die qualitativen Ansätze zur Demokratiemessung richtet, erscheint im kommenden Heft der ZSE.

gen kommen aus der Globalisierungsdebatte und damit verbundenen Bedenken zur demokratischen Steuerung politikrelevanter Entscheidungen. Auch der Abwehr des internationalen Terrorismus wird attestiert, dass sie zu viele demokratische und bürgerliche Rechte auf diesem Wege einschränke.<sup>2</sup> Diese und weitere Probleme werden in manchen Ansätzen dermaßen überzeichnet, dass bereits vom Ende der Demokratie oder von einer postdemokratischen Epoche gesprochen wird.<sup>3</sup> Auch wenn diese Einschätzung hier nicht geteilt wird, so verdeutlichen die Anmerkungen doch, dass es ausreichend Indizien für die Veränderung auch etablierter Demokratien gibt und dass es aufgrund wechselnder Forschungsperspektiven und Maßstäbe zugleich schwierig ist, eine in sich stimmige Beurteilung der Demokratieentwicklung zu geben.

Gefordert ist daher eine Demokratiemessung, die plausibel Auskunft über den Grad und die Qualität der Demokratie geben soll.<sup>4</sup> Die bisherigen Forschungsanstrengungen auf diesem Gebiet werden im nachfolgenden Beitrag aufgegriffen, um zu prüfen, inwieweit deren Verfahren und Ergebnisse für die Einschätzung demokratischer Qualität hilfreich und brauchbar sind. Dabei geht es sowohl um die Klärung des jeweiligen Maßstabs als auch um die methodische Vorgehensweise. Generelles Grundverständnis ist die Überzeugung, dass mit Demokratie ein soziales Phänomen vorliegt, das analog eines Idealtypus von *Max Weber* oftmals graduell ausgeprägt ist. Daher ist es sinnvoll, Aussagen über den Grad der Ausprägung zu treffen.

- 2 Zu den USA vgl. *Braml, J./Lauth, H.-J.*: The United States of America. A Deficient Democracy, in: Erdmann, G./Kneuer, M. (eds.): *Regression of Democracy*, Wiesbaden, 2011, i.E.; *Braml, J.*: Gewaltenteilung und Demokratie. Konzepte und Probleme der „Horizontal Accountability“ im interregionalen Vergleich, Baden-Baden, 2007, 191-212; *Lauth, H.-J.*: Deficient Democracies. Qualität und außenpolitische Relevanz der Demokratie aus Sicht der empirischen Demokratietheorie, in: Hils, J./Wilzewski, J. (Hg.): *Defekte Demokratie – Crusader State? Die Weltpolitik der USA in der Ära Bush*, Trier, 2006, 77-108.
- 3 *Guéhenno, J.M.*: Das Ende der Demokratie, München/Zürich, 1994; *Agamben, G.*: *Homo sacer. Souveräne Macht und bloßes Leben*, Frankfurt/Main, 2002; *Crouch, C.*: *Postdemokratie*, Frankfurt/Main, 2008.
- 4 *Lauth, H.-J./Pickel, G./Welzel, C.* (Hg.): *Demokratiemessung*, Opladen, 2000; *Munck, G. L./Verkuilen, J.*: Conceptualizing and Measuring Democracy. Evaluating Alternative Indices, in: *Comparative Political Studies*, 35/1 (2002), 5-34; *Landman, T.*: Map-Making and Analysis of the Main International Initiatives on Developing Indicators on Democracy and Good Governance, University of Essex – Human Rights Centre, 2003; *Lauth, H.-J.*: Demokratie und Demokratiemessung. Eine konzeptionelle Grundlegung für den interkulturellen Vergleich, Wiesbaden, 2004; *Diamond, L./Morlino, L.* (eds.): *Assessing the Quality of Democracy*, Baltimore, 2005; *Müller, T./Pickel, S.*: Wie lässt sich Demokratie am besten messen? Zur Konzeptqualität von Demokratie-Indizes, in *PVS*, 3 (2007), 511-539.

## II. Grundlagen der Demokratiemessung

Der gemeinsame Fokus für die Demokratie verdeckt oftmals die unterschiedlichen Auffassungen, die mit dieser Vorstellung verbunden werden. Eine zentrale Konfliktlinie verläuft zwischen den beiden Polen einer substantiellen oder materiellen Demokratiekonzeption auf der einen Seite und einer prozeduralistischen Demokratievorstellung auf der anderen Seite. In der ersten Version wird die Qualität maßgeblich durch das materielle Ergebnis für das Volk bestimmt, während in der zweiten Version Verfahren der Partizipation ausschlaggebend sind. Mit der Niederlage substanzieller Demokratiekonzeptionen im Kleide des Realsozialismus ist der Siegeszug der prozeduralistischen Demokratie, wie sie alle westlichen Demokratien kennzeichnet, eng verbunden. Dennoch ist die Frage nicht obsolet geworden, inwieweit der *output* bzw. der *outcome* die Qualität einer Demokratie auch prägt, prägen kann oder sogar muss.<sup>5</sup> Fragen der Gleichheit und der Gerechtigkeit werden in verschiedenen Facetten immer wieder zur Beurteilung von Demokratien herangezogen. Gleichfalls werden Kriterien von *good* bzw. *bad governance* als relevant für die Qualität der Demokratie erachtet.

Auch die ideale Form der prozeduralistischen Demokratie ist umstritten. So konkurrieren direktdemokratische mit repräsentativen Verfahren, Mehrheits- mit Konsensprozeduren oder elite- mit basisorientierten Partizipationsmodellen. Einen besonderen Akzent auf qualitative Verfahren legen die verschiedenen Varianten deliberativer Demokratie. Trotz dieser beachtlichen Divergenz beziehen sich alle prozeduralistischen Formen auf zwei basale Dimensionen der Demokratie: politische Freiheit und politische Gleichheit. Diese finden ihren ursprünglichsten Ausdruck im freien und gleichen Wahlrecht. Eine dritte fundamentale Dimension ist im Wahlakt angelegt: in der zeitlichen Begrenzung der Wahlentscheidung. Demokratie ist Ausdruck einer begrenzten Herrschaftsform. Die elementarste Begründung für diese Begrenzung der Herrschaft, dass also die Mehrheit nicht über alles verfügen darf, liefern die individuellen Menschenrechte, die nicht zur Disposition gestellt werden dürfen, ohne die Kernidee der Demokratie – die freiheitliche Selbstbestimmung aller Bürger – selbst zu verletzen. Damit erhält zugleich die *output*-Dimensionen eine demokratietheoretische Bestimmung, wenngleich *in negativo*. Ausgeschlossen werden alle Entscheidungen, welche die Grundrechte der Bürger unterminieren oder gar abschaffen wollen. Damit würde eine Demokratie zumindest an den Beschlüssen erkannt werden können, die sie nicht fällt. Aber es gibt durchaus auch Hinweise

5 Scharpf, F.W.: Demokratietheorie zwischen Utopie und Anpassung, Konstanz, 1970.

auf die Möglichkeit einer positiven Bestimmung der *output*-Gestaltung von Demokratien, ohne in Aporien der Gemeinwohlbestimmung zu verfallen – und zwar, indem die Voraussetzungen der Demokratie beachtet werden.

Ohne Zweifel bedarf die Ausübung der Partizipationsrechte bestimmter sozialer Grundvoraussetzungen, ohne die eine freie und egalitäre Ausübung der Rechte nicht möglich ist. Der Staat hat somit die individuelle Nutzungsmöglichkeit der Rechte zu garantieren, wenn die Qualität einer Demokratie (oder diese sogar in Gänze) nicht beeinträchtigt werden soll. Auch der Umfang der damit verbundenen Staatstätigkeit wird kontrovers diskutiert, gänzlich abgestritten wird er nicht. Die wohl elementarste Form der Sicherung von Rechten und Begrenzung der Herrschaft, die gerade auf die Wahrung der individuellen Rechte zielt, ist der Rechtsstaat. Er ist in diesem Sinne grundlegend mit der Demokratie verbunden, die ohne ihn ihre notwendige rechtliche Fundierung der Herrschaftsweise verliert.<sup>6</sup> Gleichwohl ist der Rechtsstaatsbezug auch nicht unumstritten, wie die verwandte Diskussion zum Konstitutionalismus zeigt.

Ein anderes Kriterium, das mit der Qualität der Demokratie in Verbindung gebracht wird, ist die Effektivität der Herrschaft. Wenn damit auf das Gewaltmonopol des Staates abgezielt werden soll, dann könnte zu Recht auf die Rechtsstaatlichkeit verwiesen werden, die diesen Aspekt bereits erfasst. Auch die mit der Staatstätigkeit notwendige Verwaltungsstruktur wäre damit zu erfassen. Im demokratietheoretischen Sinne lässt sich ein anderer Aspekt beleuchten, der auf der grundlegenden Idee der Responsivität oder *responsive rule* beruht.<sup>7</sup> Demnach sollte demokratische Herrschaft die maßgeblichen Präferenzen der Bürgerinnen und Bürger beachten. Die Effektivität ihrer Umsetzung wäre demnach ein Kriterium der Qualität einer Demokratie. Allerdings stößt diese Präferenz Erfüllung hinsichtlich der notwendigen rechtsstaatlichen Garantien an ihre Grenzen.<sup>8</sup> Gleichfalls kann das Responsivitätsprinzip mit dem Verantwortlichkeitsprinzip kollidieren, das beispielsweise auch die Interessen zukünftiger Generationen einbezieht. Weitere Qualitätskriterien werden mit der Qualität der öffentlichen Kommunikation und der politischen Rechtfertigung verbunden sowie mit der Transparenz der politischen Prozesse.

6 *Habermas, J.*: Faktizität und Geltung, Frankfurt/Main, 1992; *O'Donnell, G.*: Why the Rule of Law matters, in: *Journal of Democracy*, 15/4 (2004), 32-46.

7 *Saward, M.*: *The Terms of Democracy*, Cambridge, 1998.

8 Vgl. *Lauth, H.-J.*: Quality Criteria for Democracy. Why Responsiveness is not the Key, in: Erdmann, G.; Kneuer, M. (eds.): *Regression of Democracy*, Sonderheft 1 der ZfVP, Wiesbaden, 2011, i.E.

Die Liste aller Kriterien, die für die Qualitätsbestimmung einer Demokratie angeführt werden, ist mit den bislang angesprochenen Dimensionen sicherlich noch nicht erschöpft.<sup>9</sup> Gleichwohl verdeutlichen diese Überlegungen, dass es alles andere als selbstverständlich ist, einen allgemein akzeptablen Maßstab zu formulieren, sieht man von der allgemeinen abstrakten Formulierung der Dimensionen der Freiheit, Gleichheit und Kontrolle ab, die sich in unterschiedlichen Ausprägungen und Akzentuierungen in allen prozeduralistischen Demokratiedefinitionen findet. So geht es auch in diesem Beitrag nicht darum, einen entsprechenden Maßstab zu formulieren. Ziel ist vielmehr herauszuarbeiten, welche Maßstäbe verwendet werden und welche empirischen Ergebnisse damit erzielt werden.

Die bisherige Diskussion liefert aber zugleich auch erste Argumente, um die inhaltliche Plausibilität des Maßstabes zu beurteilen. Dazu gehört die inhaltliche Kohärenz und interne Stringenz (Modellogik) der Demokratiekonzeption.<sup>10</sup> Andere Beurteilungskriterien aus methodologischer Sicht – wie Validität und Reliabilität – sollen im Hinblick auf die Operationalisierung gleichfalls beachtet werden. Dabei ist zu bedenken, dass Demokratiemessungen auf den subjektiven Urteilen von Beobachtern beruhen. Ein gewisses Maß an Objektivität kann dennoch erreicht werden, wenn das Bewertungsverfahren, seine Kriterien (Maßstab), Quellen und Aggregationsregeln transparent dokumentiert werden. Das Messverfahren gewinnt zusätzlich an Überzeugungskraft, je besser die einzelnen Schritte auch theoretisch begründet werden. Auch die Plausibilität der empirischen Befunde ist zu beachten. So kann eine Messanlage keine differenzierten Befunde liefern, wenn alle Bewertungen hinsichtlich einer größeren Gruppe identisch sind. Betrachtet werden im Folgenden diejenigen quantitativen Konzepte, die bislang in der Forschung am stärksten beachtet wurden oder aktuell neue Innovationen versprechen; angeführt werden sie unter der jeweils gängigen Bezeichnung. Nicht berücksichtigt werden konnten Ansätze, die wie der *democratic audit* der qualitativen Forschung zugerechnet werden.

