

Berthold Kuhn

Nutzung und Wertschätzung von BTI und FHI in Wissenschaft und Praxis

Eine Expertenbefragung

Kurzfassung

Indizes zur Messung von Demokratie, Governance und Transformationsprozessen haben sich in den letzten Jahren in der Politikwissenschaft zunehmend verbreitet. Dieser Beitrag befasst sich mit dem Bertelsmann Transformation Index (BTI) und dem Freedom House Index (FHI), die zu den bekanntesten Indizes in ihren jeweiligen Themenfeldern zählen. Der Artikel stellt die Ergebnisse einer Befragung von Wissenschaftlern, Praktikern, sogenannten „Grenzgängern“ und Studierenden unterschiedlichen Alters und Geschlechts vor, die zu Themen der internationalen Politik und/oder zu Demokratisierungs- und Transformationsprozessen arbeiten. Die Befragung von 82 potenziellen deutschsprachigen Nutzern der Indizes ergab, dass der BTI trotz durchschnittlich höherer Wertschätzung etwas seltener als der FHI genutzt wird. Die Indizes bekommen durchaus Wertschätzung, aber auch vielfältige Kritik, hauptsächlich zu konzeptionellen und methodischen Aspekten. Eine wesentliche Empfehlung lautet, regionale und nationale Institutionen stärker in die Bewertungen und Berichte einzubinden.

Inhalt

1. Einleitung	578
2. Bedeutungszuwachs und Bewertung von Indizes	580
3. Der Bertelsmann Transformation Index (BTI)	583
4. Der Freedom House Index (FHI)	585
5. Ergebnisse der Expertenbefragung	586
a) Nutzung der Indizes	587
b) Allgemeine und spezifische Wertschätzung der Indizes	589
c) Beurteilung des Wertes der Indizes für Entscheidungsträger	590
d) Das Spektrum der genannten Stärken	592
e) Das Spektrum der genannten Schwächen	594
f) Das Spektrum der genannten Empfehlungen	597
6. Fazit	598

1. Einleitung

Dieser Aufsatz befasst sich mit zwei Indizes, die sowohl in der Wissenschaft als auch in der Praxis der internationalen Politik einen hohen Bekanntheitsgrad erlangt haben: dem Bertelsmann Transformation Index (BTI) und dem Freedom House Index (FHI). Mit der empirischen Erhebung von Daten zu Anwendern von Indizes sowie Anwendungspraktiken erschließt der Aufsatz ein bisher weitgehend unbeachtetes Forschungsfeld und bietet eine neue Perspektive in der politikwissenschaftlichen Diskussion zur Evaluation von Demokratie- und Governance-Indizes.

Der BTI ist ein komplexer Index jüngerer Datums, der FHI ein eher wenig komplexer Index, der bereits seit vielen Jahren international etabliert ist. Diese beiden Indizes wurden für die Befragungen ausgewählt, weil sie in Vorgesprächen sowohl mit Wissenschaftlern als auch mit Praktikern als die am weitesten verbreiteten Indizes in Deutschland angesehen wurden, die umfassende und vergleichende Analysen und Bewertungen der politischen Entwicklungen von Staaten bzw. Regionen präsentieren. Beide Indizes, speziell jedoch der BTI, bieten im Vergleich zu anderen Indizes über Skalen hinaus relativ ausführliche Länderberichte an. Der BTI analysiert und bewertet die Qualität von Demokratie, Marktwirtschaft und politischem Management in Entwicklungs- und Transformationsländern; der FHI bewertet den Grad an Demokratie und Freiheit, speziell politische Rechte und bürgerliche Freiheiten. Dem BTI liegt insgesamt betrachtet ein stärker konsensorientierter, dem FHI ein stärker wettbewerbsorientierter Demokratiebegriff zugrunde.

Mit den Expertenbefragungen, die vom Autor und seinem Team¹ durchgeführt wurden, sollten quantitative und qualitative Ergebnisse zur Nutzung, Wertschätzung sowie den Stärken und Schwächen der Indizes ermittelt werden. Darüber hinaus wurden die Befragten gebeten, sich zu Empfehlungen für Verbesserungen dieser beiden Indizes zu äußern. Insgesamt wurden rund 180 deutschsprachige Personen angeschrieben, die als potenzielle Nutzer der Indizes gelten konnten.² Darunter sollten sich sowohl Wissenschaftler als auch Praktiker unterschiedlichen Alters, Status und Geschlechts befinden, die zu Themen der internationalen Politik, der internationalen Zusammenarbeit, der Demokratisierung und/oder zu Transformationsprozessen arbeiten. Die als potenzielle Nutzer identifizierten Experten³ wurden um Interviews und/oder um das Ausfüllen eines Fragebogens gebeten. Insgesamt 82 Befragte äußerten sich im Rahmen von Interviews und/oder auf Basis des Fragebogens. Darüber hinaus gingen einige freie Kommentare zu den Indizes ein. Der Autor hat sich dafür entschieden, prägnante Äußerungen aus den Interviews und Fragebögen mit dem ausdrücklichen Einverständnis der Befragten – speziell zu den Stärken, Schwächen und Empfehlungen – zu zitieren, um das breite Spektrum der Antworten zum Umgang mit den Indizes authentisch illustrieren zu können.

Das Erkenntnisinteresse dieses Beitrags fokussiert auf die Perzeption der beiden ausgewählten Indizes in Wissenschaft und Praxis. Fragen nach der Produktion und Validität von Datensätzen werden dagegen nicht näher erörtert.⁴ Die Darstellung der Ergebnisse konzentriert sich auf die signifikantesten Erkenntnisse der Befra-

1 Urs Schrade, M.A., Doktorand, und Eliza Girod, Studentin.

2 Das Kriterium der deutschen Sprache wurde aus zwei Gründen gewählt: Erstens musste ein überschaubarer Kreis von Personen identifiziert werden, aus dem die Experten ausgewählt werden sollten. Eine internationale Befragung hätte die Möglichkeiten des Vorhabens überschritten. Zweitens hat der relativ junge BTI bisher überwiegend in Deutschland Verbreitung gefunden.

3 Potenzielle Nutzer wurden überwiegend durch Veröffentlichungen in Fachzeitschriften, teilweise auch mittels Teilnehmerlisten an Konferenzen sowie durch persönliche Netzwerke identifiziert.

4 Im Rahmen des Forschungsvorhabens wurde die Frage der Validität der Daten speziell am Beispiel der V.R. China und Indiens untersucht, unter anderem mittels vergleichenden Analysen zwischen den numerischen Bewertungen und den entsprechenden Passagen in den jeweiligen Länderberichten. Die Analysen zeigten, dass die Bewertungen deutlich von normativen Sichtweisen beeinflusst sind und teilweise nicht mit den Ausführungen in den Länderberichten korrespondieren. Es drängt sich der Eindruck auf, dass China strenger als Indien bewertet wird. Die Bertelsmann Stiftung bestätigte, dass bestimmte Werte im Bereich der Steuerungsleistungen bei autoritär eingestuften Regimen bei der Hälfte der Punktzahl gedeckelt werden. Bewertungen über verschiedene Provinzen bzw. Bundesstaaten und Megastädte hinweg sind bei großen Staaten besonders problematisch. Die Ergebnisse sind bisher noch nicht veröffentlicht. Zur Frage der Validität von Datensätzen bieten speziell die Einführung und die Bestandsaufnahme von Cas Mudde und Andreas Schedler in Political Research Quarterly einen guten Überblick (Mudde/Schedler 2010; Schedler/Mudde 2010), die durch weitere Beiträge im gleichen Heft ergänzt werden.

gung, speziell auf die tendenziell unterschiedlichen Perzeptionen der Indizes zwischen Wissenschaftlern und Praktikern.

2. Bedeutungszuwachs und Bewertung von Indizes

Indizes zur Messung von Demokratisierungs-, Entwicklungs- und Transformationsprozessen erfreuen sich speziell seit Mitte der 1990er Jahre einer zunehmenden Beliebtheit in Wissenschaft und Praxis. Diese Entwicklung ist vor allem auf die Neubelebung der Demokratiedebatte nach dem Ende des Ost-West-Konfliktes und dem Aufschwung der Governance-Forschung zurückzuführen. Die Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (eng. OECD) führt darüber hinaus den Einfluss der *New Institutional Economics* als Grund an, speziell die Arbeiten von Douglas North (Arndt/Oman 2006: 18; North 1990, 2005).

