MEASURING CIVIL LIBERTY: AN ASSESSMENT OF STANDARDS-BASED DATA SETS

Midiendo la libertad civil: Una evaluación de base de datos estandarizados

SVEND-ERIK SKAANING
Aarhus University

ABSTRACT

This article offers a critical assessment of ten civil liberty measures. The evaluation addresses their focus and scope, conceptualization, measurement, and aggregation. It demonstrates that the measures are marked by differences that are often not only a question of form but also of appropriateness. In general, the justification of choices made in relation to the index constructions are inadequate and among the particular shortcomings we find severe limitations in the years covered, conceptual conflation and redundancy, restrictions on availability of disaggregate data, unsystematic and insufficient coding rules, low discriminatory power, and unfounded aggregation rules. In addition, the measures tap into two distinct types of civil liberties. The evaluation underlines that more precaution is needed in the development (and use) of civil liberty datasets.

Key words: *Civil Liberty, conceptualization, measurement, coding rules.*

RESUMEN

Este artículo ofrece una evaluación crítica de diez medidas de libertad civil (sus respectivos objetos de estudio, su alcance, su conceptualización, su medición y su agregación). El trabajo demuestra que estas medidas están marcadas por diferencias que a menudo no son sólo una cuestión de forma, sino también de idoneidad. En general, la justificación de las decisiones tomadas en relación con los indicadores es inadecuada, y dentro de las deficiencias encontramos grandes limitaciones en la cobertura temporal, reducción conceptual y redundancia, restricciones en torno a la disponibilidad de datos desagregados, reglas de codificación insuficientes y no sistemáticas, bajo poder de discriminación y normas de agregación con poco fundamento. Además, las medidas confunden dos tipos distintos de libertades civiles. La evaluación destaca que se necesita mucha más precaución en el desarrollo (y uso) de bases de datos sobre libertades civiles.

Palabras clave: Libertades civiles, conceptualización, medición, reglas de codificación.

I. INTRODUCTION

While reports on civil liberty violations certainly suggest the existence of a human rights problem, the creation of equivalent measures is often problematic for a number of methodological reasons. First of all, the availability and reliability of data for contemporary human rights studies is rather low. The general problem is quite simply that governments do not generally publish statistics on how repressive they are, and it is virtually an axiom that the more repressive the regime, the more difficult it is for researchers to have access to information about its atrocities (Goldstein, 1992: 44-45). Nonetheless, the causes, consequences, and development of civil liberties have been the subject of considerable cross-national research. Many analysts have turned to a number of quantitative data sources to facilitate their studies. Only seldom, however, have they stepped back and systematically taken stock of the conceptualization and operationalization of the core variables framing their analysis even though the quality of academic work presupposes accurate measurement.

Kenneth Bollen (1992: 189) probably exaggerates the problems associated with not having any quantitative indicators when he argues that without them, the assessment of rights has to be based on rough impressions not allowing comparisons between and shifts within countries. On the other hand, he is right that monitoring political and civil rights issues is advanced by the development of measures helping us to track differences across space and time. Such efforts, however, only signify a step forward if the development is not followed by an inclination of paying only little attention to the quality of the data; thus risking uncritical use of measures highly suspect in terms of their reliability, validity, and equivalence (Barsh, 1993; Goldstein, 1992: 48).

The recent years have seen a great expansion in the creation and use of human rights data sets, calling for a critical overview that makes crucial aspects of such measurement more transparent. Whereas the most prominent democracy indices have been scrutinized intensively (e.g. Bollen & Paxton, 2000; Munck & Verkuilen, 2002; Lauth, 2004), this does not apply to civil liberty measures. Although recent attempts to close this gap have emerged (e.g. McCormick & Mitchell, 1997; Cingranelli & Richards, 1999; Poe *et al.*, 2001), they tend only to discuss one or two measures and to focus only on personal integrity rights. Against this background, this article provides a thorough comparative assessment of civil liberty measures –out of which a number have not previously been subjected to intense scrutiny– to clarify their respective advantages and disadvantages. This task is carried out using the analytical framework for the assessment of data sets elaborated by Munck and Verkuilen (2002). Hence, the article is divided into four parts, in which the civil liberty scales and indices are evaluated as regards their empirical focus and scope, conceptualization, measurement, and aggregation, respectively.

II. FOCUS AND SCOPE

Measurement of human rights can take various forms. Rights actually enjoyed and exercised by groups and individuals can be measured through 1) survey-based data, where people

are asked about personal experiences with and/or their estimation of rights violations; 2) event-based data charting the reported acts of violation committed against groups and individuals –thus, providing answers to the questions of what, when, and who; and, finally, 3) standards-based data mirroring how often and to what degree violations occur (Landman, 2004: 918-919).