9 Schiller, T.: Prinzipien und Qualifizierungskriterien von Demokratie, in: Berg-Schlosser, D./Giegel, H.J. (Hg.): Perspektiven der Demokratie, Frankfurt/Main, 1999, 28-56; Lauth, H.-J.: Demokratie und Demokratiemessung, a.a.O.; O'Donnell, G./Cullell, J.V./Iazzetta, O.M.: The quality of democracy: theory and application, Notre Dame, 2004; Diamond, L./Morlino, L., a.a.O.

10 Vgl. die Diskussion methodischer Prüfkriterien bei Munck, G.L./Verkuilen, J., a.a.O.; Lauth, H.-J.: Demokratie und Demokratiemessung, a.a.O.; Müller, T./Pickel, S., a.a.O.

### III. Quantitative Ansätze der Demokratiemessung

#### 1. Diskussion zentraler Messkonzepte

Zur quantitativen Demokratiemessung werden diejenigen Ansätze gezählt, die ihre Ergebnisse (Bewertung) vorrangig numerisch darstellen wollen. Grundlage dieser Darstellung bleiben trotzdem fast immer qualitative Urteile der beteiligten Forscher. Ausgangspunkt vieler folgender Demokratiemessungen ist das *polyarchy*-Modell von *Robert Dahl*, das bereits in dem Band selbst zu einer ersten Messung geführt hat (einmalige Messung in den 1960er Jahren).<sup>11</sup> Eckkoordinaten dieser Messung bilden die beiden Dimensionen Wettbewerb und Partizipation, die sich in ihrer Beschreibung als Ausdruck der Dimensionen von Freiheit und Gleichheit verstehen lassen. Inwieweit beide Dimensionen gegeben sind, macht *Dahl* an der Wirkungsweise von sieben weiteren Merkmalen fest, die sich maßgeblich auf einen freien und fairen Wahlkampf beziehen.<sup>12</sup>

##### a) *Vanhanen*

*Vanhanen* greift in seiner Studie von 1984, wie auch in seinen folgenden Untersuchungen 1990 und 1997, die letztlich 172 Länder umfasst, die beiden Dimensionen (Wettbewerbsgrad und Inklusion) von *Dahl* 1971 auf und versteht Demokratie im Sinne des Polyarchiemodells.<sup>13</sup> Die Eleganz und zugleich Schwäche seiner Untersuchung besteht darin, dass er für jede Dimension nur einen Indikator heranzieht, der ihm zur Bestimmung des Demokratiewertes ausreicht. Die beiden Indikatoren werden im Wesentlichen wie folgt festgelegt:<sup>14</sup>

(1) Das Ausmaß der Partizipation wird am Anteil der Wähler an der Gesamtbevölkerung gemessen (z.B. 20 Mio. Wähler bei 40 Mio. Gesamtbevölkerung = 50%); (2) der Wettbewerbsgrad wird durch einen Index erfasst, bei dem der Anteil der auf die stärkste Partei entfallenden Stimmen von 100 subtrahiert wird (z.B. die siegreiche Partei hat 43% der Stimmen erreicht, so lautet der Wettbewerbsindex 57). Beide Werte miteinander multipliziert und anschließend divi-

11 *Dahl, R.A.*: Polyarchy. Participation and Opposition, New Haven/London, 1971.

12 Ausführlicher zu *Dahl, Vanhanen, Polity* und *Freedom House* vgl. *Lauth, H.-J.*: Demokratie und Demokratiemessung, a.a.O., Kap. 3.1.

13 *Vanhanen, T.*: The Emergence of Democracy. A Comparative Study of 119 States, 1850-1979, Helsinki, 1984; *Vanhanen, T.*: The Process of Democratization. A Comparative Study of 147 States, 1980-88, New York, 1990; *Vanhanen, T.*: Prospects of Democracy. A Study of 172 Countries, London/New York, 1997.

14 *Vanhanen, T.*: Prospects of Democracy, a.a.O., 34f.

diert durch 100 ergeben den Demokratieindex (nun 28,5). Fällt dieser Wert unter eine gewisse Schwelle,<sup>15</sup> wird der Bereich der Demokratie bzw. Polyarchie verlassen. Der höchste Wert, der für eine Demokratie erreicht werden kann, liegt bei ca. 45. Auch wenn es zunächst als eine rein quantitative Messung erscheint, so gilt es nur bedingt, da auch Bewertungen über die Relevanz der Wahlen (Parlaments- und Präsidentschaftswahlen) getroffen werden müssen. Diese Bewertung ist Grundlage für die Auswahl und Gewichtung der Wahlergebnisse.

Verschiedene Probleme kennzeichnen den Ansatz von *Vanhanen*:<sup>16</sup> Zum einen benachteiligt er Zweiparteiensysteme gegenüber Mehrparteiensystemen, indem er Wahlsieger mit niedrigem Stimmenanteil prämiert. Zum anderen diskriminiert er Gesellschaften mit einem hohen Anteil junger, aber noch nicht wahlberechtigter Bevölkerung. Gleichfalls ist zu fragen, ob alle relevanten Demokratieaspekte in dieser schmalen Messanlage mit Bezug auf die lediglich zwei Dimensionen von *Dahl* berücksichtigt werden. Die Messung erlaubt jedenfalls – wie der Autor selbst betont – nur sehr vorsichtige Anmerkungen zu unterschiedlichen Graden der Demokratie.<sup>17</sup> Somit sind die Daten – wie die empirischen Befunde zeigen werden – nicht für die Einschätzung etablierter Demokratien geeignet, auch wenn sich die Messanlage als weitgehend reliabel erweist.

#### b) *Polity*

Ein anderer Ansatz, der in der komparativen Forschung auch aufgrund seines umfassenden Datensatzes besondere Aufmerksamkeit erfahren hat, ist *Polity*. Die Messungen im Rahmen des *Polity*-Projekts<sup>18</sup> haben inzwischen vier Aktualisierung und Erweiterungen des Messzeitraums erfahren, wobei die grundlegende Methodik nur unwesentlich verändert wurde.<sup>19</sup> Die bis heute reichenden jährlichen Messungen umfassen mehr als 150 historische und noch bestehende Regime seit 1800. Zur Bestimmung eines Regimes und seiner Qualität werden folgende drei Variablen auf ihre Ausprägung hin untersucht, die in vier bis sie-

15 Etwa unter den Wert „5“ bei *Vanhanen, T.*: *Prospects of Democracy*, a.a.O., 41.

16 *Schmidt, M.G.*: *Demokratiethorien. Eine Einführung*, Opladen, 1995, 274f.; *Lauth, H.-J.*: *Demokratie und Demokratiemessung*, a.a.O., 247ff.

17 Die folgenden Daten zu der Messung von 1998 verdeutlichen dies: Italien 42,75; Belgien 42,0; GB 30,16; USA 17,73. Weder die Reihenfolge, noch die Abstände können überzeugen.

18 *Gurr, T.R./Jagers, K./Moore, W.H.*: *The Transformation of the Western State. The Growth of Democracy, Autocracy, and State Power since 1800*, in: *Studies on Comparative International Development*, 25 (1990), 73-108.

19 Vgl. unter [www.bsos.umd.edu/cidcm/inscr](http://www.bsos.umd.edu/cidcm/inscr).

ben Stufen differenziert wird, wobei der eine Teil der Kategorien die demokratischen Ausprägungen und ein anderer Teil die autokratischen Merkmale eines politischen Systems erfassen. Die drei Variablen sind der „Wettbewerbsgrad der politischen Partizipation“, der „Wettbewerbsgrad und die Offenheit der politischen Rekrutierung“ sowie die „Begrenzung der Exekutive“. Die erste Variable wird durch die beiden Indikatoren „*competitiveness of political participation*“ und „*regulation of political participation*“ gemessen. Die zweite Variable wird gleichfalls in zwei Indikatoren aufgespalten („*competitiveness*“ und „*openness*“ der politischen Rekrutierung). Die dritte Variable wird mit einem gleichlautenden Indikator operationalisiert und zielt darauf ab, die Beschränkung der exekutiven Gewalt zu erfassen. Die Grundidee ist hierbei weniger die rechtsstaatliche Kontrolle der Regierung als die Erhöhung der Responsivität.

Die mit den Kategorien verbundenen regimespezifischen Punktwertungen ermöglichen die separate Messung von Demokratie und von Autokratie. Der Höchstwert des Demokratieindex (DEMOC) und des Autoritätsindex (AUTOC) beträgt jeweils zehn Punkte. Beide Indizes bilden demgemäß eine elfstufige Skala, der Intervall-Niveau unterstellt wird.<sup>20</sup> Zur Abgrenzung der Regime werden *thresholds* im Bereich der Demokratien angegeben: neben der *coherent democracy* mit 7-10 Punkten eine *mature and internally coherent democracy* bei 8 Punkten und eine *pure democracy* bei 10 Punkten.<sup>21</sup> Daneben wird explizit die Existenz von transitorischen und hybriden Systemen berücksichtigt. Solche politischen Systeme wären jene, die sich aufgrund der gewählten Grenzziehung nicht eindeutig einer Regimegruppe zuordnen lassen. Demnach lassen sich inkohärente Autokratien (AUTOC von minus sechs bis Null) von inkohärenten Demokratien (DEMOC von eins bis sechs) unterscheiden.

Auch die Messanlage von *Polity* kennzeichnen verschiedene methodische Probleme.<sup>22</sup> So ist die Trennschärfe zwischen den einzelnen Indikatoren genauso wenig durchgehend gegeben wie die Abstandsgleichheit der einzelnen Skalenstufen. Die Zuverlässigkeit und Validität der Messung besteht nur begrenzt. Auch bei diesem Ansatz werden wichtige Aspekte der Demokratie (z.B. bürgerliche

20 Gurr, T.R.: Persistence and Change in Political Systems, 1800-1971, in: American Political Science Review, 68/4 (1974), 1482-1504.

21 Gurr, T.R./Jagers, K./Moore, W.H., a.a.O., 80f.

22 Lauth, H.-J.: Demokratie und Demokratiemessung, a.a.O., 276ff.; Munck, G.L./Verkuilen, J., a.a.O.

Rechte und Rechtsstaat) nicht berücksichtigt.<sup>23</sup> Weder bilden die vorhandenen Indikatoren angemessen die selbst gewählte Minimaldefinition von Demokratie ab, noch ist die Indikatorenvalidität sichergestellt. Auch ist die Reliabilität begrenzt. Dies liegt sowohl an der zuweilen mangelnden Präzision der Skalenbeschreibung und in der fehlenden Transparenz der Quellendokumentation. So führt der Einbezug neuer Quellen in der Vergangenheit zu mehrfachen Rekodierungen der Befunde. Für die differenzierte Erfassung der Demokratiequalität in etablierten demokratischen Systemen ist die Operationalisierung daher nur begrenzt geeignet, wie zudem der empirische Befund zeigt, der wenig Varianz der Messung erkennen lässt. Darauf verweisen aber bereits die Urheber der *Polity*-Daten, indem sie feststellen: „*It is not designed to register more detailed aspects of democracy*“.<sup>24</sup>

### c) *Freedom House*

*Freedom House* (FH) misst explizit nicht die Qualität der Demokratie, sondern den Grad politischer Rechte und bürgerlicher Freiheiten anhand von zwei Checklisten zu *political rights* (PR) und *civil liberties* (CL).<sup>25</sup> Trotz bestehender Unterschiede sind die Überlappungen mit einem prozeduralen Demokratiemodell so groß, dass die Messungen oftmals ‚demokratieanalog‘ verstanden werden. Alle drei Dimensionen der Demokratie werden durch die Variablen erfasst. Allerdings wird die Kontrolldimension maßgeblich nur durch Variablen der bürgerlichen Freiheiten gemessen. Die Nähe zur Demokratie ist jedoch in der Skala zu politischen Rechten größer als in der Skala zu bürgerlichen Freiheiten, in der einige Fragen über eine Demokratiedefinition im engeren Sinne hinausgehen.<sup>26</sup> Seit Anfang der 1970er Jahre werden in jährlichen Untersuchungen nun 193 Staaten anhand von Fragen zu beiden Themen untersucht, die im Laufe der Jahre leicht variieren.<sup>27</sup> Die Fragen werden von den Beobachtern in fünf Einstufungsniveaus beantwortet. Anhand der damit ermittelbaren Punktezahl wird ein Land als *free*, *partly free* oder *not free* eingestuft. Eine Klassifikation als *electoral democracy*

23 Darauf weisen die Autoren selbst hin. Sie begründen den Verzicht jedoch mit der historischen Dimension ihrer Messung. Da speziell für die Anfangszeiten der Messung solche Daten zur Bestimmung der bürgerlichen Rechte nicht zur Verfügung stehen, wurde generell darauf verzichtet, diese einzubeziehen.