Als Ausgangspunkt der modernen empirischen Demokratieforschung kann Robert Dahls Polyarchiemodell angesehen werden (Dahl 1971). Die empirische Demokratieforschung hat sich in den darauf folgenden Jahrzehnten differenziert; die meisten Konzepte und Indizes stammen aus den 1990er Jahren. Die Literatur konstatiert zudem, dass Demokratiemessung ein insgesamt noch relativ junges Unterfangen ist.⁵

Die Überwindung des Ost-West-Konfliktes und die nachfolgende Beteiligung internationaler Institutionen an Transformationsprozessen haben einen neuen und tieferen Einblick in vormals schwer zugängliche politische Systeme ermöglicht. Bessere Statistiken und neue Informationstechnologien erleichterten den Vergleich von Daten und die Erstellung von Indizes. Nachhaltig wirkten in dieser Zeit die Thesen von Fukuyama (1992) vom weltweiten Siegeszug der liberalen Demokratie und Marktwirtschaft. Theoretische und methodische Probleme sowie unstete und wenig transparente Transformationsprozesse machen es jedoch schwer, glaubwürdige und gültige Aussagen von gewisser Dauer zu generieren (Patz 2010: 33). Die OECD sieht in dem enorm ansteigenden Interesse an der Qualität von Governance seitens der Entscheidungsträger in Politik und Wirtschaft den Grund für die Entwicklung der zahlreichen Indikatoren zur vergleichenden Messung von Governance in verschiedenen Staaten und Regionen (Arndt/Oman 2006: 13).

5 Vergleiche Patz (2010: 32) mit Hinweis auf Pickel/Pickel (2006: 151-270) und Lauth (2004: 227-318).

Die Anliegen von Bertelsmann Stiftung und Freedom House sind in ihrer Stoßrichtung politischer als die von Weltbank⁶ und OECD. Letztere verfolgen Agenden, die vor allem verlässliche politische Rahmenbedingungen für wirtschaftliche Investitionen und die internationale Zusammenarbeit im Blick haben. Besonders der BTI steht dabei in einer Reihe mit komplexeren Indizes, die in der Wissenschaft und Praxis eine zunehmende Verbreitung gefunden haben.⁷ Angesichts des Aufwands und der Ressourcen, mit denen viele Indizes entwickelt wurden, ist die Anzahl der unabhängigen Forschungsarbeiten zu Indizes allerdings eher dürftig.

Die Analyse von Munck/Verkuilen aus dem Jahr 2002 gilt international als das meist zitierte Vorhaben einer kritischen Auseinandersetzung mit den Methoden der Indizes. Die Autoren untersuchten Vorhaben der Konzeptionalisierung, der Messung und der Aggregation von Daten zur Messung von Demokratie. Sie stellen ihre Ergebnisse in einer Übersicht zu Stärken und Schwächen dar (ebd.: 6). Die Autoren kommen zu dem Ergebnis, dass viele Indizes mit einer zu minimalistischen Definition von Demokratie arbeiten und sich zu sehr auf kompetitive Verfahren der Regierungsbildung konzentrieren. Die Qualität der Ausübung und die Begrenzung von Regierungsmacht würden nicht ausreichend berücksichtigt. Auch würden die meisten Indizes den Aspekt der Partizipation vernachlässigen. Dies träfe besonders auf den Polity IV-Index zu, in abgeschwächter Form auch auf den von Alvarez, Cheibub, Limongi und Przeworski (1996) so benannten ACLP Index. Insgesamt kommen sie zu einer eher kritischen Einschätzung, was die Qualität und die Zusammenführung der Daten und damit die Bewertungen betrifft. Freedom House betreffend verweisen Munck/Verkuilen auf die Kritik von Ryan (1994) und Bollen (1986: 584), dass die Bezüge zwischen den zahlreichen Kriterien und deren Relevanz für Demokratie nicht ausreichend reflektiert würden. Sie bemängeln zudem, dass Freedom House die Regeln für die Kodierung nicht offenlege und damit die Ergebnisse nicht ausreichend nachvollziehbar seien (Munck/Verkuilen 2002: 19,

6 Kaufmann und Kollegen haben für die Weltbank umfangreiche Research Papers zu Governance-Indikatoren entwickelt. Exemplarisch hierzu Kaufmann/Kraay/Zoido-Lobaton 1999, die zum Teil kritisch diskutiert wurden (vgl. Knack 2006 und Thomas 2006).

7 Dazu gehören Indizes wie das Demokratiebarometer (Bühlmann u. a. 2012; www.democracybarometer.org), der Failed State Index von Foreign Policy (2011), der Mo Ibrahim Index (Mo Ibrahim Foundation 2011), der Economic Freedom of the World Index des Fraser Instituts (Gwartney/Hall/Lawson 2010), der Civil Society Index von CIVICUS (2011) und der Rule of Law Index des World Justice Projects (Agrast/Botero/Ponce 2011). Der Rule of Law Index ist das jüngste Beispiel für einen Index, der international in Wissenschaft und Praxis rasch Verbreitung gefunden hat. Das World Justice Project hat nach mehrjähriger Entwicklungsphase diesen Index aufgelegt, der Rechtsstaatlichkeit auf der Basis von acht bzw. neun Faktoren untersucht, denen insgesamt 52 Indikatoren zugeordnet sind. Der Index Report für 2011 umfasst 66 Staaten. Die Berichterstattung und Bewertung für 2012 soll auf 100 Staaten ausgeweitet werden.

21).⁸ Coppedge u. a. (2011) kritisieren den FHI, weil er nicht ausreichend zwischen verschiedenen Formen von Demokratien differenziere, die Indikatoren nicht trennscharf genug seien und es keine ausreichend strengen und transparenten Kodierungsregeln gäbe. Der BTI wird weder von Munck/Verkuilen noch von Coppedge u. a. näher untersucht.

In Deutschland unterzogen Müller/Pickel (2007) in einem Forumsbeitrag der PVS sechs Indizes einer Bewertung, die empirische Beiträge zur Bestimmung der Qualität von Demokratien geliefert haben, darunter BTI (Bertelsmann Stiftung), Polity Index (Jagers/Gurr 1995), FHI (Freedom House und Gastil 1991), Vanhanen-Index (Vanhanen 1990), Index Defekter Demokratien (Croissant/Thierry 2000; vgl. Merkel u. a. 2003) sowie den Democratic Audit (Beetham 1994). Der Index Defekter Demokratie von Croissant/Thierry (2000) schneidet bei der Evaluierung am besten ab, der FHI am schlechtesten (Müller/Pickel 2007: 532). Die Autoren konzentrieren sich primär auf die Konzeptvalidierung und zeigen mit einem Quervergleich der Indizes Schwachstellen in der Konzeptionalisierungsphase auf. Sie stellen fest, dass die Demokratiekonzepte häufig intern nicht hinreichend trennscharf sind und es zwischen den Dimensionen inhaltliche Überschneidungen gibt (ebd.: 534). Die Kritik zielt insgesamt darauf ab, dass zugunsten von Ergebnissen, die zwischen Autokratien und Demokratien unterscheiden, zuweilen das Verfahren der Indexkomposition die Sorgfalt vermissen lässt, die für eine durchgängige Validität und Reliabilität der Messung nötig wäre (Müller/Pickel 2007, 2008). Bühlmann u. a. (2008: 116) bemerken dazu in einer Replik, dass die Unzulänglichkeiten der meisten bisherigen Indizes auf dem Umstand beruhen, dass sie – beeinflusst von dem Dahlschen Polyarchie-Konzept (Dahl 1971) – sich auf ein zu minimalistisches Demokratiekonzept berufen.

Dirk Berg-Schlosser (2007) liefert eine kurze kritische Betrachtung des BTI und anderer Indizes. Seine Kritik bezieht sich auf die fehlenden *hard data* und damit die Subjektivität der Daten (Perzeptionsproblem), die Methode der Aggregation dieser und die fehlende Präzision bei der Bestimmung von Bewertungen. Er weist auch darauf hin, dass politische Systeme mit Kriterien gemessen werden, die für sie selbst nicht als Referenzrahmen gelten. Systeme, die sich liberale Demokratien nicht zum Orientierungsmaßstab nehmen, können kaum an der Erreichung von Maßstäben gemessen werden, die sie nicht anstreben (36).

Eine Publikation der OECD (Arndt/Oman 2006) stellt eine Reihe von Governance-Indikatoren vor, erläutert die Relevanz von Indikatoren, diskutiert deren Potenziale und Grenzen und hat einen Leitfaden zum Umgang mit diesen ver-

8 Freedom House veröffentlicht seit 2003 die Bewertungen der Subkategorien.

öffentlich (OECD 2006). Die Bertelsmann Stiftung hat im Jahr 2009 selbst umfangreiche Evaluierungen ihres Index in Auftrag gegeben (vgl. Ahrens 2010; Bühlmann/Hänni/Zumbach 2010).

Insgesamt betrachtet scheint das Spektrum der Wertschätzung, Kritik und der Empfehlungen hinsichtlich der Indizes nicht umfangreich ausgeleuchtet zu sein. Eine Befragung von Experten aus Wissenschaft und Praxis zu Governance-Indizes, die quantitativ und qualitativ ausgewertet wurde, liegt für deutschsprachige Experten nicht vor und ist dem Autor auch auf internationaler Ebene nicht bekannt. Diese Lücke möchte der vorliegende Beitrag in bescheidenem Umfang versuchen zu schließen.