The present evaluation concerns more or less up-to-date, original, and standards-based datasets focusing on the actual respect for different civil liberty aspects in an extensive number of countries from different world regions. Thus, several well-established datasets concerning civil liberties are excluded from the discussion. For example, the Worldwide Governance Indicators (Kray *et al.*, 2005) and the Database of Liberal Democratic Performance by Joe Foweraker and Roman Krznaric (2001) are not included because their measures merely combines data from other measures, whereas others have a very limited scope either in relation to the countries covered[1] or to the aspects of civil liberty considered.[2]

Ten measures linked to seven datasets fulfil these criteria and are subjected to further evaluation.[3] A closer look at the different datasets listed in table 1 shows that none of them cover the years before 1972.

Although the datasets do not go back to the beginning of the nineteenth century as some democracy measures do (cf. Gurr *et al.*, 1990; Vanhanen, 2000), most of them display fairly impressive empirical scopes.[4] The dataset provided by Freedom House has the widest extension both concerning the number of countries and the years covered, whereas the Human Rights Data Set (CIRI) and the Political Terror Scales (PTS) demonstrate almost the same range. The Bertelsmann and the DIHR indices constitute partly exceptions though, as they do not include the old Western democracies and many small-populated countries. In addition, these datasets and the Contestation Scale only provide data on two years and the EIU Civil Liberties Index just for one year. Such restrictions on time and space variation imply that their value and relevance for research questions pertaining to different contexts is limited.

Even though the datasets share a focus on civil liberty violations, their centre of attention diverge somewhat. Not just concerning the specific rights they assess –as will be clear from the discussion of conceptualization below– but also as to whether the scores exclusively

1 E.g., Nations in Transit (Freedom House, 2006a) and Polyarchy Plus (Baker & Koesel, 2001).

2 E.g., the State Capacity Survey (State Failure Task Force, 2000); torture and fair trials (Hathaway, 2002), the Political Regime Change Dataset (Gasiorowski, 1996; Reich, 2002), Countries at the Crossroads (Freedom House, 2006b), and various measures of press freedom (Freedom House, 2006c; Reporters Without Borders, 2006). Notice, furthermore, that the assessment focuses on data (and methodology) regarding year 2006 and back. I am aware of the (minor) recent developments in the methodology of the CIRI measures but for reasons of space and to avoid confusion I have chosen not to consider the changes that are fully documented on the CIRI webpage.

3 The Bertelsmann political participation and the rule of law scales constitute parts of an overall status index measuring the state of democracy, quite broadly understood (Bertelsmann, 2005b; 2006). Also the EIU civil liberties index is part of an overall democracy measure (EIU, 2007). Due to the focus of this review, though, only the explicitly civil liberty related measures are considered.

4 Number of countries concerns the latest year covered. Some data sets have increased the number of countries covered during their lifetime.

Table 1: Civil Liberty Measures and Their Empirical Scope

Dataset	Generator(s)	Index/Scale	Countries	Years
Annual Survey of Political Rights and Civil Liberties (Freedom House)	Raymond D. Gastil Freedom House	Civil Liberties Rating	193	1972-2006
Political Terror Scales (PTS)	Michael Stohl Mark Gibney *et al.*	Political Terror Scale (AI) Political Terror Scale (SD)	177	1976-2006
CIRI Human Rights Dataset (CIRI)	David Cingranelli David Richards	Physical Integrity Index Empowerment Index	192	1981-2006
Data on Democracy and Democratization (Coppedge)	Michael Coppedge (Wolfgang Reinicke)	Contestation Scale	193	1985, 2000
Human Rights Indicators (DIHR)	Danish Institute for Human Rights	Civil and Political Rights Index	112	1998, 2001
Bertelsmann Transformation Index (Bertelsmann)	Bertelsmann Stiftung	Political Participation Scale Rule of Law Scale	119	2003, 2005
Democracy Index (EIU)	Economist Intelligence Unit	Civil Liberties Index	165	2006

reflect violations committed by the government/state or also take the overall rights condition in society into consideration. This issue is rarely addressed by the scholars using the measures and even though both foci are relevant, the choice among them must depend on the particular research question.