24 Gurr, T.R./Jagers, K./Moore, W.H., a.a.O., 100.

25 Gastil, R.D.: The Comparative Survey of Freedom. Experiences and Suggestions, in: Inkeles, A. (ed.): On Measuring Democracy. Its Consequences and Concomitants, New Brunswick/New Jersey, 1991, 21-36.

26 Schmidt, M.G., a.a.O., 280.

27 Vgl. unter [www.freedomhouse.org](http://www.freedomhouse.org).

ist von dieser Einstufung weitgehend frei und bezieht sich maßgeblich auf das Wahlregime.

Die Fragen zu politischen Rechten sind in drei Untergruppen unterteilt, wobei diese jeweils durch zehn weitere Leitfragen präzisiert werden, in denen insgesamt 40 Punkte erreicht werden können:

- *electoral process* (drei Fragen),
- *political pluralism and participation* (vier Fragen),
- *functioning of government* (drei Fragen).

Die bürgerlichen Freiheiten werden in vier Untergruppen unterschieden, die gleichfalls mit nun 15 Leitfragen (analog 60 Punkten) konkretisiert werden:

- *freedom of expression and belief* (vier Fragen),
- *associational and organizational rights* (drei Fragen),
- *rule of law* (vier Fragen),
- *personal autonomy and individual rights* (vier Fragen).

Die insgesamt 25 Leitfragen werden wiederum durch weitere Fragen (Indikatoren) spezifiziert, wobei die Anzahl der Indikatoren pro Leitfrage von zwei bis zehn variiert. Insgesamt liegen für die sieben Kategorien (Untergruppen) 134 Fragen vor, die für die Einschätzung des Freiheitsgrades zu bearbeiten sind.

Ein zentrales Problem des *Freedom-House-Ratings* ist die fehlende Transparenz der Einstufungen, die sich allerdings seit 2006 verbessert hat.<sup>28</sup> Das Problem verschärft sich ferner dadurch, dass die jeweilige Leitidee, die den Einstufungen zugrunde liegen muss, nur unzureichend präzisiert wird. Wenn der Maßstab nicht kommuniziert wird, dann können die einzelnen Fragen sehr unterschiedlich beantwortet werden. Die Nachvollziehbarkeit und somit die Reliabilität der Bewertungen ist daher nur sehr begrenzt sichergestellt. Zudem ist zu bedenken, dass sich nicht alle Variablen oder Fragen aufgrund ihrer Mehrdimensionalität eindeutig beantworten lassen und damit wird ein Aggregationsproblem geschaffen, das von den Autoren nicht thematisiert wird. Somit kann man das Urteil von *Munck* und *Verkuilen* nur unterstreichen: „*In the end, the aggregate data offered by Freedom House has to be accepted largely on faith*“.<sup>29</sup> Außerdem gibt es –

28 Die über viele Jahre sehr geringe Transparenz der Forschung reflektiert auch den politischen Verwendungszusammenhang der Befunde. Allerdings werden seit 2006 detailliertere Teilergebnisse auf der Ebene der sieben Untergruppen zur Verfügung gestellt, vgl. [www.freedomhouse.org](http://www.freedomhouse.org) unter „*Freedom in the World Aggregate and Subcategory Scores*“. Dies ermöglicht sowohl eine bessere Nachvollziehbarkeit der Ergebnisse als auch eine differenziertere Wahrnehmung der etablierten Demokratien.

29 *Munck, G.L./Verkuilen, J., a.a.O., 33.*

ähnlich wie bei *Polity* – wenig Möglichkeiten, die Qualität von etablierten Demokratien differenziert zu erfassen, da ein Großteil der Skala in den Bereich autokratischer Regime reicht und für die etablierten Demokratien nur wenig Differenzierung zulässt. Doch dieser Einwand lässt sich auch positiv wenden. Denn mit beiden Messanlagen kann nicht nur der Bereich der Demokratie erfasst, sondern die gesamte Bandbreite politischer Regime ausgeleuchtet werden.

d) *Economist Intelligence Unit: Democracy Index*

Der *Democracy Index* (DI) ist eines der jüngeren Instrumente, die zur Regime-messung vorgelegt wurden. Verantwortlich zeigt sich die *Economist Intelligence Unit/EIU*, die als Teil der *Economist Group* mit dem Magazin „*The Economist*“ verbunden ist.<sup>30</sup> Es existieren bislang zwei Messreihen (2006 und 2008), die 167 Länder umfassen. Das Demokratieverständnis ist weiter als bei *Freedom House* gefasst und schließt neben den politischen Rechten und bürgerlichen Freiheiten auch die faktische politische Partizipation und politische Kultur ein. Vier Regimeunterscheidungen werden getroffen: Volle Demokratien, Defizitäre Demokratien, Hybride Regime und Autoritäre Regime. Die Zuordnung der empirischen Befunde erfolgt durch eine Setzung von Schwellenwerten (8-10, 6-8, 4-6, 0-4) auf der Skala, die von 0 bis 10 reicht. Die Werte der Skala ergeben sich aus der Aggregation von 60 Indikatoren, die in fünf Kategorien geordnet sind: Wahlprozess und Pluralismus, bürgerliche Freiheiten, effiziente Regierung, politische Partizipation und politische Kultur. Die Indikatoren sind entweder dichotom (1, 0) oder nach einer Dreierskala (1, 0.5, 0) ausgerichtet. Für vier Indikatoren gibt es Sonderregeln, die zu Abwertungen in zwei Kategorien führen können.

Die Datenerhebung, an denen rund 500 Personen beteiligt sind, erfolgt durch Experten-*ratings* und durch empirische Umfragedaten (maßgeblich *World Value Survey* (WVS) plus weitere regionale und nationale *surveys*). Einige spezifische Bewertungen sind in diesem Messinstrument enthalten. So trägt eine steigende Wahlbeteiligung (Frage 27), die parteiliche und zivilgesellschaftliche Organisationsquote (Frage 31) und ein höherer Anteil von Frauen im Parlament (Frage 29) ebenso zu höheren *ratings* bei wie eine Vormacht des Parlaments gegenüber der Exekutiven (Frage 14). Gleichfalls werden Mehrparteiensysteme höher bewertet als fest etablierte Zweiparteiensysteme (Frage 10) oder reduziert Regierungseinfluss in „*private business*“ die Demokratiequalität.

30 Vgl. unter [www.eiu.com/public](http://www.eiu.com/public).

Die Definition der Demokratie ist nicht sehr präzise, da sie maßgeblich bereits durch die genannten Komponenten (Kategorien) bestimmt wird, die dann anhand der Fragen (Indikatoren) präzisiert werden. Überwiegend (nicht Frage 10) messen die Indikatoren die jeweiligen Kategorien valide, auch wenn die Antwortalternativen zum Teil kaum erläutert werden und nicht jegliche Grenzziehung für die Zuordnung zum Indikatorenniveau plausibel ist; dies gilt beispielsweise für die stark divergierenden Schwellenangaben bei den politischen Kulturdaten (Fragen 37-42). Gleichfalls finden sich Fragen, die zwar für die Demokratie relevant sind, aber nur in einer Richtung gewertet werden sollten. So reduziert ein Mangel an *stateness* (Frage 20) sicherlich die Qualität der Demokratie, doch eine umfassend gegebene Staatlichkeit ist eine notwendige Bedingung für eine Demokratie und kein Merkmal, wie ein Blick auf stabile Autokratien leicht verdeutlichen kann. Zu kritisieren ist auch das Demokratieverständnis. Dies betrifft zum einen das Insistieren auf hohen Partizipationsraten, ohne die Qualität der politischen Beteiligung näher zu betrachten. Zum anderen wirft der Einbezug der politischen Kulturdaten Fragen auf, die sich auf *items* beziehen wie z.B. Vertrauen in Regierung (Frage 25), in politische Parteien (Frage 26) oder der Glaube, dass Demokratie besser für wirtschaftliche Entwicklung ist (Frage 41) oder das bessere politische System (Frage 42). Etliche der vom WVS übernommenen *items* werden in der Forschung für die Messung der Konsolidierung bzw. Stabilität und kaum für die Qualität einer Demokratie verwendet. So sollte beispielsweise ein geringes Vertrauen in eine Regierung nicht als Ausdruck einer geringen Demokratiequalität verwendet werden, da solche Einschätzungen oftmals mit Performanzleistungen verknüpft sind.

Die empirischen Befunde des DI zeigen deutliche Unterschiede zu den Daten von FH und *Polity*, die bei den etablierten Demokratien wenig bis sehr wenig variieren. So erfasst der DI zumindest die Reduktion der bürgerlichen Freiheiten, die in den USA und in GB im Kampf gegenüber dem Terrorismus festzustellen ist. Allerdings bleiben beide im Status voller Demokratien. Überraschend ist, dass die Türkei schlechter als Singapur bewertet wird und sich in der Kategorie hybrider Regime befindet. Negativ zu kommentieren ist die fehlende Transparenz bei der Bewertung. Vorgestellt werden nur die bereits aggregierten Daten auf Niveau der Kategorien. Es bleibt daher unklar, welche Ausprägungen die einzelnen Indikatoren aufweisen. Gleichfalls finden sich in den nicht-kommerziellen Ausführungen zum DI keine Angaben über die Experten und das konkrete Vorgehen bei der Datenerhebung.

e) *Weltbank: Voice and Accountability*

Eine weitere Möglichkeit der Demokratiemessung bieten die Daten zu *Voice and Accountability* der *governance*-Indikatoren der Weltbank. Das zentrale Anliegen der Weltbank ist die Messung von *governance*, um auf dieser Grundlage bestimmen zu können, inwieweit *good governance* in den untersuchten Ländern realisiert ist. Zur Operationalisierung des Konzepts wird *governance* in sechs Dimensionen unterteilt, die zusammen alle wesentlichen Bereiche erfassen sollen.<sup>31</sup> Dies sind: (1) Mitspracherecht/Partizipation und Rechenschaftspflicht; (2) politische Instabilität und Gewalt; (3) effektive Regierung; (4) Regelungsqualität; (5) Rechtsstaat, (6) Kontrolle der Korruption. Erfasst werden inzwischen 212 Länder und Territorien seit 1996 – zunächst im zweijährigen Rhythmus, ab 2002 dann jährlich.<sup>32</sup> Die Angaben basieren auf 37 verschiedenen Datenquellen, die von 31 Organisationen (z.B. BTI, FH, EIU) erhoben werden. Zwei Datentypen werden zur Beurteilung der verschiedenen Aspekte von *governance* genutzt: Experten-*ratings*, für die Personen aus Wirtschaft, Medien und Zivilgesellschaft befragt wurden, und repräsentative Umfragen. Vier Quellentypen werden unterschieden, die unterschiedliche Anteile an den Gesamtdaten haben (Angaben für 2008): Kommerzielle Geschäftsdienste (42%), Zensuserhebungen von Firmen und Haushalten (18%), NGO-Berichte (18%) und offizielle (staatliche) Daten und Angaben (22%).