3. Der Bertelsmann Transformation Index (BTI)

Der Bertelsmann Transformationsindex wird von der Bertelsmann Stiftung in Gütersloh erstellt, die zu den größten deutschen Unternehmensstiftungen gehört.⁹ Sie befasst sich seit 1996 mit Entwicklungs- und Transformationsforschung und veröffentlichte den umfassenden Index erstmals im Jahr 2003. Seit dem Jahr 2006 sind alle zwei Jahre Veröffentlichungen vorgesehen. Die Zusammenarbeit mit dem Centrum für angewandte Politikforschung (CAP), das wesentlich in die Entwicklung des Index involviert war, wurde von der Bertelsmann Stiftung inzwischen beendet. Seit 2005 arbeitet die Stiftung zusätzlich an einem Index zu Sustainable Governance-Indikatoren (SGI), der den Reformbedarf und die Reformfähigkeit im Sinne der Regierungsleistung in 31 entwickelten Industriestaaten bewertet, die der OECD angehören.

Der BTI untersucht und bewertet Transformationsprozesse und politische Gestaltungsleistungen in inzwischen 128 Staaten und erstellt vergleichende Ranglisten von Entwicklungs- und Transformationsländern. Er besteht aus den Indikatorenkomplexen „Demokratie“ und „Marktwirtschaft“, die in den Status Index einfließen, sowie einem Management Index. In den Indikatorenkomplex Demokratie fließen fünf Kriterien ein: Staatlichkeit, politische Partizipation, Rechtsstaatlichkeit, Stabilität demokratischer Institutionen sowie politische und gesellschaftliche Integration. Die fünf Kriterien werden anhand von insgesamt 18 Fragen beurteilt, davon jeweils vier Fragen im Bereich Staatlichkeit, politische Partizipation, Rechtsstaatlichkeit, politische und gesellschaftliche Integration sowie zwei Fragen für das Kriterium Stabilität demokratischer Institutionen. In den Indikatorenkomplex Marktwirtschaft fließen sieben Kriterien ein, denen vierzehn Fragen zugeordnet sind.

9 Im Geschäftsjahr 2010 betrug der Etat der Bertelsmann Stiftung 60,3 Mio. Euro. Ende des Jahres 2010 beschäftigte die Stiftung 316 Mitarbeiter.

Diese sieben Kriterien sind sozioökonomisches Entwicklungsniveau, Markt- und Wettbewerbsordnung, Währungs- und Preisstabilität, Privateigentum, Sozialordnung, Leistungsstärke der Volkswirtschaft und Nachhaltigkeit.

Neben dem Status Index wird ein Management Index erstellt, der die Qualität der Steuerungsleistung politischer Entscheidungsträger misst. Dieser wird anhand von vier Kriterien, denen insgesamt 14 Fragen zugeordnet sind, ermittelt. Bei den Kriterien handelt es sich um Gestaltungsfähigkeit, Ressourceneffizienz, Konsensbildung und internationale Zusammenarbeit.

Darüber hinaus wird der Schwierigkeitsgrad politischer Steuerungsleistungen ermittelt, dem sechs Fragen zugrunde liegen. Die politischen Steuerungsleistungen sollen umso höher bewertet werden, je schwieriger die Ausgangsbedingungen und je geringer die verfügbaren Ressourcen sind (Bertelsmann 2010 a: 219).

Die Prüfsystematik kumuliert im Transformation Index. Ländergutachter beschreiben und bewerten den Entwicklungsstand für 128 Länder anhand eines standardisierten Codebuchs danach, inwieweit die 17 Kriterien, denen 52 Indikatoren zugeordnet sind, erreicht werden. Ein zweiter Gutachter, laut Bertelsmann Stiftung meist aus dem betreffenden Land, überprüft die Darstellungen und Bewertungen. Anschließend wird die Stimmigkeit im regionalen und internationalen Vergleich vom BTI Board,¹⁰ das laut Bertelsmann aus renommierten Wissenschaftlern besteht, überprüft und verabschiedet. Die Ländergutachter werden nicht genannt; die Mitglieder des 28-köpfigen Boards werden aufgeführt, darunter viele amtierende und zwei emeritierte Universitätsprofessoren.

Da der BTI sukzessive weiterentwickelt wurde, sind die Ergebnisse aus den Jahren 2003 mit denen ab den Jahren 2006 nur eingeschränkt vergleichbar. Mit der Veröffentlichung des BTI 2012 im März dieses Jahres hat die Bertelsmann Stiftung eine eigene Homepage für das Projekt eingerichtet (www.bti-project.de in deutscher und www.bti-project.org in englischer Sprache). Die Darstellung der Ergebnisse wurde nochmals überarbeitet; die Länderberichte sind nun in der Regel etwas ausführlicher.¹¹

10 Beim BTI Board handelt es sich nicht um einen „Vorstand“ mit autoritativen Entscheidungs- oder direkten Leitungsbefugnissen, sondern um einen Beirat, der Beratungs- und Überprüfungsfunktionen erfüllt, strategische Projektentscheidungen trifft und die BTI-Bewertungen beschließt.

11 Die BTI-Länderberichte zu Brasilien (32 Seiten), China (29 Seiten) und Indien (25 Seiten) enthalten beispielsweise im Jahr 2012 eine bis drei Seiten mehr als die jeweiligen Berichte im Jahr 2010 (Bertelsmann Stiftung 2012).

4. Der Freedom House Index (FHI)

Der Freedom House Index bzw. der Freedom of the World Survey von Freedom House, einer Forschungseinrichtung mit Sitz in Washington, liefert eine jährliche Bewertung und Analyse von Fort- und Rückschritten in Staaten und Gebieten weltweit in den Kategorien politische Rechte und bürgerliche Freiheiten. Die Bewertungen beziehen sich in erster Linie auf die Rechte und Freiheiten von Individuen und nicht direkt auf die Qualität von Regierungshandeln.¹² Freiheiten können durch Staaten, aber auch durch nichtstaatliche Akteure, etwa Aufständische oder bewaffnete Gruppen, begrenzt werden. Freedom House konstatiert daher, dass das Zusammenspiel von verschiedenen staatlichen und nichtstaatlichen Akteuren entscheidend für den ermittelten Freiheitsgrad sei. Die Analysen werden von internen Mitarbeitern der Organisation sowie von externen Wissenschaftlern und Experten erstellt.

Freedom House begann seine vergleichende Bewertung von Staaten in den fünfziger Jahren mit den sogenannten Balance Sheets of Freedom. Auf Grundlage einer neuen Methode, die maßgeblich von Raymond Gastil entwickelt wurde, veröffentlichte Freedom House den ersten Index im Jahr 1973. Die Arbeit von Freedom House wird nach unterschiedlichen Angaben zu etwa 80 Prozent aus Haushaltsmitteln der US-Regierung und zusätzlich durch Zuwendungen verschiedener Stiftungen, darunter die Open Society Foundation, finanziert.¹³ Freedom House veröffentlicht analytische Berichte und numerische Bewertungen zu insgesamt 194 Staaten und 14 umstrittenen Gebieten.

Die Einstufungen im Bereich der politischen Rechte (*political rights*) basieren auf den drei Unterkategorien *electoral process*, *political pluralism and participation* und *functioning of government*. Die Einstufungen im Bereich bürgerliche Freiheiten (*civil liberties*) fußen auf vier Unterkategorien: *freedom of expression and belief*, *associational and organizational rights*, *rule of law* sowie *personal autonomy and individual rights*.

12 „The survey does not rate governments or government performance per se, but rather the real-world rights and freedoms enjoyed by individuals [...]“ (Freedom House, Methodology, 2010 Edition, http://www.freedomhouse.org/template.cfm?page=351&ana_page=363&year=2010; Stand: 29.3.2012).

13 Entsprechende Hinweise finden sich im Internet auf der Homepage von Freedom House (www.freedomhouse.org) sowie auf der englischsprachigen Wikipedia-Seite http://en.wikipedia.org/wiki/Freedom_House: „Freedom House receives the majority (80%) of its funding from the U.S. government through the National Endowment for Democracy (NED), USAID, and the State Department“ (Stand: 28.2.2012).

Jedes Land bzw. Gebiet wird in den beiden Kategorien politische Rechte und bürgerliche Freiheiten auf einer Skala von eins bis sieben bewertet. Eins, vollständig frei, ist die Höchstnote, sieben, vollständig unfrei, die niedrigste Bewertung. Die numerische Bewertung auf der Skala von eins bis sieben ergibt sich aus der Anzahl von Punkten (bis zu 100), die jedes Land auf insgesamt 25 Fragen erhält – zehn Fragen zu politischen Rechten und fünfzehn Fragen zu bürgerlichen Freiheiten, die jeweils mit null bis vier Punkten bewertet werden. Null ist die niedrigste, vier die höchste Bewertung. Im Ergebnis wird ein Staat oder ein Gebiet dann als Wahl-demokratie bezeichnet, wenn er bzw. es mindestens sieben von zwölf Punkten im Hinblick auf die Fragen zu politischen Rechten und mindestens 20 von 40 Punkten hinsichtlich der Fragen zu bürgerlichen Freiheiten erhält. Darüber hinaus werden auch Trendpfeile vergeben, die generell positive oder negative Trends im Laufe des Jahres beschreiben.