The Freedom House rating provides an evaluation of the state of freedom as experienced by individuals. It does not rate governments or government performance per se, but rather the real-world rights and freedoms enjoyed by individuals, so both actions by states/governmental and non-state/non-governmental actors, such as guerrillas and other armed groups, influence the assessment (Freedom House, 2004). In contrast, the CIRI dataset is solely concerned with the governments' human rights practices and any and all of its agents and nothing else (Cingranelli & Richards, 2004: 5-6). The DIHR, the Coppedge, and the Bertelsmann measures also reflect the extent of the state's/regime's willingness and/or ability to respect certain rights in practice, so they are not directed to the quality of a country's civil society with regard to the functioning of the political system as a whole (Sano & Lindholdt, 2000: 72; Bertelsmann, 2005a: 5-7; Coppedge & Reinicke, 1990: 52). Finally, the PTS adopts an intermediate position because, albeit its primary aim is to measure government terror, the coders are also instructed not to ignore other forms of terror from non-governmental actors in order to reflect human rights violations existing in a country more generally (Gibney & Dalton, 1996: 79). Although the EIU does not address this issue in explicit terms its focus also appears to be bound to an intermediate position.

III. CONCEPTUALIZATION

Another crucial part of any comparison of measures is the evaluation of the attributes and the components of the attributes (hereafter just components) singled out to reflect the core concept under consideration. It is questionable whether there is such a thing as *the* correct specification of civil liberty, and researchers should therefore refrain from claiming the settlement of conceptual disputes - thus rejecting alternatives as nothing but redundant. On the other hand, not all stipulations of constitutive elements are of equal value. As regards civil liberty, we should require –as a minimum– that the defining attributes are rooted in liberal political theory and fall under what Benjamin Constant (1988[1816]) called modern liberty, Isaiah Berlin negative liberty (1997[1958]), and David Miller (1991) liberal freedom, meaning that freedom is as a property of individuals and consists in the absence of constraint or interference by public authorities and/or other persons in general.[5]

In addition, awareness of conceptual logic and the observance of certain guidelines can reduce the likeness of conceptual pitfalls. Recalling that the extension and intension of concepts[6] are inversely related (Sartori, 1975: 64) or, in other words, more defining attributes means less referents and vice versa. Thus, a maximalist definition characterized by many defining attributes runs the risk of reducing the number of empirical referents so much that a concept's analytical worth is practically non-existing (Munck & Verkuilen, 2002: 9). Another problem frequently related to maximalist definitions is that they include attributes that for theoretical reasons are not constitutive elements of the key concept. Hence, they are flawed by low conceptual coherence besides reducing the number of testable propositions as the internal relationship between different aspects is settled by definitional fiat (Munck & Verkuilen, 2002: 9). The obvious alternative is to make use of minimalist definitions. This procedure has the advantage that it becomes easier to find empirical referents and to ask and investigate more questions. Then again, a concept covering too many cases has no great value because it lacks discriminatory power. Moreover, a minimalist definition also risks excluding aspects of a concept that are theoretically and/or empirically relevant.

A concept's constitutive attributes not only have to be identified. Apart from organizing the attributes systematically according to their abstraction level and, in so doing, redundancy and conflation must be evaded, that is, the different aspects have to be mutually exclusive and groupings of manifestations have to be connected to the same overarching attribute (Munck & Verkuilen, 2002: 12-13).

With these criteria in mind, an overview of conceptualization of the civil liberty scales and indices in question demonstrates that the constitutive attributes and their respective components of the basic concept show much convergence. However, flagrant divergences also exist as shown in table 2.

5 For a detailed discussion of the definition and normative grounding of civil liberty, see Skaaning 2006.

6 This relationship applies to classical categories where the relation among categories is expressed in terms of a taxonomic hierarchy –each category having clear boundaries and defining properties shared by all members– but not for family resemblances or radial categories (Collier & Mahon, 1993).

Table 2: Attributes and Components of Civil Liberty Scales and Indices

Name of Index/Scale	Attribute(s)	Components
Physical Integrity Index (CIRI)	Physical integrity	Political and other extrajudicial killings Disappearances Torture Political imprisonment
Empowerment Index (CIRI)	Political rights Civil liberties	Freedom of speech Freedom of religion Freedom of movement (Freedom of political participation) Workers' rights
Civil and Political Rights Index (DIHR)	Civil rights Political rights	Extrajudicial killings/disappearances Torture and ill-treatment Detention without trial Unfair trial (Participation in the political process) Association Expression Discrimination
Civil Liberties Rating (Freedom House)	Freedom of expression and belief	Free media Free religious institutions Academic freedom Open and free private discussion
	Associational and organizational rights	Freedom of assembly, demonstration, and open public discussion Freedom of political or quasi-political organizations Free trade unions, peasant organizations, and private organizations
	Rule of law	Independent judiciary Rule of law in civil and private matters, police under civilian control Freedom from police terror, unjustified imprisonment, torture, war, etc. Equal treatment under the law
	Personal autonomy and individual rights	Personal autonomy - free travel, residence, and employment Right to own property and establish private businesses Personal social freedoms Equality of opportunity and absence of economic exploitation