Die Datenbasis für die einzelnen Dimensionen und Ländergruppen variiert deutlich; beispielsweise stehen für die Industrieländer in der Regel mehr Angaben zur Verfügung als für wenig entwickelte Länder. Um fehlende Daten zu kompensieren, arbeiten die Autoren (*Kaufmann/Kraay/Zoido-Lobaton*) mit der *unobserved components method*, einer Variante der Faktoranalyse.<sup>33</sup> Für die Messung

31 Kaufmann, D./Kraay, A./Zoido-Lobaton, P.: Aggregating governance indicators, Policy Research Working Paper Series 2195, Washington DC, 1999.

32 Die methodischen Grundlagen für die Erhebung der Governance-Indikatoren wurden Ende des letzten Jahrzehnts gelegt und haben sich bis heute nicht geändert (*Kaufmann, D./Kraay, A./Zoido-Lobaton, P.: Aggregating governance indicators, a.a.O.; Kaufmann, D./Kraay, A./Zoido-Lobaton, P.: Governance Matters, World Bank Policy Research Working Paper No. 2196, Washington DC, 1999*); allerdings haben sich die Daten geringfügig geändert, da in der letzten Version zusätzliche Quellen aufgenommen wurden (*Kaufmann, D./Kraay, A./Mastruzzi, M.: Governance Matters III: Governance Indicators for 1996, 1998, 2000, and 2002, in: World Bank Economic Review, 18 (2004), 253-287; analog Kaufmann, D./Kraay, A./Mastruzzi, M.: Governance Matters VIII: Aggregate and Individual Governance Indicators, 1996-2008, in: World Bank Policy Research Working Paper No. 4978, 2009.*

33 "which expresses the observed data in each cluster as a linear function of the unobserved common component of governance, plus a disturbance term capturing perception errors and/or sampling variation in each indicator" (*Kaufmann, D./Kraay, A./Zoido-Lobaton, P.: Governance Matters, a.a.O., 9*).

einer einzigen Dimension werden die jeweiligen Ergebnisse der einzelnen Datenquellen additiv zusammengefasst, wobei sie je nach Brauchbarkeit unterschiedlich gewichtet werden.<sup>34</sup> In diesem Sinne bilden die Weltbankangaben einen Superindex, der auf einer Fülle unterschiedlicher Einzelmessungen basiert. Die Messwerte verlaufen von  $-2,5$  bis  $2,5$ , wobei letzterer die optimale *governance*-Ausprägung anzeigt.

In der kritischen Einschätzung dieser Datensätze lässt sich die Entscheidung über die Auswahl der einzelnen Quellen und deren Gewichtung diskutieren. Weiterhin ist zu beachten, dass der Maßstab, der den einzelnen Messungen zugrunde liegt, nicht abgestimmt werden kann und die Dimensionen auch aufgrund überlappender Indikatoren nicht trennscharf gemessen werden. *Landman*<sup>35</sup> unterstreicht zudem die „Gefahr“, dass sich die Bewertung der *governance*-Dimensionen – vor allem bei „effektiver Regierung“ und „Regelungsqualität“ – stärker an der allgemeinen Wirtschaftslage als an den gemeinten Sachverhalten orientiert. Unabhängig von der geleisteten Kritik erscheinen die vorliegenden Daten als brauchbar, weil (1) bei der Auswahl der Quellen in der Regel die wichtigsten existenten Datensätze berücksichtigt werden, (2) die Eindimensionalität der Indizes für die sechs Dimensionen im Allgemeinen gewährleistet ist und (3) die Autoren auf die Ungenauigkeiten der Messung selbst hinweisen und versuchen, diese einzuschätzen.<sup>36</sup> Um die Ungenauigkeit nachvollziehbar zu machen, präsentieren sie *margins of error* für die Messungen der Dimensionen für jedes Land. Auch wenn der durchschnittliche *standard error* im Laufe der Jahre sinkt (von  $0,25$  im Jahr 1996 auf  $0,17$  im Jahr 2008), bleibt die Fehlerbreite so groß, so dass einzelne Länderergebnisse nur schwer direkt verglichen werden können. Erst deutliche Abstände, bei denen sich die Fehlerbandbreiten nicht mehr überlappen, lassen sich sinnvoll als Unterschiede interpretieren.

Im Anschluss an diese Überlegungen zu den *governance*-Indikatoren lässt sich die Verwendung der *governance*-Dimension „*voice and accountability*“ (V&A) als Demokratieindikator diskutieren, da dort etliche Rechte und Freiheiten, die für die Demokratie relevant sind, erfasst werden. Hierzu gehören maßgeblich kommunikative Rechte („*voice and accountability* (V&A) – *capturing percep-*

34 Vgl. hierzu die Erläuterungen bei *Thiery, P./Sehring, J./Muno, W.*: Die Messung von Rechtsstaatlichkeit, in: *Estermann, J.* (Hg.): *Interdisziplinäre Rechtsforschung zwischen Rechtswirklichkeit, Rechtsanalyse und Rechtsgestaltung*, Bern, 2009, 211-230.

35 *Landman, T.*, op. cit, 32 f.

36 *Kaufmann, D./Kraay, A./Zoido-Lobaton, P.*: *Aggregating governance indicators*, a.a.O., 11.

tions of the extent to which a country's citizens are able to participate in selecting their government, as well as freedom of expression, freedom of association, and a free media.”).<sup>37</sup> Die Daten von V&A stammen maßgeblich von der *Economist Intelligence Unit* (hier neben dem besprochenen DI auch andere, nur kommerziell zugängliche Angaben zu Rechten und *accountability*), dem *BTI* (hier die Daten zu Politischer Partizipation, Stabilität der demokratischen Institutionen und politischer und sozialer Integration) und vor allem von *Freedom House* (fast 35% der Gesamtgewichtung). Bei FH werden die Daten zu PR und CL, zu Pressefreiheit und *Nation in Transition* zusammengefasst und zusätzlich die Angaben aus den Berichten zu *Countries on the Crossroads* aufgegriffen.

Ein zentrales Problem dieser Messanlage ist, dass keine Definition der Demokratie vorliegt und das Verständnis von *Voice and Accountability* auch nicht näher erläutert wird. Die Präzisierung dieser Dimension erfolgt letztlich durch die Auswahl der Datenquellen. Hierbei werden im Wesentlichen anerkannte Quellen verwendet, die aber – gerade mit Bezug zu *Freedom House* – mehr Inhalte transportieren, als die Dimension zum Ausdruck bringen möchte. So ist die Validität der Messung begrenzt. Zudem beruhen andere Daten auf Experteneinschätzungen, die eher wirtschaftliche denn politische Kompetenz haben und die im Unterschied zu *Polity* und *Freedom House* noch weniger repräsentativ sind. Die Reliabilität ist aufgrund der transparenten Datenpräsentation der Weltbank weitgehend gegeben, wobei sich dieses Urteil nicht durchgängig auf die Primärdaten übertragen lässt. Schwierig gestaltet sich aus den genannten Validitätsproblemen die Interpretation der Datenschatzung. Diese ist zudem problematisch, weil die Zeiträume der Beobachtungen nicht bei allen Quellen gleich sind. Gleichfalls fehlen Angaben zu Schwellenwerten, die eine Demokratie von einer Autokratie unterscheiden lassen.

#### f) *BTI*

Eine weitere Möglichkeit, quantitative Angaben zur Bestimmung der Qualität einer Demokratie zu verwenden, bietet der *Bertelsmann-Transformations-Index* (BTI), dessen Daten jedoch auf ‚Transformationsländer‘ begrenzt sind.<sup>38</sup> Insgesamt werden inzwischen 128 Transformationsländer erfasst, darunter alle Staa-

37 Kaufmann, D./Kraay, A./Mastruzzi, M.: *Governance Matters VIII*, a.a.O., 8.

38 Vgl. *Bertelsmann Stiftung* (Hg.): *Codebuch. Bertelsmann Transformation Index 2003. Auf dem Weg zur marktwirtschaftlichen Demokratie*, Gütersloh, 2004; *dies.* (Hg.): *Bertelsmann Transformation Index 2006. Auf dem Weg zur marktwirtschaftlichen Demokratie*, Gütersloh, 2005; zur Kommentierung vgl. Müller, T./Pickel, S., a.a.O.

ten, die im Zuge der Osterweiterung EU-Mitglieder geworden sind. Die theoretischen Grundlagen des BTI beruhen maßgeblich auf dem Konzept der *embedded democracy*.<sup>39</sup> Für die Messung sind folgende fünf Teilregime für die Beurteilung einer Demokratie beziehungsweise der Ergebnisse der politischen Transformation relevant: Staatlichkeit, politische Partizipation, Rechtsstaatlichkeit, institutionelle Stabilität sowie politische und gesellschaftliche Integration.

Die fünf Teilregime werden durch weitere Fragen präzisiert. Die empirischen Befunde werden auf einer Skala von eins bis zehn (höchster Wert für eine Demokratie) eingeordnet. Um die Zuordnung zu erleichtern und die Reliabilität zu erhöhen, werden in einem detaillierten Codebuch die Skalen auf vier Bewertungsstufen (jeweils für Skalenstufe 1-2, 3-5, 6-8 und 9-10) näher erläutert. Die insgesamt 18 Fragen werden zunächst von zwei Länderexperten bearbeitet und anschließend im intraregionalen und interregionalen Vergleich abgestimmt. Die Indexbildung erfolgt durch die Mittelwertbildung der Bewertungen der fünf Teilregime, die ihrerseits auf den Mittelwerten der einzelnen Fragen basiert. Eine Demokratie liegt vor, wenn ein Land bei jeder der Fragen ein bestimmtes Bewertungsniveau erreicht hat. Das Mindestniveau liegt mit Ausnahme von Frage 2.1 zu allgemeinen Wahlen, in der sechs Punkte erreicht werden müssen, stets bei drei Punkten. Unter diesem Niveau wird es als Autokratie klassifiziert.

Mit diesem Vorgehen ist bereits ein erstes Problem angesprochen. Für die Reliabilität einer Klassifikation ist es nicht ganz überzeugend, dass der Schwellenwert innerhalb einer Bewertungsstufe liegt, die die Beschreibung der Skalenwerte 3-5 umfasst. Klarere Grenzen wären zwischen den vier Bewertungsstufen gegeben. Ein weiteres Problem ergibt sich durch die Skalierung. Zunächst ist eine Zehnerskalierung generell weniger reliabel als eine Skala mit weniger Stufen. Doch, so ließe sich dagegen halten, steige damit die Validität der Messung. Aber dies kann nur begrenzt gelten, da die Beschreibungen der vier Bewertungsstufen nicht immer die gleichen Kriterien deklinieren. Im Unterschied zu *Freedom House*, das generell auf die Formulierung eines Maßstabes verzichtet, werden nun in der Beschreibung von vier Skalenniveaus ähnlich wie bei *Polity* mehrere Zielkorridore beschrieben, die einer stringenten Auffächerung der einzelnen Stufe nicht immer förderlich sind.