Freedom House veröffentlicht seit 2006 die Werte für die Subkategorien und bereits seit 2003 die aggregierten Bewertungen. Die Gutachter und Einzelheiten zum Bewertungsprozess können inzwischen auf der Internetseite von Freedom House eingesehen werden.¹⁴ Die Methode und die Kriterien werden in regelmäßigen Abständen überprüft und geringfügig angepasst. Die Vergleichbarkeit zwischen den Jahren soll dabei gewahrt bleiben.¹⁵

5. Ergebnisse der Expertenbefragung

Mit der Expertenbefragung, bestehend aus einem Fragebogen und begleitenden Interviews, wurde der Anspruch verfolgt, die Nutzung und Wertschätzung der Indizes zu ermitteln und das Spektrum an Zustimmung, Kritik und Empfehlungen auszu-leuchten. Die Expertenbefragung unterscheidet zwischen den Kategorien Senior Researcher, Junior Researcher, Praktiker, Student/in und Grenzgänger/in. Promovierte Wissenschaftler über 35 Jahre oder solche, die bereits Junior-Professor sind, wurden als Senior Researcher eingestuft, jüngere Wissenschaftler mit Promotion oder in der Promotionsphase als Junior Researcher. Als Praktiker wurden Experten

14 Siehe zum Beispiel zu Indien http://www.freedomhouse.org/report/freedom-world-2010/survey-team?page=351&ana_page=366&year=2010 (Stand: 29.3.2012).

15 „The survey’s methodology is reviewed periodically by an advisory committee of political scientists with expertise in methodological issues. Over the years, the committee has made a number of modest methodological changes to adapt to evolving ideas about political rights and civil liberties. At the same time, the time series data are not revised retroactively, and any changes to the methodology are introduced incrementally in order to ensure the comparability of the ratings from year to year.” (<http://www.freedomhouse.org/report/freedom-world-2011/methodology>; Stand: 19.4.2012).

angesehen, die bei internationalen Organisationen, EU-Kommission, Bundestag, Ministerien, entwicklungspolitischen Organisationen, Stiftungen oder in der Beratung tätig sind. In der Kategorie Studenten wurden Bachelor- und Masterstudenten berücksichtigt, die Politikwissenschaft mit Schwerpunkten in den Bereichen Internationale Beziehungen oder Regionalstudien studieren. Mit Grenzgängern sind Experten gemeint, die sowohl in der Wissenschaft als auch in der Praxis zu verorten sind. Das Geschlecht der Befragten wurde ebenfalls als unabhängige Variable erfasst, um Aussagen über die Nutzung, die allgemeine und spezifische Wertschätzung und über verschiedene Dimensionen der Stärken, Schwächen sowie Empfehlungen zu den Indizes zu erfassen. Die Fragen zu Stärken, Schwächen und Empfehlungen waren offene Fragen, die mit einem Hinweis auf mögliche Antwortbereiche, zum Beispiel Konzept, Methode oder Darstellung, ergänzt wurden.

Der Autor und sein Team konnten zunächst vermuten, dass erfahrene Nutzer und solche, die die Indizes besonders wertschätzen, eine größere Bereitschaft für Antworten und Interviews zeigen würden. Um die Repräsentativität der Stichprobe zu erhöhen, wurde versucht, gegenüber den potenziellen Nutzern deutlich zu kommunizieren, dass eine Antwort unabhängig von Nutzung und Wertschätzung erwünscht sei. Die Ergebnisse sprechen gewissermaßen dafür, dass die Bemühungen des Autors und seines Teams erfolgreich waren, denn die Ergebnisse präsentieren – speziell unter Wissenschaftlern – ein breites Spektrum an Nutzungsverhalten und Wertschätzung.

a) Nutzung der Indizes

Bei der Nutzung der zwei näher untersuchten Indizes ergeben sich für den BTI und den FHI jeweils ähnliche Werte. Die Mittelwerte liegen jeweils knapp unterhalb der „seltenen Nutzung der Indizes“. Beim Nutzungsverhalten zeigen die verfügbaren Antwortdaten keine signifikanten Unterschiede zwischen der Nutzung der numerischen Bewertungen und der Länderberichte. Es wurden jedoch einige Aussagen zum BTI getroffen, wonach speziell die Länderberichte wertgeschätzt werden. Nicht vermutet wurde, dass der FHI von Senior Researchers in Deutschland etwas häufiger genutzt wird als der BTI. Dies entspricht nicht dem vom Autor und seinem Team im wissenschaftlichen Umfeld und der beruflichen Praxis beobachteten Trend hin zu komplexeren Indizes. In den Interviews wurde deutlich, dass Wissenschaftler, bei deren Arbeit die Auswertung von Daten eine bedeutende Rolle spielt, dazu neigen, verschiedene Indizes und eigene Datensätze zu kombinieren, weil sie sich die Zusammenstellung und Gewichtung unterschiedlicher Kriterien im Zuge von Erkenntnisprozessen nicht durch einen umfassenden Index, wie etwa den BTI, ab-

nehmen lassen wollen. Studierende und Grenzgänger nutzen dagegen häufiger den BTI als den FHI; Ersterer hat bei Studierenden offensichtlich eine gewisse Verbreitung gefunden. Nach einigen Rückmeldungen und Interviews zu urteilen, kann davon ausgegangen werden, dass die ausführlichen Länderberichte des BTI bei Studierenden beliebt sind. Allerdings können aus der kleinen Stichprobe bei diesen beiden Kategorien kaum belastbare Schlussfolgerungen gezogen werden.

Tabelle 1: Nutzung Bertelsmann Transformation Index (BTI) und Freedom House Index (FHI)

Kategorie		Wie oft nutzen Sie den BTI?	Wie oft nutzen Sie den FHI?
Senior Researcher	Mittelwert	1,96	2,16
	N	25	25
	Standardabweichung	,889	,800
Praktiker	Mittelwert	1,91	1,96
	N	23	23
	Standardabweichung	,596	,638
Junior Researcher	Mittelwert	1,77	2,08
	N	13	13
	Standardabweichung	,725	,862
Student/in	Mittelwert	1,58	1,50
	N	12	12
	Standardabweichung	,515	,522
Grenzgänger/in	Mittelwert	2,11	1,78
	N	9	9
	Standardabweichung	,782	,667
Insgesamt	Mittelwert	1,88	1,95
	N	82	82
	Standardabweichung	,727	,735

Werte: 1 = (noch) nicht, 2 = selten, 3 = häufig.

Die Unterschiede zwischen den Geschlechtern sind bei der Nutzung der Indizes wie erwartet gering ausgeprägt. Die etwas häufigere Nutzung der Indizes von männlichen Befragten ist beim FHI geringfügig stärker ausgeprägt.¹⁶

Wissenschaftler arbeiten außerdem vor allem mit dem Polity IV-Index. Der Vanhanen Index und der Ease of doing Business Index der Weltbank¹⁷ erhielten ebenso einige Nennungen. Praktiker erwähnten vor allem den Human Development Index

16 Der Mittelwert der Nutzung liegt für Männer beim BTI bei 1,92, bei Frauen bei 1,82; beim FHI liegt der Mittelwert bei Männern bei 2,02 und bei Frauen bei 1,85. Die Werte sind folgendermaßen zugeordnet: 1 = (noch) nicht, 2 = selten, 3 = häufig; siehe auch Tabelle 1.

17 <http://www.doingbusiness.org> (Stand: 12.4.2012).

der Vereinten Nationen¹⁸ und den Corruption Perception Index von Transparency International¹⁹ als weitere Indizes, die von ihnen genutzt werden.

Beim Zweck der Nutzung ist das gesamte Spektrum von Forschung, Lehre, Vorträgen, Präsentationen und Beratung vertreten. Viele Befragte gaben Mehrfachnennungen ab. Viele Praktiker benutzen die Indizes als Einstieg und/oder als grobe Orientierung bei der Beschäftigung mit Staaten, die sie weniger gut kennen, zum Beispiel im Kontext neuer beruflicher Arbeitsfelder, Zuständigkeiten, anstehender Dienstreisen oder der Erstellung von Gutachten. Die Indizes werden auch von einigen Senior Researchers gerne zusammen mit anderen Indizes in der Lehre verwendet. Entsprechende Äußerungen wurden unter anderem von Aurel Croissant, Claudia Derichs, Andrea Fleschenberg, Wolfgang Merkel, Ingo Peters, Susanne Pickel und Uwe Wagschal abgegeben.²⁰

b) Allgemeine und spezifische Wertschätzung der Indizes

Hinsichtlich der Frage nach der allgemeinen Wertschätzung der Indizes für die Einstufung und Vergleichbarkeit von Transformationsprozessen (siehe Tab. 2) wurde das gesamte Spektrum der Skala von 0 (keine Antwort) bzw. 1 (sehr gering) bis 5 (sehr hoch) ausgeschöpft. Der BTI wird insgesamt mehr wertgeschätzt als der FHI, wobei sich beim BTI praktisch keine Unterschiede zwischen den Expertenkategorien ergeben. Die weiblichen Befragten schätzen den BTI etwas mehr als die männlichen Befragten.

Erfahrene Nutzer der Indizes, überwiegend Senior Researchers,²¹ gaben trotz zahlreicher Kritikpunkte insgesamt eine mittlere bis hohe Wertschätzung für die Indizes an. Starke Kritik kam unter anderem von einigen Befragten, die selbst nur geringe Kenntnisse über die Indizes hatten.²² Die Praktiker gaben bei Wertschätzungen überwiegend mittlere Werte an; nur wenige gaben an, über vertiefte Kenntnisse in Bezug auf die Indizes zu verfügen.