Aber auch nicht alle Teilregime können als Indikator für die Qualität einer Demokratie überzeugen. So misst Frage 1.1 zu Staatlichkeit den Umfang des Ge-

39 Merkel, W./Puhle, H.-J./Croissant, A./Eicher, C./Thiery, P.: Defekte Demokratien, Bd. 1: Theorie, Wiesbaden, 2003.

waltmonopols. Nun sinkt mit diesem durchaus die Qualität einer Demokratie, aber diese steigt nicht notwendigerweise mit einer Verbesserung des staatlichen Gewaltmonopols, wie bereits in der Diskussion der DI/EIU verdeutlicht wurde. Das Gewaltmonopol ist eine notwendige Bedingung für, aber kein Merkmal der Demokratie. Interessant ist auch, dass Staatlichkeit im Zielkorridor als säkulare verstanden wird (1.3.). Doch dies erschließt durchaus den demokratischen Gehalt eines Staates. Problematischer ist dagegen der Einbezug des fünften Teilregimes, der stark die Stabilitätsaspekte betont und in vielen Teilelementen (Fragmentierung des Parteiensystems, Zustimmung zur Demokratie (Umfragedaten) und Ausprägung des Sozialkapitals) Kriterien aufgreift, die bei der Konsolidierung der Demokratie herangezogen werden.<sup>40</sup> Schließlich ist die Frage von 3.1 und 3.2 sehr ähnlich ausgerichtet; beides zielt auf die Unabhängigkeit der Justiz. Auch hinsichtlich der Fragen 1.2, 3.4 und 4.2 gibt es im Fokus auf die Anerkennung bürgerlicher Rechte und Freiheiten durchaus Überlappungen. Sehr positiv ist die transparente Dokumentation der Länderberichte und der Skalierung zu beurteilen, die eine detaillierte Auseinandersetzung mit den jeweiligen Skalenstufen ermöglichen. Auch das Demokratiekonzept ist in seiner konzeptionellen Grundlegung und theoretischen Erläuterung plausibel. Weniger überzeugend ist es hingegen, die Demokratiemessung zusammen mit den Befunden zur Marktwirtschaft in einem Statusindex zu verbinden.

g) *Sustainable Governance Indikatoren/Democracy (SGI)*

Da sich der BTI explizit nur auf die Transformationsstaaten bezieht und die Differenziertheit der Messung auf diese Ländergruppe ausrichtet, ist er für die Messung etablierter Demokratien weniger geeignet. Daher wurde durch die Bertelsmann Stiftung für die Untersuchung dieser Ländergruppe (30 OECD-Länder) mit den *Sustainable Governance Indicators* (SGI) eine veränderte Messanlage entwickelt. Bei den SGI wird die Qualität der Demokratie als Teil des Status-Index erfasst, der zudem das politische und ökonomische Leistungsniveau beurteilt. Die Forschung ist somit eng verbunden mit der Einschätzung der Leistungs- und Reformfähigkeit von Staaten.<sup>41</sup> Vier Kategorien werden in der Demokratiemessung aufgegriffen: Wahlprozess, freie Medien, bürgerlicher Rechte und Rechtsstaat. Diese Kategorien werden durch insgesamt elf Indikatoren erschlos-

40 Merkel, W.: Systemtransformationen, Wiesbaden, 2010, 110-127.

41 Zu einer umfassenden Diskussion des Gesamtkonzepts vgl. Jäckle, S./Bauschke, R.: Lässt sich Reformfähigkeit messen? Eine kritische Würdigung der *Sustainable Governance Indicators*, in: ZPol, 19/3 (2009), 359-386.

sen, die auf einer Zehnerskala bewertet werden.<sup>42</sup> Hierbei werden inhaltliche Bezugspunkte zu verschiedenen Punkten (1-2, 3-5, 6-8, 9-10) gegeben, um die Transparenz und Reliabilität der Zuordnung zu erhöhen. Die Bewertung erfolgt wiederum durch Länderexperten, wobei stets einheimische und externe Beurteiler einbezogen werden. Die Anpassung der Beurteilung verläuft in einem dem BTI-analogem Verfahren. Die Ergebnisse werden darüber hinaus durch eine lineare Transformation standardisiert, um die Unterschiede zwischen den Befunden zu akzentuieren.<sup>43</sup> Dabei werden die jeweiligen niedrigsten und höchsten Werte der empirischen Befunde den Skalenwerten ‚eins‘ beziehungsweise ‚zehn‘ zugeordnet. Auf dieser numerischen Basis erfolgt dann die weitere Aggregation der Daten.

Das Demokratiekonzept ist deutlich schmäler als beim BTI aufgestellt. Nicht mehr einbezogen werden die Faktoren, die für die Konsolidierung einer Demokratie relevant sind (wie institutionelle Stabilität sowie politische und gesellschaftliche Integration); auch das Kriterium der Staatlichkeit fehlt. Letztlich entspricht es dem Konzept von *Dahl* plus Rechtsstaat. Verzichtet wird auf die Festlegung von Regimegrenzen. Es wird anscheinend davon ausgegangen, dass alle OECD-Demokratien in der Tat Demokratien sind; selbst auf eine Bestimmung eines Schwellenwerts einer defizitären Demokratie wird verzichtet. Dies ist nicht unproblematisch, da Länder wie Mexiko und Türkei betrachtet werden. Problematisch ist gleichfalls die Transformation der Befunde in eine Zehnerskala, die eine künstliche Vergrößerung der Differenzierung bewirkt. Denn dabei können sehr unterschiedliche Ausgangspunkte (also deutlich unterschiedliche Skalenwerte) gleich gesetzt werden, die in der weiteren Aggregation dann als gleich interpretiert werden. Zudem müssen auch die niedrigsten Werte – also auch Werte die deutlich unter der korrespondierenden BTI-Messung liegen – als voll demokratisch interpretiert werden. Auch unterschiedlich große Abstände der ersten Messung werden durch die Transformation gleich geschaltet. Die ursprünglichen Skalenbefunde werden auf diese Weise nicht mehr sichtbar. Daher erscheint die Transformation nicht „*intuitively plausible*“, wie die Autoren meinen.<sup>44</sup> Letztlich macht die Darstellung der Befunde die Ergebnisse mit den bislang vorgestellten Messungen nur schwer kompatibel.

42 Vgl. unter <http://www.sgi-network.org>.

43 Vgl. *Brusis, M.*: Designing Sustainable Governance Indicators – Assessment Criteria and Methodology, in: *Stiftung, Bertelsmann* (Hrsg.), *Sustainable Governance Indicators*. Gütersloh 2009, 71-100, hier 87 f.

44 *Brusis, M.*, a.a.O., 87.

Auch die Reliabilität lässt zu wünschen übrig. Auch wenn der Standardfehler (zwischen 0.471 und 0.88) im Bereich der Demokratiemessung<sup>45</sup> akzeptabel ist, gilt dies nicht hinsichtlich der Abweichungen bei etlichen Einzelmessungen; die Aggregation der Werte der Einzelindikatoren ist gleichfalls nicht plausibel. Betrachten wir nur den Indikator ‚*free election process*‘ hinsichtlich USA und Türkei (SGI 2009) anhand der angeführten Daten: Zunächst überrascht die hohe Diskrepanz der Bewertung durch die Gutachter (Türkei 9/8/4; USA 5/9/10) von jeweils fünf (!) Skalenpunkten. Als zusammenfassende Länderwerte werden dann sechs (Türkei) und neun (USA) für diesen Indikator angeboten. Für die beiden weiteren Indikatoren des Wahlprozesses (*fair electoral campaign, inclusive electoral process*) lauten die Werte fünf und acht (Türkei) sowie sieben und acht (USA). In der abschließenden *score*-Tabelle, die nun die transformierten Werte anzeigt, erhält die Türkei den Wert 1 und die USA den Wert 4,5. Die Gesamtwerte für beide Länder betragen für die Türkei 1,5 und für die USA 7.3. Die Skalierungsangaben werden in diesem aggregierten Wert nicht mehr sichtbar.

Generell ist unverständlich, dass für die Messung von verschiedenen Ländergruppen (Transformationsstaaten und OECD-Länder) verschiedene Demokratiekonzepte und divergente Indikatoren verwendet werden, zumal einige Länder beiden Gruppen angehören. Hier ist die Anlage von *Polity* und FH plausibler, die die Idee eines Regimekontinuums, die an sich auch der Konzeption der *embedded democracy* zugrunde liegt, stringenter umsetzen. Hier bleiben Konzept und Kriterien konstant, nur die Skalierung wird dem Untersuchungsbereich angepasst. Dieser Weg wäre auch für BTI und SGI sinnvoller gewesen.

#### *h) NCCR-Demokratiebarometer (DB)*

Auch der Demokratiebarometer steht in einem ähnlichen theoretischen Kontext wie BTI und SGI, die ihren Ursprung im Konzept der *embedded democracy* haben. Er ist am *National Center of Competence in Research* (NCCR) im Modul 5 (Qualität der Demokratie) angesiedelt. Dies ist ein interdisziplinäres Forschungsprogramm, das vom Schweizer Nationalfonds seit 2005 gefördert wird. Untersucht wird neben der Qualität einer Demokratie auch die Wirkungen unterschiedlicher Demokratiegrade sowie Möglichkeiten, die Qualität zu verbessern.

45 Ebd., 85

Die Untersuchungen möchten 75 etablierte Demokratien möglichst differenziert erfassen.<sup>46</sup>

Wie BTI und SGI erstreckt sich das Demokratiekonzept auf drei Dimensionen: Freiheit, Gleichheit und Kontrolle. Diese Dimensionen werden durch jeweils drei Funktionen präzisiert, wobei diese in den einzelnen Entwicklungsstadien des Projekts leicht variieren. In der letzten Formulierung<sup>47</sup> werden der Freiheit die Funktionen der individuellen Freiheiten, des Rechtsstaats und der Öffentlichkeit zugeordnet; der Kontrolle nun Wettbewerb, Gewaltenkontrolle, Regierungsfähigkeit; der Gleichheit die Funktionen der Transparenz, Partizipation und Repräsentation. Alle Funktionen werden durch weitere Komponenten und anschließend Indikatoren operationalisiert. Insgesamt werden mehr als 300 Indikatoren gesichtet und schließlich nach methodischer Prüfung 95 Indikatoren verwendet, die von bereits bestehenden Datensätzen entnommen werden. Der Demokratiebarometer bildet somit analog der *governance*-Indikatoren der Weltbank einen Meta-Index und verzichtet auf eine Primärdatenerhebung. Im Unterschied zum BTI versucht der DB möglichst auf Daten zurückzugreifen, die nicht aufgrund von Experten-*ratings* entstanden sind. Um die Daten kombinieren und aggregieren zu können, werden diese jeweils auf eine Skala von Null bis 100 (höchster Demokratiewert) transformiert. Mit Ausnahme der untersten Ebene, in der die Aggregation mittels Faktorenanalyse vorgenommen wird, erfolgt die Datenbestimmung durch die Berechnung der Mittelwerte.<sup>48</sup> In der aktuellsten Version wurden diese Verfahren wiederum geändert. Nun werden die Aggregationen auf den beiden untersten Stufen durch Mittelwerte ermittelt. „*In the following steps (components to functions, functions to principles, principles to ‘Quality of Democracy’ [QoD]), the idea of the optimal balance is implemented: the value of the higher level is calculated using a formula that rewards high values at the lower level but penalizes incongruence between pairs of values*“.<sup>49</sup>

46 Vgl. Bühlmann, M./Merkel, W./Wessels, B.: The Quality of Democracy: Democracy Barometer for Established Democracies, National Centre of Competence in Research (NCCR) *Challenges to Democracy in the 21st Century* Working Paper No. 10a, Zürich, 2008; Bühlmann, M.: The Beauty and the Beast? A Tale of Democratic Crises and Globalization, National Centre of Competence in Research (NCCR) *Challenges to Democracy in the 21st Century* Working Paper No. 46. Zürich, 2010; Merkel, W./Giebler, H.: Measuring Social Justice in the OECD, in: Bertelsmann Stiftung (Hg.): Sustainable Governance Indicators 2009. Policy Performance and Executive Capacity in the OECD, Gütersloh, 2009, 187–215.; vgl. auch <http://www.nccr-democracy.uzh.ch>.

47 Bühlmann, M., a.a.O.

48 Merkel, W./Giebler, H., a.a.O.

49 Bühlmann, M., a.a.O., 13.

In einer ersten Studie wurden 30 etablierte Demokratien gemessen und die Werte über einen längeren Zeitraum (1995-2005) zusammengefasst, wobei aber auch jährliche Messungen verfügbar sind. Die Werte verdeutlichen starke Unterschiede in der Demokratiequalität in der Gruppe etablierter Demokratien und zugleich unterstreichen sie, dass auch die Ausprägung der drei Dimensionen deutliche Akzentuierungen aufweisen kann. Überraschend sind nicht die hohen Werte für die skandinavischen Demokratie, die zwischen 77 und 84 Punkten liegen, gleichfalls nicht die niedrigeren Werte für südeuropäische Demokratien (Portugal 52, Spanien 55), sondern die niedrigen Werte von zwei anderen etablierten Demokratien: Frankreich 44,9 und GB 47,7. Beide liegen unter den Werten der Tschechischen Republik (48,2) und Ungarn (49).

Die Dimensionen und auch die angeführten Funktionen sind für die Entfaltung eines dreidimensionalen Demokratiekonzepts plausibel. Dies gilt jedoch nicht für die interne Modelllogik. So ist die Zuordnung der einzelnen Funktionen zu den einzelnen Dimensionen weder ausreichend begründet noch überzeugend. Warum sollte gerade Transparenz der Gleichheit zugeordnet werden und nicht der Kontrolle oder der Freiheit? Oder warum wird der Wettbewerb unter Kontrolle und nicht unter Gleichheit geführt oder Rechtsstaat nicht unter Kontrolle oder Gleichheit? Diese Fragen ließen sich problemlos fortsetzen und möchten nur ein Argument unterstreichen. Es ist nicht überzeugend, die einzelnen Funktionen nur einer Dimension zuzurechnen, wenn sie gleichfalls mit den anderen verbunden ist. Dies wird auch im Konzept deutlich, wenn beispielsweise auf der Ebene der Komponenten der Indikator ‚Gleichheit vor dem Gesetz‘ bei der Funktion Rechtsstaat und somit bei der Freiheitsdimension auftaucht. Eine differenzierte Diskussion der methodischen Grundlagen ist zurzeit nicht möglich, da weder die Indikatoren noch die verwendeten Quellen dokumentiert werden.<sup>50</sup> Wenn die Daten öffentlich zugänglich sein sollten und die Aggregationsregeln dargelegt werden, sollte die Messung reliabel sein. Die Validität kann dagegen zurzeit nicht eingeschätzt werden. Problematisch erscheint jedoch, dass auf *Expertennratings* weitgehend verzichtet wird. Auch das aktuelle Aggregationsverfahren, das Messunterschiede ‚künstlich‘ vergrößert, wäre genauer zu prüfen und erscheint ähnlich problematisch wie dasjenige von den SGI.

50 Eine Liste früherer Vorschläge (*Bühlmann, M./Merkel, W./Weßels, B., a.a.O.*) wird nicht diskutiert, da nicht klar ist, inwieweit diese noch aktuell sind. Dort finden sich jedoch Indikatoren (wie Anteil der Exporte und Importe am BSP), die wenig überzeugen.

i) *Neuer Index der Demokratie (NID)*

Diese Messanlage versucht mittels der gezielten Kombination von vorhandenen Datenreihen deren Defizite zu kompensieren.<sup>51</sup> *Freedom House/Political Rights* und *Polity* konzentrieren sich weitgehend auf die Dimensionen der Freiheit und der Gleichheit. Um die Inhaltsvalidität dieser Messungen auf der Basis einer dreidimensionalen Demokratiekonzeption zu verbessern, müssten diese durch die Kontrolldimension ergänzt und damit Aspekte der horizontalen *accountability* und der Rechtsstaatlichkeit aufgenommen werden. Bislang fehlt jedoch gleichfalls ein umfassender Datenbestand zur politischen und rechtlichen Kontrolle. Daher gilt es funktionale Äquivalente zu finden. Ein solches Äquivalent bietet der Index „*rule of law*“, der zu den *governance*-Indikatoren der Weltbank gehört.<sup>52</sup> Eine Kombination mit den beiden erstgenannten Demokratiemessungen ermöglicht es, alle drei Dimensionen von Demokratie zu berücksichtigen. Diese Möglichkeit wird im Folgenden mit der Bildung zweier Indizes (NID3D und NID) aufgegriffen.<sup>53</sup>

Der neue dreidimensionale Index der Demokratie (NID3D) setzt sich aus den Befunden von *Polity*, *Freedom House/PR* und der Weltbank (*governance indicators: rule of law*) zusammen (NID3D = 3. Wurzel des Produktes der drei Komponenten). Das Aggregationsverfahren der Multiplikation wurde gewählt, um auszuschließen, dass schwache Werte auf einer Dimension durch hohe Werte auf einer anderen Dimension kompensiert werden können, wie das bei einem additi-

51 Vgl. ähnlich *Berg-Schlosser, D.*: The Quality of Democracies in Europe as Measured by current Indicators of Democratization and Good Governance, in: *Journal of Communist Studies and Transition Politics*, 20/1 (2004), 28-55. Er verwendet verschiedene Ansätze der Messung von Demokratie und Menschenrechten (*Vanhanen, Polity, Freedom House/Political Rights, Worldbank Governance Indicators: Voice and Accountability* sowie *Rule of Law, Gross Human Rights Violations*), um demokratische Regime anhand des Einbezugs von Schwellenwerten in verschiedene Subtypen zu klassifizieren. Er bildet auf diese Weise keinen neuen Index, sondern erzielt seine Klassifikationen durch die Zuordnung der vorliegenden Befunde in ein Klassifikationsschema. Er kann damit die Mitgliedstaaten der EU begrenzt differenziert bewerten, indem er Werte aus 1996 und 2003 kombiniert. Außer Bulgarien und Rumänien, die unter der Kategorie B als Demokratien mit gewissen Mängeln klassifiziert werden, werden alle anderen Länder der höchsten Kategorie A zugeordnet. Diese wird nochmals unterschieden (A, A–), wobei A als volle und konsolidierte Demokratien bezeichnet wird. Kleinere Abstriche (A–) weisen die mitteleuropäischen Staaten sowie Griechenland, Italien und Zypern auf.

52 *Kaufmann, D./Kraay, A./Mastruzzi, M.*: Governance Matters III. Governance Indicators for 1996-2002, Washington DC, 2003.

53 *Lauth, H.-J.*: Die Qualität der Demokratie. Der NID als pragmatischer Vorschlag für die komparative Forschung, in: Schnapp, K.-U./Behnke, N./Behnke, J. (Hg.): *Datenwelten. Datenerhebung und Datenbestände in der Politikwissenschaft*, Baden-Baden, 2008, 373-390. Weitere Informationen und die Datensätze finden sich auf [www.politikwissenschaft.uni-wuerzburg.de](http://www.politikwissenschaft.uni-wuerzburg.de) unter „Neuer Index der Demokratie“.

ven Indexbildungsverfahren möglich wäre. Da jedes Aggregationsverfahren voraussetzt, dass die Basisinformationen auf einer gleichen Skala verfügbar sind, ist eine Transformation der Daten notwendig. Bezugspunkt für die Bildung der Teilindizes ist die Elf-Punkte-Skala von *Polity*. In dieser Skala gibt der Wert ‚10‘ die höchste Demokratiequalität an. Die Daten von *Freedom House/PR* und von der Weltbank (*governance*-Indikatoren: *rule of law*) werden durch eine z-Transformation entsprechend umgeformt. Hierzu werden die Werte von FH zunächst umgepolt und bei den *governance*-Indikatoren die jeweils faktisch höchsten und tiefsten Werte als Endpunkte (0, 10) der Skala gewählt. Die Werte der Variablen aus den beiden transformierten Datensätzen liegen zwischen null und zehn.

Neben der Kontrolldimension ist bei einer globalen Perspektive der Faktor „Staatlichkeit“ zusätzlich in den Demokratieindex einzubeziehen. Bei *Freedom House* und *Polity* sowie bei den *governance*-Indikatoren zu *rule of law* ist dies aufgrund der Indikatoren nur sehr begrenzt gegeben. Auch in diesem Fall liefert die *governance*-Datenreihe der Weltbank einen Indikator. Für ‚Staatlichkeit‘ kann der Indikator „*political stability*“ verwendet werden, der den Aspekt eines funktionsfähigen Gewaltmonopols aufgreift. Da das Fehlen von Staatlichkeit alle Dimensionen der Demokratie tangiert, wird auch der neue Index der Demokratie (NID) durch Multiplikation kombiniert und auf der Grundlage des bereits vorgestellten dreidimensionalen Index gemäß der nachfolgenden Gleichung erstellt:  $NID = \sqrt{NID3D * \text{Faktor von Staatlichkeit}}$ . Die Transformation der Daten der *governance*-Indikatoren geschieht analog zum NID3D. Der Einbezug der Staatlichkeitsdaten kann zu einer Reduktion der Demokratiewerte (NID3D) führen, kann diese jedoch nicht verbessern, da Staatlichkeit eine notwendige Bedingung, jedoch kein Merkmal von Demokratie darstellt. Dieser Vorschlag kann die Inhaltvalidität für die Messung eines dreidimensionalen Demokratiekonzepts erhöhen; die oben genannten methodischen Probleme der einzelnen Datenreihen können jedoch damit nicht kompensiert werden. Die Werte liefern eine differenziertes Bild der etablierten Demokratien und lassen sich mit einer externen Validitätsprüfung (Korruptionsmessung von *Transparency International*: CPI) bestätigen. Grundlage für die Regimeklassifikation sind bei dieser Messanlage die Werte von NID3D und NID. Die Schwellenwerte werden von den jeweiligen Quellen entsprechend dem jeweiligen Aggregationsverfahren übertragen.<sup>54</sup> Bei der Klassifizierung ist zu beachten, dass die beim NID3D vor-

54 Vgl. Lauth, H.-J.: Die Qualität der Demokratie, a.a.O.

genommenen Klassifikationen nicht mehr rückgängig gemacht werden. Bei Berücksichtigung dieser Angaben liegen die NID3D-Schwellenwerte für eine defizitäre beziehungsweise funktionierende Demokratie bei den Werten sechs und acht; beim NID bei fünf und sieben.

## 2. Empirische Befunde

An dieser Stelle kann der empirische Befund der verschiedenen Messanlagen nur kurz behandelt werden. Verschiedene Aspekte lassen sich festhalten, wenn wir FH, *Polity*, EUI/DI, V&A, BTI und NID betrachten (vgl. Tab. 1, Anhang). Während FH und *Polity* bei den etablierten Demokratien weitgehend gleiche Bewertungen aufweisen, bieten die anderen vier Messanlagen generell differenzierte Befunde.<sup>55</sup> Dabei zeigt sich, dass die Qualität der Demokratie in verschiedenen Regionen unterschiedlich ausgeprägt ist. Auch bei etablierten Demokratien gibt es beachtlich Unterschiede. Inzwischen haben einige junge Demokratien bereits bessere Werte als manche gestandene Demokratie aufzuweisen. Wenn die vorgeschlagenen Schwellenwerte für die Regimeklassifikation aufgegriffen werden, dann lassen sich die etablierten Demokratien durchweg als funktionierende Demokratien begreifen, während etliche junge Demokratien als defizitäre Demokratie zu bezeichnen wären. Es zeigt sich aber auch, dass die verschiedenen Demokratiemessungen hier keine eindeutigen Lösungen anbieten. Klare Fälle werden übereinstimmend klassifiziert, während unklare Fälle im Grenzbereich zwischen Demokratie und Autokratie oftmals kontrovers beurteilt werden.

Wenn wir die Entwicklung der Demokratiequalität anhand der NID-Daten (vgl. Tab. 2, Anhang) für die Jahre 1998 und 2008 betrachten, dann zeigt sich in der OECD-Welt keine eindeutige Entwicklung.<sup>56</sup> Zunächst sind eine Reihe von Verschlechterungen festzustellen, die etwas kleiner (Ungarn, GB, USA) oder etwas größer ausfallen (am deutlichsten Belgien, Italien und Spanien). Dem stehen etliche Verbesserungen gegenüber, die am stärksten in jungen Demokratien (Südkorea, Mexiko und Türkei) ausfallen; allerdings bleiben diese dann noch immer unter dem OECD-Durchschnitt.<sup>57</sup> Auch die Entwicklung in den neuen EU-Mitgliedstaaten im mittelosteuropäischen Raum ist weitgehend positiv verlaufen. Der Durchschnittswert aller Staaten weltweit hat sich von 4,58 (1998) auf

55 Dies ist bei FH erst ab 2006 möglich.

56 Zur Berechnung der NID-Datensätze und der allgemeinen Aufbereitung der Daten der anderen Messungen hat Oliver Kauff maßgeblich beigetragen.

57 Allerdings erscheinen manche Werte auf den ersten Blick nicht plausibel, wie die Verbesserung der Bewertung der USA von 8.15 (2002) auf 9.02 (2008).

5,02 (2008) verbessert. Dabei ist die Anzahl der Autokratien mit 91 fast konstant geblieben, während die Anzahl der Demokratien leicht gestiegen ist (von 61 auf 66). Von einer Krise der Demokratie kann aus dieser Perspektive nicht gesprochen werden.