18 <http://hdr.undp.org/en/statistics/hdi> (Stand: 12.4.2012).

19 <http://cpi.transparency.org/cpi2011> (Stand: 12.4.2012).

20 Wolfgang Merkel und Uwe Wagschal sind an der Erstellung des BTI beteiligt.

21 Darunter Claudia Derichs, Andrea Fleschenberg, Harald Fuhr und Susanne Pickel.

22 Der Autor möchte aus Rücksicht auf die befragten Personen hier auf die Nennung von Namen verzichten. Einige Befragte erteilten auch explizit nicht die Erlaubnis zum Zitieren.

Table 2: Allgemeine Wertschätzung von BTI und FHI

Kategorie		Wie hoch ist Ihre allgemeine Wertschätzung des BTI?	Wie hoch ist Ihre allgemeine Wertschätzung des FHI?
Senior Researcher	Mittelwert	3,43	2,90
	N	21	21
	Standardabweichung	1,121	,944
Praktiker	Mittelwert	3,35	3,11
	N	17	19
	Standardabweichung	,786	,809
Junior Researcher	Mittelwert	3,50	2,90
	N	10	10
	Standardabweichung	1,080	,738
Student/in	Mittelwert	3,63	2,56
	N	8	9
	Standardabweichung	,744	1,014
Grenzgänger/in	Mittelwert	3,57	2,83
	N	7	6
	Standardabweichung	,787	,753
Insgesamt	Mittelwert	3,46	2,91
	N	63	65
	Standardabweichung	,930	,861

Werte: 1 = sehr gering, 2 = gering, 3 = mittel, 4 = hoch, 5 = sehr hoch.

Bei der spezifischen Wertschätzung der Qualität der Indizes, die mit Fragen zur Vergleichbarkeit von Ländern und Zeiträumen, zu den Bewertungen und Analysen sowie zu den Länderberichten abgefragt wurde, ist ebenso das gesamte Spektrum von 0 (keine Antwort) bzw. 1 (trifft nicht zu) über 2 (trifft teilweise zu) bis zu 3 (trifft zu) vertreten. Nur wenige Antworten differenzieren deutlich bei der spezifischen Wertschätzung. Unter den Wissenschaftlern unterscheiden sich die Werte stärker als bei den Praktikern, den Studierenden und den Grenzgängern. Speziell beim BTI ist die Standardabweichung bei den befragten Senior Researchers (1,121) deutlich höher als bei den Praktikern (0,786). Die Kritik der Wissenschaftler an Konzeption und Methode fällt differenziert aus: Für Praktiker ist die Aktualität der Indizes wichtiger als für Wissenschaftler, die eher an einer Vergleichbarkeit über lange Zeitreihen interessiert sind.

c) Beurteilung des Wertes der Indizes für Entscheidungsträger

Hinsichtlich des Wertes der Indizes für Entscheidungsträger bewegen sich die Einschätzungen auf einer Skala von 1 bis 5 überwiegend im Mittelfeld. Der BTI schneidet mit einem Mittelwert von 3,14 leicht besser ab als der FHI mit einem Mittelwert von 2,84 (vgl. Tab 3).

Tabelle 3: Einschätzung des Wertes von BTI und FHI für Entscheidungsträger

Kategorie		Wie hoch schätzen Sie den Wert des BTI für Entscheidungsträger ein?	Wie hoch schätzen Sie den Wert des FHI für Entscheidungsträger ein?
Senior Researcher	Mittelwert	3,26	2,82
	N	19	19
	Standardabweichung	1,046	,960
Praktiker	Mittelwert	3,05	2,82
	N	19	19
	Standardabweichung	,970	,989
Junior Researcher	Mittelwert	2,88	2,60
	N	8	10
	Standardabweichung	1,126	,699
Student/in	Mittelwert	3,40	3,25
	N	10	8
	Standardabweichung	,699	,886
Grenzgänger/in	Mittelwert	3,00	2,86
	N	7	7
	Standardabweichung	,816	,690
Insgesamt	Mittelwert	3,14	2,84
	N	63	63
	Standardabweichung	,948	,888

Werte: 1 = sehr gering, 2 = gering, 3 = mittel, 4 = hoch, 5 = sehr hoch.

Im Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung (BMZ) befasst man sich nur im Governance-Referat genauer mit den Indizes. Direkte Empfehlungen für die Zusammenarbeit mit Partnerländern werden von den Regionalreferaten daraus nicht abgeleitet. Die Erstellung länderspezifischer polit-ökonomischer Kurzanalysen, in denen in der Regel auf die Indizes Bezug genommen wird, erfolgt im Auftrag des BMZ durch externe Gutachter. Für Staaten mit schwierigen Rahmenbedingungen, in denen keine oder nur eine begrenzte Zusammenarbeit stattfindet, wird laut einigen Praktikern eher auf Indizes zurückgegriffen als bei Staaten, mit denen bereits eine Zusammenarbeit existiert, auf deren Grundlage dann spezifische Einschätzungen vorgenommen werden können. Im Einzelnen wurde auch auf andere Aspekte eingegangen. So äußerte sich Helmut Reisen zum Wert der Indizes für Entscheidungsträger wie folgt:

„Regierungsqualität und insbesondere Korruption können nie direkt vermessen werden. Es werden also in jedem Fall Proxys gebraucht, die auf subjektiven Urteilen und Meinungen aufbauen (sowohl facts-based als auch perception-based Indikatoren). Insbesondere kleine Unterschiede in Hunderten von Ländern, entweder über Zeit oder zwischen Ländern, sind somit nicht vertrauenswürdig. Die Kausalität kann nie eindeutig durch simple Regressionen nach-

gewiesen werden, insbesondere wenn der Fehlerterm des Indikators nicht zufällig ist, was mit großer Wahrscheinlichkeit der Fall ist.“²³

Marius Eckert, Student, führte an, dass der Corruption Perception Index für Entscheidungsträger vermutlich einen höheren Wert als der BTI und der FHI habe, da er sich klar und nuanciert dem Teilproblem der Korruption annehme und nicht wie BTI und FHI versuche, sehr komplexe und normative Themen wie Demokratie und Freiheit in letztendlich eine einzige Kennzahl zu verpacken.²⁴

d) Das Spektrum der genannten Stärken

Die Fragen nach den Stärken und Schwächen wurden als offene Fragen formuliert. Dabei wurden beispielhaft Bereiche – darunter Konzept, Methode und Darstellung – angegeben, die den Befragten als Leitlinien dienen sollten (vgl. Tab. 4). Einige Antworten griffen speziell das Thema der Relevanz und Eignung für den Politikdialog auf, weil dies ausdrückliche Anliegen sowohl der Bertelsmann Stiftung als auch von Freedom House sind.

Das Vorhaben der Erstellung von Indizes wurde von vielen grundsätzlich wertgeschätzt. Die Formulierung von Markus Karbaum, Junior Researcher, bringt die persönliche Einschätzung zu den Stärken der Indizes auf den Punkt: „Man kann sich Überblickswissen verschaffen. Zeitreihen-, Länder- und Regionalvergleiche bieten sich an. Globale Trends sind erkennbar.“²⁵

Die amerikanische Nichtregierungsorganisation Freedom House und die Bertelsmann Stiftung genießen als Institutionen bei vielen Befragten einen überwiegend seriösen Ruf. Alexander Pieske (Deutscher Bundestag, CDU/CSU-Fraktion, als Praktiker eingestuft) hebt ausdrücklich das „Renommee der dahinter stehenden Institutionen“ hervor.²⁶ Hans Jörg Schmidt, Junior Researcher, John Stuart Mill Institute, betont, dass „renommierte Gutachter und kompetente Regional-Koordina-

23 Reisen, Helmut (mit Unterstützung von Christina Wolf, OECD), ausgefüllter Fragebogen vom 20.6.2011, der nur persönliche Meinungen wiedergibt. Helmut Reisen ist Head of Research des OECD Development Centre und Professor für Volkswirtschaft an der Universität Basel. Er wurde als Grenzgänger eingestuft.

24 Eckert, Marius, ausgefüllter Fragebogen vom 30.6.2011.

25 Karbaum, Markus, ausgefüllter Fragebogen vom 12.6.2011. Markus Karbaum ist promovierter Politikwissenschaftler mit Schwerpunkt Südostasien. Er wurde als Junior Researcher eingestuft.

26 Pieske, Alexander, ausgefüllter Fragebogen vom 27.6.2011. Alexander Pieske arbeitet als Referent der CDU/CSU-Fraktion im Deutschen Bundestag. Die Antwort gibt seine persönliche Meinung wieder.

toren“ beteiligt seien. Schmidt hebt ebenso hervor, dass die Länderberichte ein „inhaltslich gelungenes feines Raster verwenden, das Differenzierungen erlaubt“.²⁷

Häufig wurde geantwortet, dass die Indizes Basis- und Überblickswissen in ansprechender Darstellung vermitteln und einen guten Einstieg in die politische Analyse von Staaten anbieten. Beim BTI wurden die umfangreichen Kriterien und die ausführlicheren Länderberichte positiv erwähnt; beim FHI wurden die längeren Zeitreihen und die jährliche Veröffentlichung geschätzt, die Vergleiche über zurückliegende Jahre möglich machen.