Generell sollten in Anbetracht der geäußerten methodischen und theoretischen Kritik an den verschiedenen Messansätzen die präsentierten Daten nicht zu ‚wörtlich‘ genommen werden. Die Daten mit oftmals sehr präzisen Zahlenangaben täuschen eine Genauigkeit vor, die die Messanlagen nicht einlösen können. Wenn wir divergierende Länderbeurteilungen betrachten und beispielsweise die relativ guten Bewertungen von Frankreich und GB in den meisten Messungen mit den doch schwachen Bewertungen beim DB/NCCR kontrastieren, dann sagen die Datenunterschiede weit mehr über die Unterschiede der Messanlagen als über die real existenten Differenzen aus.

#### IV. Fazit

Zusammenfassend lässt sich festhalten: Die Vorstellungen zur Demokratie variieren ebenso stark wie die Anstrengungen in der Formulierung einer präzisen Definition. Das schmalste Demokratieverständnis finden wir bei *Vanhanen*. Bei *Polity* ist es zwar umfassender formuliert, wird aber nur in einer verkürzten Form für die empirische Forschung aufgegriffen. Beide entbehren somit zentraler Merkmale. Vor allem die Dimension der Kontrolle findet keine angemessene Berücksichtigung. *Freedom House*, die Weltbank (*Voice and Accountability*) und die *Economist Unit* (DI) haben zwar eine breitere Messanlage, um verschiedene Dimensionen der Demokratie zu erfassen. Doch sie verzichten auf eine präzise Definition und generell auf eine theoretische Begründung der skizzierten Demokratievorstellung. Dies ist nicht der Fall beim BTI und den SGI, die sich maßgeblich auf die Teilregime der *embedded democracy* und des entsprechenden Begründungszusammenhangs beziehen. Allerdings integriert der BTI mit Blick auf die Stabilität der Demokratie etliche Kriterien der Konsolidierungsforschung, während der SGI letztlich ein um den Rechtsstaat ergänzte Version des *Polyarchy*-Modells von *Dahl* darstellt. Gleichfalls theoretisch fundiert ist das Demokratiekonzept des DB (NCCR), das auf den Dimensionen Freiheit, Gleichheit und Kontrolle basiert; allerdings kann es in seiner Modelllogik weniger überzeugen.

Der NID bezieht sich auf die gleichen Dimensionen, die in einem anderen Kontext umfassender begründet wurden.<sup>58</sup>

Die Validität der Messungen kann bei den Messanlagen nur begrenzt gegeben sein, denen ein unpräziser Maßstab konstatiert wurde (*Polity*, *Freedom House*, V&A, DI). Zudem finden sich dort etliche Indikatoren, die nur sehr bedingt auf die Qualität der Demokratie bezogen werden können. Dies gilt auch für die zweite Gruppe (BTI, SGI, DB und NID), deren Messungen bis auf den SGI auf sehr vielen Indikatoren beruhen. Entsprechend konstatieren *Müller/Pickel*<sup>59</sup> größere Validitätsprobleme in der Operationalisierung. Viele Validitätstests berechnen die Korrelationskoeffizienten, um die Güte der Messung zu belegen. Dabei ignorieren sie den Tatbestand, dass oftmals ähnliche Quellen verwendet werden oder auch in der Anfertigung der Ländergutachten bereits die Messangaben konkurrierender Ansätze einfließen. Zugleich operieren sie mit der gewagten und nicht zutreffenden Annahme, dass der Maßstab gleich und gleiche Werte entsprechend identisch zu interpretieren sind.

Die Reliabilität ist bei *Vanhanen* und bei den Meta-Indizes (V&A, DB und NID) weitgehend gegeben – zumindest, wenn die Quellengrundlage ausreichend dokumentiert wird, was beim DB bislang jedoch nicht der Fall ist. Die Zuverlässigkeit ist dagegen bei den Messanlagen, die selbst Primärdaten erheben, geringer, zumal mit subjektiven *ratings* (Experteneinschätzungen) gearbeitet werden muss. Umso problematischer gestaltet sich der Ansatz, wenn die Transparenz beim Prozess der Datenerhebung und -auswertung gering ist, wie dies bei *Freedom House* und auch in etwas geringerem Umfang bei *Polity* der Fall ist. Sehr positiv besticht hier die Transparenz beim BTI und beim SGI. Allerdings eröffnet diese Transparenz dann große Diskrepanzen in den Bewertungen der SGI, deren Reliabilität bislang nicht überzeugen kann. Auch die Aggregationsverfahren von SGI und DB mit der numerischen Vergrößerung der Differenzen ist aus Validitätsgründen nicht unproblematisch; zudem erschwert sie die Vergleichbarkeit mit anderen Datensätzen.

58 Lauth, H.-J.: Dimensionen der Demokratie und das Konzept einer defekten Demokratie, in: Pickel, G./Pickel, S./Jacobs, J. (Hg.): Demokratie – Entwicklungsformen und Erscheinungsbilder im interkulturellen Vergleich, Frankfurt/Oder, 1997, 33-54; Lauth, H.-J.: Demokratie und Demokratiemessung, a.a.O.

59 Müller, T./Pickel, S., a.a.O.

Angesichts dieser eher skeptischen Einschätzungen scheint es erstaunlich, dass viele Datenbestände in der Politikwissenschaft breite Verwendung finden. Dennoch können die Daten von Nutzen für die Forschung sein:

- Erstens können die Resultate vorsichtig und in einem begrenzten Sinne für eine erste Sichtung der Ausgangssituation verwendet werden, sie sind jedoch nicht ausreichend für eine vertiefende Untersuchung.
- Zweitens können die Messanlagen herangezogen werden, um zur Weiterentwicklung der bestehenden Verfahren beizutragen oder um die Entwicklung neuer Messverfahren zu unterstützen.
- Drittens können sie bereits jetzt deutlich machen, dass die Qualität der Demokratie vielfältig variiert. Dies betrifft junge wie etablierte Demokratien.
- Viertens besteht aufgrund der empirisch gegebenen Differenzen ein wachsender Bedarf an Demokratieforschung. Die theoretische Forschung, in der die Qualität der Demokratie als unabhängige oder abhängige Variable eingesetzt wird, steht noch in den Anfängen.

## V. Anhang

*Tabelle 1: Ergebnisse Demokratiemessung 2008 (alle Länder)*

Country	Polity 2008 DEMOC (2008, Dec)	Freedom House 2009 PR (2008, Jan-Dec)	Democracy Index 2008 (2008, Sept)	BTI 2008 (2007, Jan)	Voice & Accountability 2008 (2008)	NID 3D (2008)	NID (2008)
<b>Afghanistan</b>	-66,00	5,00	3,02	3,57	2,59	..	..
<b>Albania</b>	9,00	3,00	5,91	7,50	6,29	6,45	6,45
<b>Algeria</b>	3,00	6,00	3,32	4,27	3,16	2,77	2,77
<b>Andorra</b>	..	1,00	..	..	9,49	..	..
<b>Angola</b>	2,00	6,00	3,35	3,97	3,11	2,16	2,16
<b>Argentina</b>	8,00	2,00	6,63	7,85	6,80	6,67	6,67
<b>Armenia</b>	5,00	6,00	4,09	6,00	4,20	3,47	3,47
<b>Australia</b>	10,00	1,00	9,09	..	9,56	9,85	9,62
<b>Austria</b>	10,00	1,00	8,49	..	9,54	9,97	9,92
<b>Azerbaijan</b>	0,00	6,00	3,19	3,80	2,68	0,00	0,00
<b>Bahamas</b>	..	1,00	..	..	8,95	..	..
<b>Bahrain</b>	0,00	5,00	3,38	4,63	3,77	0,00	0,00
<b>Bangladesh</b>	0,00	4,00	5,52	5,95	4,34	0,00	0,00
<b>Barbados</b>	..	1,00	..	..	9,02	..	..

Country	Polity 2008 DEMOC (2008, Dec)	Freedom House 2009 PR (2008, Jan-Dec)	Democracy Index 2008 (2008, Sept)	BTI 2008 (2007, Jan)	Voice & Accountability 2008 (2008)	NID 3D (2008)	NID (2008)
Belarus	0,00	7,00	3,34	3,93	1,71	0,00	0,00
Belgium	8,00	1,00	8,16	..	9,59	8,88	8,62
Belize	..	1,00	..	..	7,90	..	..
Benin	7,00	2,00	6,06	7,90	6,85	6,45	6,45
Bhutan	3,00	4,00	4,3	3,75	4,02	4,62	4,62
Bolivia	8,00	3,00	6,15	6,40	5,93	5,64	5,24
Bosnia-Herzegovina	-66,00	4,00	5,7	6,70	5,96	..	..
Botswana	8,00	2,00	7,47	8,45	7,39	7,81	7,81
Brazil	8,00	2,00	7,38	7,95	7,29	6,99	6,90
Brunei	..	6,00	..	..	3,30	..	..
Bulgaria	9,00	2,00	7,02	8,70	7,52	7,45	7,45
Burkina Faso	2,00	5,00	3,6	6,25	5,07	3,21	3,21
Burundi	7,00	4,00	4,51	5,80	4,19	4,96	4,45
Cambodia	3,00	6,00	4,87	4,13	3,46	2,58	2,58
Cameroon	1,00	6,00	3,46	4,13	3,24	1,82	1,82
Canada	10,00	1,00	9,07	..	9,72	9,88	9,58
Cape Verde	..	1,00	7,81	..	8,46	..	..
Central African Rep.	1,00	5,00	1,86	4,28	3,30	2,07	2,07
Chad	1,00	7,00	1,52	2,83	2,10	0,00	0,00
Chile	10,00	1,00	7,89	9,30	8,55	9,46	8,85
China	0,00	7,00	3,04	3,15	1,38	0,00	0,00
Colombia	7,00	3,00	6,54	5,75	5,26	6,03	4,58
Comoros	9,00	3,00	3,58	..	4,79	5,98	5,41
Congo, Brazzaville	0,00	6,00	2,94	3,63	2,86	0,00	0,00
Congo, Kinshasa	6,00	6,00	2,28	3,72	2,04	2,78	2,37
Costa Rica	10,00	1,00	8,04	9,50	8,55	8,76	8,51
Cote d'Ivoire	..	6,00	3,27	2,90	2,67	..	..
Croatia	9,00	2,00	7,04	8,85	7,23	7,64	7,64
Cuba	0,00	7,00	3,52	3,42	1,04	0,00	0,00
Cyprus	10,00	1,00	7,7	..	8,57	9,28	8,71
Czech Republic	8,00	1,00	8,19	9,55	8,66	8,47	8,47
Denmark	10,00	1,00	9,52	..	9,87	9,97	9,58
Djibouti	..	5,00	2,37	..	2,98	..	..
Dominican Republic	8,00	2,00	6,2	7,60	6,32	6,69	6,69
Ecuador	5,00	3,00	5,64	6,15	5,36	4,71	4,71
Egypt	1,00	6,00	3,89	4,40	2,78	2,10	2,10
El Salvador	7,00	2,00	6,4	7,20	6,10	6,37	6,37
Eritrea	0,00	7,00	2,31	2,60	0,11	0,00	0,00
Estonia	9,00	1,00	7,68	9,55	8,68	8,98	8,63
Ethiopia	3,00	5,00	4,52	4,13	2,51	3,55	3,38
Fiji	0,00	6,00	5,11	..	4,22	0,00	0,00
Finland	10,00	1,00	9,25	..	9,85	9,93	9,93

Country	Polity 2008 DEMOC (2008, Dec)	Freedom House 2009 PR (2008, Jan-Dec)	Democracy Index 2008 (2008, Sept)	BTI 2008 (2007, Jan)	Voice & Accountability 2008 (2008)	NID 3D (2008)	NID (2008)
France	9,00	1,00	8,07	..	9,24	9,25	8,77
Gabon	0,00	6,00	3,00	..	3,73	0,00	0,00
Gambia	0,00	5,00	4,19	..	3,37	0,00	0,00
Georgia	7,00	4,00	4,62	6,85	5,27	5,61	5,25
Germany	10,00	1,00	8,82	..	9,48	9,82	9,60
Ghana	8,00	1,00	5,35	8,10	7,23	7,63	7,41
Greece	10,00	1,00	8,13	..	8,28	9,04	8,37
Guatemala	8,00	3,00	6,07	5,90	5,27	5,66	5,66
Guinea	1,00	7,00	2,09	3,98	2,45	0,00	0,00
Guinea-Bissau	6,00	4,00	1,99	..	3,85	4,33	4,33
Guyana	6,00	2,00	6,12	..	6,39	5,98	5,91
Haiti	6,00	4,00	4,19	4,40	4,06	4,42	4,24
Honduras	7,00	3,00	6,18	6,65	5,17	5,64	5,64
Hong Kong	..	5,00	5,85	..	7,24	..	..
Hungary	10,00	1,00	7,44	9,35	8,60	9,10	8,71
Iceland	..	1,00	9,65	..	9,78	..	..
India	9,00	2,00	7,8	8,10	7,15	7,68	6,15
Indonesia	8,00	2,00	6,34	6,45	5,56	6,63	5,70
Iran	0,00	6,00	2,83	3,73	2,02	0,00	0,00
Iraq	-66,00	6,00	4	3,30	2,62	..	..
Ireland	10,00	1,00	9,01	..	9,67	9,84	9,70
Israel	10,00	1,00	7,48	..	7,78	9,15	6,10
Italy	10,00	1,00	7,98	..	8,49	8,75	8,34
Jamaica	9,00	2,00	7,21	8,55	7,56	7,07	6,77
Japan	10,00	1,00	8,25	..	8,47	9,57	9,33
Jordan	2,00	5,00	3,93	3,98	4,05	3,57	3,57
Kazakhstan	0,00	6,00	3,45	4,23	3,27	0,00	0,00
Kenya	7,00	4,00	4,79	7,00	5,53	5,04	4,69
Korea, North	0,00	7,00	0,86	2,70	0,08	0,00	0,00
Korea, South	8,00	1,00	8,01	8,85	7,50	8,43	8,18
Kosovo	8,00	6,00	..	..	4,23	3,89	3,89
Kuwait	0,00	4,00	3,39	4,08	4,54	0,00	0,00
Kyrgyzstan	3,00	5,00	4,05	5,95	4,04	3,13	3,13
Laos	0,00	7,00	2,1	2,78	1,42	0,00	0,00
Latvia	8,00	2,00	7,23	8,70	8,23	7,88	7,88
Lebanon	8,00	5,00	5,62	6,25	4,89	4,82	3,73
Lesotho	8,00	2,00	6,29	..	6,05	7,00	7,00
Liberia	7,00	3,00	5,25	5,25	5,17	5,27	5,10
Libya	0,00	7,00	2	2,98	0,92	0,00	0,00
Lithuania	10,00	1,00	7,36	9,35	8,21	8,89	8,76
Luxembourg	..	1,00	9,1	..	9,92	..	..
Macedonia	9,00	3,00	6,21	7,75	6,37	6,73	6,56

Country	Polity 2008 DEMOC (2008, Dec)	Freedom House 2009 PR (2008, Jan-Dec)	Democracy Index 2008 (2008, Sept)	BTI 2008 (2007, Jan)	Voice & Accountability 2008 (2008)	NID 3D (2008)	NID (2008)
Madagascar	7,00	4,00	5,57	7,45	5,53	5,51	5,51
Malawi	6,00	4,00	5,13	6,60	5,46	5,36	5,36
Malaysia	6,00	4,00	6,36	5,33	4,41	5,90	5,90
Mali	7,00	2,00	5,87	7,25	6,69	6,64	6,63
Malta	..	1,00	8,39	..	9,16	..	..
Mauritania	0,00	6,00	3,91	4,57	3,50	0,00	0,00
Mauritius	10,00	1,00	8,04	8,80	8,28	9,16	9,01
Mexico	8,00	2,00	6,78	7,45	6,14	6,64	6,16
Moldova	8,00	4,00	6,5	6,85	5,23	5,76	5,76
Mongolia	10,00	2,00	6,6	7,00	6,57	7,27	7,27
Montenegro	8,00	3,00	6,43	7,85	6,61	6,68	6,68
Morocco	0,00	5,00	3,88	4,40	4,10	0,00	0,00
Mozambique	6,00	3,00	5,49	6,55	5,90	5,58	5,58
Myanmar (Burma)	0,00	7,00	1,77	1,70	0,00	0,00	0,00
Namibia	6,00	2,00	6,48	8,10	7,46	6,89	6,89
Nepal	7,00	4,00	4,05	3,90	3,86	5,25	4,24
Netherlands	10,00	1,00	9,53	..	10,00	9,85	9,48
New Zealand	10,00	1,00	9,19	..	9,86	9,92	9,74
Nicaragua	9,00	4,00	6,07	6,65	5,57	5,61	5,61
Niger	7,00	3,00	3,41	6,43	4,85	5,74	5,60
Nigeria	4,00	5,00	3,53	6,05	4,36	3,55	3,12
Norway	10,00	1,00	9,68	..	10,00	10,00	9,96
Oman	0,00	6,00	2,98	3,63	3,11	0,00	0,00
Pakistan	5,00	4,00	4,46	3,65	3,27	4,56	2,55
Panama	9,00	1,00	7,35	7,80	7,49	7,83	7,56
Papua New Guinea	4,00	4,00	6,54	6,10	6,19	4,22	4,22
Paraguay	9,00	3,00	6,4	6,75	5,08	5,98	5,85
Peru	9,00	2,00	6,31	6,60	5,99	6,80	5,98
Philippines	8,00	4,00	6,12	6,30	5,41	5,74	4,80
Poland	10,00	1,00	7,3	8,80	8,23	8,81	8,79
Portugal	10,00	1,00	8,05	..	9,09	9,27	9,27
Qatar	0,00	6,00	2,92	..	3,90	0,00	0,00
Romania	9,00	2,00	7,06	8,55	7,22	7,52	7,52
Russia	5,00	6,00	4,48	5,35	3,38	3,17	3,17
Rwanda	0,00	6,00	3,71	3,67	2,67	0,00	0,00
Saudi Arabia	0,00	7,00	1,9	2,72	1,35	0,00	0,00
Senegal	8,00	3,00	5,37	7,10	5,52	6,48	6,48
Serbia	9,00	3,00	6,49	7,75	6,46	6,60	6,28
Sierra Leone	8,00	3,00	4,11	6,05	5,19	5,75	5,75
Singapore	2,00	5,00	5,89	5,37	4,87	3,98	3,98
Slovakia	10,00	1,00	7,33	9,20	8,31	8,83	8,83
Slovenia	10,00	1,00	7,96	9,70	8,65	9,18	9,18

Country	Polity 2008 DEMOC (2008, Dec)	Freedom House 2009 PR (2008, Jan-Dec)	Democracy Index 2008 (2008, Sept)	BTI 2008 (2007, Jan)	Voice & Accountability 2008 (2008)	NID 3D (2008)	NID (2008)
<b>Solomon Islands</b>	9,00	4,00	..	..	6,44	5,69	5,69
<b>Somalia</b>	-77,00	7,00	..	1,43	1,05	..	..
<b>South Africa</b>	9,00	2,00	7,91	8,60	7,75	7,68	7,31
<b>Spain</b>	10,00	1,00	8,45	..	8,92	9,38	8,10
<b>Sri Lanka</b>	7,00	4,00	6,61	6,30	4,78	5,86	3,95
<b>Sudan</b>	0,00	7,00	2,81	2,13	1,27	0,00	0,00
<b>Swaziland</b>	0,00	7,00	3,04	..	2,76	0,00	0,00
<b>Sweden</b>	10,00	1,00	9,88	..	10,00	9,95	9,72
<b>Switzerland</b>	10,00	1,00	9,15	..	9,80	9,93	9,82
<b>Syria</b>	0,00	7,00	2,18	2,60	1,32	0,00	0,00
<b>Taiwan</b>	10,00	2,00	7,82	9,45	7,81	8,52	8,52
<b>Tajikistan</b>	1,00	6,00	2,45	3,73	2,45	1,78	1,78
<b>Tanzania</b>	2,00	4,00	5,28	6,85	5,71	3,73	3,73
<b>Thailand</b>	5,00	5,00	6,81	5,10	4,47	4,57	4,53
<b>Timor, East</b>	7,00	3,00	7,22	..	6,35	5,36	4,98
<b>Togo</b>	1,00	5,00	2,43	3,90	2,95	2,38	2,38
<b>Trinidad &amp; Tobago</b>	10,00	2,00	7,21	..	7,36	7,58	7,41
<b>Tunisia</b>	1,00	7,00	2,96	3,95	2,60	0,00	0,00
<b>Turkey</b>	8,00	3,00	5,69	7,05	5,44	6,83	6,12
<b>Turkmenistan</b>	0,00	7,00	1,72	2,78	0,47	0,00	0,00
<b>Uganda</b>	1,00	5,00	5,03	6,80	4,69	2,50	2,50
<b>Ukraine</b>	7,00	3,00	6,94	7,35	5,87	5,91	5,91
<b>United Arab Emirits.</b>	0,00	6,00	2,6	3,47	3,36	0,00	0,00
<b>United Kingdom</b>	10,00	1,00	8,15	..	9,47	9,79	9,00
<b>United States</b>	10,00	1,00	8,22	..	8,92	9,77	9,02
<b>Uruguay</b>	10,00	1,00	8,08	9,90	8,64	8,82	8,82
<b>Uzbekistan</b>	0,00	7,00	1,74	3,22	0,91	0,00	0,00
<b>Venezuela</b>	5,00	4,00	5,34	5,65	4,29	3,89	3,89
<b>Vietnam</b>	0,00	7,00	2,53	3,15	1,66	0,00	0,00
<b>Yemen</b>	1,00	5,00	2,95	3,90	2,83	2,33	2,33
<b>Zambia</b>	5,00	3,00	5,25	6,80	5,71	5,39	5,39
<b>Zimbabwe</b>	1,00	7,00	2,53	3,97	1,91	0,00	0,00

Anmerkungen: Bei allen Daten wurden bis auf Voice and Accountability die Originaldaten übernommen. Die Angaben von V&A wurden durch eine z-Transformation in der Darstellung (0 bis 10) angeglichen. Bei den FH-Daten repräsentiert der Wert 1 den höchsten Freiheitsgrad. Die Messergebnisse des BTI wurden im Jahr 2007 erhoben; Quelle: vgl. die entsprechenden Angaben im Text.

Tabelle 2: Messergebnisse OECD-Länder (ausgewählte Jahre)

Country	SGI 2009 (2005 Jan -2007 March)	Demokratie- barometer (Average 1995-2005)	NID (1998)	NID (2008)
Australia	7,36	6,72	9,63	9,62
Austria	7,17	6,46	9,74	9,92
Belgium	7,02	7,41	9,21	8,62
Canada	8,56	6,59	9,51	9,58
Czech Republic	6,55	4,82	8,73	8,47
Denmark	8,74	8,42	9,66	9,58
Finland	9,91	8,16	9,78	9,93
France	6,71	4,49	8,86	8,77
Germany	8,89	6,19	9,62	9,60
Greece	4,90	..	8,03	8,37
Hungary	5,79	4,92	8,90	8,71
Ireland	7,80	5,98	9,59	9,70
Island	7,89	7,98	..	..
Italy	5,14	4,29	8,71	8,34
Japan	5,70	5,08	9,48	9,33
Korea, South	5,02	..	7,38	8,18
Luxembourg	6,39	7,44	..	..
Mexico	3,10	..	5,47	6,16
Netherlands	9,36	7,73	9,84	9,48
New Zealand	9,14	6,97	9,84	9,74
Norway	9,85	7,74	9,81	9,96
Poland	4,42	3,40	8,31	8,79
Portugal	6,94	5,16	9,36	9,27
Slovakia	5,87	..	7,59	8,83
Spain	6,32	5,46	8,57	8,10
Sweden	9,36	8,14	9,78	9,72
Switzerland	8,95	7,15	10,00	9,82
Turkey	1,48	..	5,12	6,12
United Kingdom	7,50	4,77	9,26	9,00
United States	7,32	6,47	9,31	9,02

Anmerkung: Der Wert 10 repräsentiert die höchste Ausprägung der Qualität der Demokratie.