Aurel Croissant, Senior Researcher, der selbst als Ländergutachter an der Erarbeitung des BTI mitwirkt, bescheinigt diesem eine „hohe Konzeptvalidität“;²⁸ Andrea Fleschenberg, Senior Researcher, schätzt das umfassende Konzept des BTI.²⁹ David Kühn, Junior Researcher, hebt die „klare Rückbindung an demokratietheoretische Überlegungen, die stringent formulierten Kriterien und die Schwellenwerte der Kriterien“ hervor. Kühn schätzt auch die „breite Darstellung des Transformationsstands, speziell den expliziten Einbezug von sozio-ökonomischen Aspekten, die Abdeckung des Prozess-Charakters der Transformation durch den Management-Index und den dreistufigen anonymisierten Peer Review Prozess bei der Anfertigung der Länderstudien“.³⁰

Sandra Destradi, Junior Researcher, antwortete zu den Stärken von BTI und FHI mit einem Vergleich der Indizes:

„Die Stärke des BTI liegt, im Vergleich zu Freedom House, in der breiteren Definition von Transformation und der Berücksichtigung der wirtschaftlichen Komponente. Zentral ist ebenfalls die Differenziertheit durch die Berücksichtigung der politischen Steuerungsleistung und dabei insbesondere des Schwierigkeitsgrads. Das Ranking des BTI erhöht zudem die Vergleichbarkeit. Für den Fall Indien hat der Länderbericht den Vorteil, dass er eine viel breitere Palette an Entwicklungen in Betracht zieht, z. B. durch die Berücksichtigung von Themen wie der Funktionsfähigkeit des Wohlfahrtsstaats, Inflation usw. Der Vorteil von Freedom House liegt in der Übersichtlichkeit der Einteilung in *free*, *partial free* und *not free*. Zudem werden die Länderberichte jährlich aktualisiert. Je nach Untersuchungsgegenstand und Fragestellung sind somit diese Indizes unterschiedlich geeignet.“³¹

27 Schmidt, Hans Jörg, ausgefüllter Fragebogen vom 8.6.2011 mit der Erlaubnis zum Zitieren. Hans Jörg Schmidt ist nicht in die Indizes-Projekte des Instituts persönlich involviert.

28 Croissant, Aurel, Interview für die Beantwortung des Fragebogens am 11.7.2011.

29 Fleschenberger, Andrea, Interview für die Beantwortung des Fragebogens am 22.6.2011.

30 Kühn, David, ausgefüllter Fragebogen vom 4.7.2011.

31 Destradi, Sandra, ausgefüllter Fragebogen vom 8.6.2011.

Ingo Peters, Senior Researcher, erwähnt zur Methodik beider Indizes positiv, dass die Kriterien nachvollziehbar operationalisiert seien, allerdings nicht von der Einzelfallbetrachtung befreien könnten.³² Subrata K. Mitra, Senior Researcher und ehemaliger Ländergutachter des BTI, bescheinigt diesem, dass die Darstellung gut zugänglich sei. Dies ermögliche eine übersichtliche und einfache Vergleichbarkeit.³³

Tabelle 4: Stärken des BTI und des FHI

		BTI Häufigkeit	BTI in %	FHI Häufigkeit	FHI in %
Gültig	Konzept	15	27,8	6	15,0
	Methode	5	9,3	10	25,0
	Darstellung	7	13,0	5	12,5
	Politik	4	7,4	4	10,0
	Multiple	23	42,6	15	37,5
	Gesamt	54	100,0	40	100,0
Fehlend	keine Antwort	28		42	

Anmerkung: Mit „Konzept“ sind im Verständnis der Autoren vor allem die Auswahl und Zusammenstellung der Kriterien und die damit verbundenen Werthaltungen gemeint. Fragen der Methodik beziehen sich vor allem auf die Verfahren der Ermittlung und Aggregation von Daten. Mit der Darstellung sind die sprachliche und die graphische Präsentation der Ergebnisse gemeint. Politik bezieht sich auf die Relevanz und Sinnhaftigkeit für den Politikdialog. Multiple Nennungen bedeutet, dass sich die Antworten auf mindestens zwei Bereiche beziehen.

Relativ wenig positive Bemerkungen finden sich zur Relevanz des BTI für den Politikdialog. Wolfgang Merkel, Senior Researcher und Mitglied im Board des BTI, hat die Entwicklung des BTI von Anfang an begleitet. Er führt an, dass das Ranking Diskussionen belebe, Rückmeldungen generiere und dazu führe, dass sich die Politik mit dem Index und seinen Inhalten befasse.³⁴

e) Das Spektrum der genannten Schwächen

Bei der Nennung von Schwächen wurde vor allem Kritik an den zugrundeliegenden Leitbildern der Indizes hinsichtlich Demokratie, Freiheit sowie des Verlaufs von Transformationsprozessen geübt (vgl. Tab. 5). Der BTI orientiere sich demnach stärker am Leitbild einer konsensorientierten und inklusiven Demokratie mit kor-

32 Peters, Ingo, ausgefüllter Fragebogen vom 10.7.2011.

33 Mitra, Subrata K., Interview für die Beantwortung des Fragebogens am 11.7.2011.

34 Merkel, Wolfgang, Interview für die Beantwortung des Fragebogens am 4.7.2011. Die Projektmanager des BTI sprechen auch davon, dass Bertelsmann zahlreiche Rückmeldungen zu dem Index aus der Politik bekommen hat, einschließlich Kommentare von hohen Regierungsvertretern (Kuhn 2011).

poratistischen Strukturen sowie einer sozial flankierten Marktwirtschaft. Die Verknüpfung von Bewertungen in den Bereichen Demokratie und Marktwirtschaft für die Ermittlung der Gesamtbewertung der Länder wurde von manchen Experten als problematisch angesehen, selbst wenn die jeweiligen Einstufungen für politische und wirtschaftliche Transformation separat ausgewiesen würden. Der FHI fokussiere dagegen zu sehr auf US-amerikanisch geprägte Leitbilder, die mehr auf bürgerliche und politische Freiheiten und Rechte als auf wirtschaftliche und soziale Rechte achteten. Diese normativen Leitbilder würden nicht ausreichend kritisch reflektiert, speziell bei Freedom House. Der FHI wurde von einigen Befragten zudem als „unterkomplex“ bezeichnet.

Von den Befragten kamen dazu unter anderem folgende Aussagen: Markus Lederer, Senior Researcher, führt an, dass Freedom House zu sehr an einen bestimmten Freiheitsbegriff gebunden sei. Die Darstellung sei zwar hilfreich, aber normativ mehr aufgeladen als notwendig.³⁵ Susanne Pickel, Senior Researcher, kritisiert beim BTI, dass die Kriterien nicht ausreichend trennscharf seien. Bei früheren Versionen des FHI kritisiert sie eine zu hohe Aggregation der Werte; die Bewertung sei zwar für die Subkategorien, nicht aber für einzelne Fragen einsehbar. Auch seien die Gutachter nicht bekannt.³⁶ Lars Brozus, Senior Researcher, moniert am BTI, dass Messungen „traditioneller Formen politischer Partizipation“ und die Erfassung von „traditionellen Formen von Governance“ kaum berücksichtigt werden.³⁷ Auch Claudia Derichs, Senior Researcher, sieht Defizite bei der Berücksichtigung von Partizipationsaspekten nichtwestlicher Spielart.³⁸ Sowohl sie als auch Andrea Fleschenberg, Senior Researcher, kritisieren den BTI im Hinblick auf geschlechterdemokratische Aspekte und sprechen von der „Gender-Blindheit des BTI“.³⁹

Marianne Braig, Senior Researcher, führte an, dass die Vorstellungen über Transformationsprozesse des BTI zu stark von einem modernisierungstheoretisch geprägten Evolutionsmodell geprägt seien. Europa habe sich nicht unbedingt nach den Kriterien des BTI entwickelt. Die Reflektion über multi-ethnische und multi-nationale Aspekte der Entwicklung von Gesellschaften komme zu kurz. Es ließe sich bei beiden Indizes ein Mangel an Erkenntnisinteresse mit Blick auf Pluralität und alternative Entwicklungswege beobachten.⁴⁰ Hans Jörg Schmidt, Junior Researcher,

35 Lederer, Markus, ausgefüllter Fragebogen vom 8.6.2011.

36 Pickel, Susanne, Interview für die Beantwortung des Fragebogens am 7.7.2011.

37 Brozus, Lars, Interview für die Beantwortung des Fragebogens am 11.7.2011.

38 Derichs, Claudia, Interview für die Beantwortung des Fragebogens am 24.6.2011.

39 Derichs, Claudia, Interview für die Beantwortung des Fragebogens am 24.6.2011, und Fleschenberg, Andrea, Interview für die Beantwortung des Fragebogens am 22.6.2011.

40 Braig, Marianne, Interview für die Beantwortung des Fragebogens am 7.7.2011.

merkte an, dass der Status der *electoral democracy* nicht zwangsläufig mit liberaler Demokratie kompatibel sei. Die subjektive Einstellung der Bevölkerung würde nicht erhoben und die ökonomische Komponente sei beim FHI kaum berücksichtigt. Die Stufenkategorisierung sei problematisch. Immerhin hätte sich die Vergleichbarkeit durch einen Leitfaden seit 2005 verbessert, Methodenänderungen seien aber nicht rückprojiziert worden. Insgesamt gab Schmidt jedoch eine mittlere Wertschätzung des FHI an.⁴¹

Kritik an der Methodik übten vor allem Wissenschaftler, speziell zur Trennschärfe der Kriterien, zur Subjektivität der Bewertungen und zur mangelnden Transparenz bei der Auswahl der Experten. Praktiker kritisierten mit Ausnahme des Aspektes der Transparenz die Methodik nicht. Bei den Schwächen der Indizes wurde auch das grundsätzliche Problem der Perzeptionsorientierung der Indikatoren thematisiert. Diese würden eigentlich eher das Problembewusstsein messen, was zur Folge habe, dass sich die Indikatoren nach Kampagnen oder Maßnahmen zur Abhilfe von Problemen verschlechtern würden, weil Defizite stärker gegenwärtig seien.⁴²

Grundlegende Kritik am BTI wie am FHI kam von Christoph Wilke von der Menschenrechtsorganisation Human Rights Watch:

„Bei Human Rights Watch setzen wir ausdrücklich darauf, keine Ländervergleiche über das Spektrum der Menschenrechte zu erstellen (A ist besser als B) oder dies in einem Punktesystem zu tun [...]. Unsere Ablehnung von Bewertungsskalen beruht auf folgenden Überlegungen: Skalen sind Futter für die Medien, deren Wirksamkeit verpufft (jedoch kurzfristig kann dies sehr nützlich sein); Skalen geben nicht unbedingt die notwendigen Impulse zu Reformen (es sei denn, es gibt z. B. eine wirksame parlamentarische Opposition). Wer gute Noten erhält, will sich nicht verbessern, wer schlechte Noten aus Amerika oder Deutschland erhält sieht sich gekränkt, und, wie oft der Fall, sieht sich wahrscheinlich nicht in einem regionalen Wettkampf um bessere Noten; wir versuchen, gleich an die Substanz einer Menschenrechtsthematik zu gehen und diese mit der Regierung über Monate und Jahre zu besprechen. Dazu sind auch die Medien und *name and shame* notwendig“.⁴³

41 Schmidt, Hans Jörg, ausgefüllter Fragebogen vom 8.6.2011.

42 Dieser Punkt wurde speziell von Harald Fuhr angesprochen. Fuhr, Harald, Interview für die Beantwortung des Fragebogens am 27.6.2011.

43 Wilke, Christoph, Textantworten vom 14. und 15. Juni 2011.

Tabelle 5: Schwächen des BTI und des FHI

		BTI Häufigkeit	BTI in %	FHI Häufigkeit	FHI in %
Gültig	Konzept	20	39,2	17	33,3
	Methode	11	21,6	14	27,5
	Darstellung	0	0,0	1	2,0
	Politik	2	3,9	2	3,9
	Multiple	18	35,3	17	33,3
	Gesamt	51	100	51	100
Fehlend	keine Antwort	31		31	

Anmerkungen: Siehe Tabelle 4.

f) Das Spektrum der genannten Empfehlungen

Im Hinblick auf Empfehlungen und Verbesserungspotenziale wünschen sich insbesondere Wissenschaftler beim BTI häufiger längere Zeitreihen und damit die Möglichkeit zur Rückbewertung der Staaten. Bei sich verändernden Kriterien sehen sie außerdem die Gefahr des Strukturbruchs. Praktiker sind mehr an einer Verfeinerung der Indikatoren und an Aktualität interessiert; die Betrachtung von Entwicklungen über Zeitreihen ist für sie von nachrangiger Bedeutung. Insgesamt äußerten sich nur wenige Praktiker zu Empfehlungen; ebenso äußerten sich Studierende kaum hierzu. Mehrere Experten wünschten sich die bessere Einbeziehung einheimischer Institutionen. Speziell mit Blick auf den FHI wurde empfohlen, die Darstellung weiter zu verbessern. Helmut Reisen, Grenzgänger, äußerte sich – unterstützt von seiner Kollegin Christina Wolf (OECD) – ausführlich zu Empfehlungen für die Indizes:

„Insbesondere in der Entwicklungszusammenarbeit haben wir es mit Ländern in verschiedenen Entwicklungsphasen zu tun. Um nun beispielsweise ihre Wachstumsperformance zu verstehen, macht es mehr Sinn, vor der empirischen Analyse die Länder in Gruppen wie „entwickelt“, „konvergierend“, „divergierend“ einzuteilen. Die meisten eindeutigen empirischen Zusammenhänge entstehen nämlich nur, wenn alle Länder gepoolt werden (reich, konvergierend und divergierend). Eine Analyse nach Gruppen bringt oft verblüffende Ergebnisse ans Tageslicht. Ebenfalls empfiehlt sich in jeder Analyse, auf mehrere Indikatoren zurückzugreifen und insbesondere sogenannte *per-*

ception based-Indikatoren und sogenannte *facts based*-Indikatoren zu kombinieren.“⁴⁴ ⁴⁵

Eine andere Grenzgängerin, Léonie Wagner von der Universität Frankfurt und der Gesellschaft für Internationale Zusammenarbeit (GIZ), merkte an, dass der BTI auch Industrieländer mit berücksichtigen solle, um anzuerkennen, dass die Qualität von staatlichen Transformationsleistungen in jedem Länderkontext von Bedeutung ist und die klare Abgrenzung zwischen Entwicklungsländern einerseits und Industrieländern andererseits aus verschiedenen Gründen zusehends wenig zeitgemäß erscheint.⁴⁶ Dieser Aspekt wurde in einigen Gesprächen vertieft.⁴⁷ Als Argumente für die Berücksichtigung von sogenannten konsolidierten Demokratien wurde angeführt, dass sich Staaten nicht nur in eine Richtung entwickelten, sondern in einigen OECD-Staaten auch Rückschritte bei Rechtsstaatlichkeit, Freiheitsrechten und Steuerungsleistungen beobachtet werden könnten. Als Gegenargument wurde angeführt, dass eine Berücksichtigung aller Staaten eine erhebliche Ausweitung der Skala zur Folge haben müsste, damit noch ausreichend zwischen diesen differenziert werden kann. Dies sei aber problematisch, denn ein noch höherer Grad an Ausdifferenzierung würde dem Nutzer eine Genauigkeit vorspiegeln, die perzeptionsbasierte Indizes nicht vorspielen sollten.⁴⁸

6. Fazit

Beide Indizes – Bertelsmann Transformation Index und Freedom House Index – können die unterschiedlichen Bedürfnisse von Wissenschaft und Praxis schwerlich gleichzeitig erfüllen. Wissenschaftler äußern sich differenzierter und kritischer zur Konzeption und Methodik der Indizes. Sie schätzen vergleichende Daten über lange Zeitreihen. Praktiker schätzen dagegen besonders Aktualität und sind eher an der Aufnahme neuer Indikatoren interessiert. Aus der Perspektive der Wissenschaft stellt dies jedoch die Vergleichbarkeit über Zeitreihen in Frage. Die Indizes bleiben für viele Experten eine Einstiegs- und Überblickslektüre, mit der sie sich durchschnittlich eher selten befassen.

Die Befragung unter deutschsprachigen potenziellen Nutzern der Indizes zeigt, dass der BTI den FHI bei der Nutzung offenbar nicht eingeholt hat. Unterschiede

44 Reisen, Helmut mit Christina Wolf, ausgefüllter Fragebogen vom 20.6.2011.

45 Das World Justice Project (Agrast/Botero/Ponce 2011) berücksichtigt den Aspekt der Gruppierungen von Staaten nach Entwicklungsniveaus beim Rule of Law Index.

46 Wagner, Léonie, ausgefüllter Fragebogen vom 21.6.2011.

47 Siehe dazu Stoiber 2011 zur Berücksichtigung des Kontextes in der Demokratiemessung.

48 Merkel, Wolfgang, Interview und ausgefüllter Fragebogen vom 4.7.2011.

zwischen den Geschlechtern sind nicht sehr signifikant. Die Wertschätzung der Indizes liegt insgesamt im Mittelfeld der Skala, der BTI wird jedoch durchschnittlich etwas mehr wertgeschätzt als der FHI. Dem BTI wird auch eine höhere Relevanz für Entscheidungsträger zugewiesen. Bei den Senior Researchers gibt es die größten Unterschiede in der Wertschätzung: Je informierter die Befragten sind, desto vielseitiger und differenzierter fallen Wertschätzung, Kritik und Empfehlungen aus. Die Wertschätzung der erfahrensten Nutzer, die die Indizes pragmatisch als Einstiegslektüre nutzen, bewegt sich im Mittelfeld.

Einigkeit herrschte bei den meisten Experten darüber, dass die Indizes ein gutes Einstiegswissen vermitteln. Es wurde bei beiden Indizes Handlungsbedarf dahingehend gesehen, regionale und nationale Institutionen stärker zu beteiligen, auch im Hinblick auf eine weitere Verbreitung und verbesserte Akzeptanz der Indizes.⁴⁹

Literatur

- Agrast, Mark David/Botero, Juan Carlos/Ponce, Alejandro*, 2011: Rule of Law Index, World Justice Project, <http://www.worldjusticeproject.org/rule-of-law-index/> (Stand: 1.9.2011)
- Ahrens, Joachim*, 2010: Evaluation des Transformation Index (BTI). Dimension Status-Index: Wirtschaftliche Transformation; beauftragt von der Bertelsmann Stiftung, Gütersloh.
- Alvarez, Michael/Cheibub, José Antonio/Limongi, Fernando/Przeworski, Adam*, 1996; Classifying political regimes, in: *Studies in Comparative International Development* 31 (2), 1-37.
- Arndt, Christiane/Oman, Charles*, 2006: Uses and Abuses of Governance Indicators, OECD Development Center Study, Paris, <http://www.oecd.org/dataoecd/21/16/40037762.pdf> (Stand: 10.4.2012).
- Beetham, David*, 1994: Defining and measuring Democracy, London u. a.
- Berg-Schlosser, Dirk*, 2007: Concepts, Measurement and Sub-Types in Democratization Research, in: ders. (Hrsg.), *Democratization. The State of the Art*, Opfaden, 31-43.
- Bertelsmann Stiftung (Hrsg.)*, 2010 a: Transformation Index. Political Management in International Comparison, Gütersloh.

49 Ein Beispiel für die maßgebliche Beteiligung von einheimischen Institutionen ist beispielsweise der Civil Society Index von CIVICUS (2011). Darunter leidet allerdings die Vergleichbarkeit, da eine zeitliche Abstimmung zwischen den vielen verschiedenen einheimischen Institutionen kaum gelingt.

- Bertelsmann Stiftung (Hrsg.)*, 2010b: Transformation Index. Politische Gestaltung im internationalen Vergleich, Gütersloh, http://www.bertelsmann-stiftung.de/bst/de/media/xcms_bst_dms_30284_30286_2.pdf (Stand: 1.9.2011).
- Bertelsmann Stiftung (Hrsg.)*, 2012: Transformation Index BTI 2012, <http://www.bti-project.org> und <http://www.bti-project.de> (Stand: 30.3.2012).
- Bollen, Kenneth A.*, 1986: Political rights and political liberties in nations: An evaluation of human rights measures 1950 to 1984, in: *Human Rights Quarterly* 8 (4), 567-591.
- Bühlmann, Marc/Hänni, Miriam/Zumbach, David*, 2010: Evaluation des Bertelsmann Transformation Index, Zentrum für Demokratie Aarau (im Auftrag der Bertelsmann Stiftung).
- Bühlmann, Marc/Merkel, Wolfgang/Müller, Lisa/Giebler, Heiko/Weßels, Bernhard*, 2012: Demokratiebarometer. Ein neues Instrument zur Messung von Demokratiequalität, in: *Zeitschrift für Vergleichende Politikwissenschaft* (i. E.).
- Bühlmann, Marc/Merkel, Wolfgang/Müller, Lisa/Weßels, Bernhard*, 2008: Wie lässt sich Demokratie am besten messen? Zum Forumsbeitrag von Thomas Müller und Susanne Pickel, in: *Politische Vierteljahresschrift* 49 (1), 114-122.
- Civicus*, 2011: Civil Society Index, Johannesburg, <http://csi.civicus.org> (Stand: 29.3.2012).
- Coppedge, Michael/Gerring, John u. a.*, 2011: Conceptualizing and Measuring Democracy. A New Approach, in: *Perspectives on Politics* 9 (2), 247-267.
- Croissant, Aurel/Thiery, Peter*, 2000: Defekte Demokratie. Konzept, Operationalisierung und Messung, in: Hans-Joachim Lauth/Gert Pickel/Christian Wetzel (Hrsg.), *Demokratiemessung*, Wiesbaden, 112-132.
- Dahl, Robert A.*, 1971: *Polyarchie. Participation and Opposition*. New Haven/London.
- Foreign Policy*, 2011: Failed State Index, <http://www.foreignpolicy.com/failed-states> (Stand: 1.9.2011).
- Fukuyama, Francis*, 1992: *The End of History and the Last Man*, New York.
- Gastil, Raymond*, 1991: The comparative survey of freedom: Experiences and suggestions, in: Alex Inkeles (Hrsg.), *On measuring democracy: Its consequences and concomitants*, New Brunswick, 21-46.
- Gwartney, James/Hall, Joshua/Lawson, Robert*, 2010: Economic Freedom of the World. 2010 Annual Report, Fraser Institute, Vancouver (<http://www.amfot.tj/tj/publication/EconFreeWorld2010.pdf>; Stand: 29.3.2011).
- Jagers, Keith/Gurr, Ted Robert*, 1995: Tracking Democracy's Third Wave with the Polity III Data, in: *Journal of Peace Research* 32 (4), 469-482.

- Kaufmann, Daniel/Kraay, Aart/Zoido-Lobaton, Pablo*, 1999: Aggregating Governance Indicator, World Bank Policy Research Working Paper 2195, October 1999, Washington D.C.
- Knack, Steven*, 2006: Measuring Corruption in Eastern Europe and Central Asia: A Critique of the Cross-Country Indicators. World Bank Policy Research Working Paper 3968, Washington D.C.
- Kuhn, Berthold*, 2011: Gesprächsprotokoll des Treffens mit Sabine Donner und Hauke Hartmann von der Bertelsmann Stiftung in Berlin am 30.6.2011 (unveröffentlicht).
- Lauth, Hans-Joachim*, 2004: Demokratie und Demokratiemessung, Wiesbaden.
- Merkel, Wolfgang/Puhle, Hans-Jürgen/Croissant, Aurel/Eicher, Claudia/Thiery, Peter*, 2003: Defekte Demokratie. Band 1: Theorie, Opladen.
- Mo Ibrahim Foundation*, 2011: Mo Ibrahim Foundation Index, <http://www.moibrahimfoundation.org/en/section/the-ibrahim-index> (Stand: 1.9.2011).
- Mudde, Cas/Schedler, Andreas*, 2010: Introduction: Rational Data Choice, in: Political Research Quarterly 63 (2), 410-416.
- Müller, Thomas/Pickel, Susanne*, 2007: Wie lässt sich Demokratie am besten messen? Zur Konzeptqualität von Demokratie-Indizes, in: Politische Vierteljahresschrift 48 (3), 511-539.
- Müller, Thomas/Pickel, Susanne*, 2008: Antwort auf die Replik von Marc Bühlmann, Wolfgang Merkel, Lisa Müller und Bernhard Weßels zum Forumsbeitrag von Thomas Müller und Susanne Pickel, in: Politische Vierteljahresschrift 49 (1), 127-128.
- Munck, Gerardo L./Verkuilen, Jay*, 2002: Conceptualizing and Measuring Democracy: Evaluating Alternative Indices, in: Comparative Political Studies 35 (1), 5-35.
- North, Douglass C.*, 1990: Institutions, Institutional Change and Economic Performance, Cambridge.
- North, Douglass C.*, 2005: Understanding the Process of Institutional Change, New Jersey.
- OECD*, 2006: Measuring Human Rights and Democratic Governance to inform key policies: light version of the synthesis report on results and lessons from the two-year pilot phase of the Metagora project, Paris.
- Patze, Peter*, 2010: Demokratie und Demokratiemessung am Beispiel der Russischen Föderation. Ein Beitrag zur Standortbestimmung eines Herrschaftssystems im Wandel. Dissertation an der Universität Würzburg, Institut für Politik-

wissenschaft und Sozialforschung (veröffentlicht 2011 mit dem Titel „Wie demokratisch ist Russland?“ ,Baden-Baden).

Pickel, Susanne/Pickel, Gert 2006: Politische Kultur- und Demokratieforschung, Wiesbaden.

Ryan, Joseph E., 1994 : Survey methodology. Freedom Review 25 (1), 9-13.

Schedler, Andreas/Mudde, Cas, 2010: Data Usage in Quantitative Comparative Politics, in: Political Research Quarterly 63 (2), 417-433.

Stoiber, Michael, 2011: Die Qualität von Demokratien im Vergleich. Zur Bedeutung des Kontextes in der empirisch vergleichenden Demokratietheorie, Baden-Baden.

Thomas, Melissa, 2006: What Do The Worldwide Governance Indicators Measure? Manuscript, Johns Hopkins University, Baltimore.

Vanhanen, Tatu, 1990: The Process of Democratization. A Comparative Study of 147 States 1980-1988, New York u. a.

Korrespondenzanschrift:

Prof. Dr. Berthold M. Kuhn

Social Governance and Environmental Participation

School of Public Affairs

Xiamen University

P.R. China

E-Mail: Berthold.Kuhn@t-online.de, KB@xmu.edu.cn

Web: <http://www.innovateCo.de>