(it continues in following page)

(continuation table 3)

Name of Index/Scale	Attribute(s)	Components
Civil Liberties Index (EIU)	Basic human rights	Free electronic media
		Free print media
		Freedom of expression and protest
		Robust, open, free, and diverse media coverage
		Political restrictions on access to the internet
		Freedom to form professional organization and trade unions
		Opportunity to successfully petition government to redress grievances
		No use of torture
		Judicial independence of government influence
		Religious tolerance and freedom of religious expression
		Equally treatment under the law
		Basic security
		Protection of private property rights and private business
		Gender equality, right to travel, choice of work and study
		Proportion of population thinking that basic human rights are protected
		No discrimination on the basis of people's race, colour, or creed
		Government does not uses new risks and threats to curb civil liberties
Political Participation Scale (Bertelsmann)	(Popular determination)	(Rulers are determined by general, free, and fair elections)
		(Democratically elected rulers have effective power to govern)
	Political liberties	Political and/or civic groups associate and assemble freely
		Citizens, organizations, and the mass media express opinions freely
Rule of Law Scale (Bertelsmann)	(Checks and balances)	(Independence and interdependence of state powers)
		(Independent judiciary)
		(Legal or political penalties for power abusing office holders)
	Civil rights	Civil rights and citizen ability to seek redress for violations
Contestation Scale (Coppedge)	Polyarchy	(Free and fair elections)
		Freedom of organization
		Freedom of expression
		Availability of alternative sources of information
Political Terror Scales (PTS)	Political terror	Political imprisonment
		Execution
		Disappearances
		Torture

The major difference mirrors a widely acknowledged, although disputed, theoretical division between personal integrity rights and what I call personal exertion rights, such as freedom of speech, assembly, and association. The group of personal exertion rights have elsewhere been labelled basic freedoms (Saward, 1994: 16) or –sometimes with a slightly different meaning– political and/or civil rights/liberties (Shue, 1980: 19; Milner *et al.*, 1999: 405; Landman, 2004: 927; Diamond & Morlino, 2005: xxv). The distinction between personal integrity rights and personal exertion rights is furthermore reflected in the disposition of the UN human rights conventions. The Physical Integrity Index and the Empowerment Index,[7] based on different parts of the same dataset, (CIRI) reflect this partition, whereas the Political Terror Scales only concerns the first and the Bertelsmann Political Participation Scale in addition to the Coppedge Contestation Scale exclusively the second of these issues.

The Freedom House rating, the EIU index, and the DIHR index, however, constitute exceptions as the attributes of these measured are associated with both conceptual dimensions. Actually, most of the components connected to the Freedom House attribute personal autonomy and individual rights –i.e., personal social freedoms, absence of economic exploitation, and the right to own property and establish private business– do not even fall straight into any of these categories thus indicating that they are constitutive parts of other core concepts. This critique also applies to –at least– the two components of the EIU index concerning the extent to which government uses new risks and threats to curb civil liberties and the protection of private property rights and private business.

Whereas steering clear of the problem of maximalist definitions is one objective to consider, the avoidance of too minimalist definitions also have to be observed. The absence of a component concerning freedom of assembly and association by the CIRI Empowerment Index and the lack of incorporation of religious freedom in the DIHR Index seem to be illustrative examples of unjustified omissions, even though both aspects almost always are referred to as core members of basic civil liberties. Also worth mentioning, the DIHR index does not take account of freedom of movement.

Whereas the examined measures explicitly make distinctions between different levels of abstraction and make efforts to sort the attributes and components accordingly, they are not equally successful in doing this. The CIRI Empowerment Index, for instance, includes the component worker's rights, which obviously belongs to a lower level of abstraction than the other components. It mainly refers to the rights to unionize already covered by the component right to associate, which (as mentioned above) could be included instead.[8] Quite similarly, Freedom House includes the components free trade unions and peasant organizations as well as free political organizations, which on that level of abstraction could be replaced by an overall component considering the general level of associational

[7] Regarding the CIRI Empowerment Index, the association between the attributes and the components is not clearly specified. This weakness also applies to the DIHR and Bertelsmann measures.

[8] Especially as such component is already coded and part of the dataset.

freedom.[9] Such step would bring it more in line with the component freedom of assembly and demonstration covered by the same attribute. In addition, the rule of law in civil and private matters makes up a component of the attribute rule of law in the same index which signifies lack of attention to distinctions between different levels of generalization.

Another example illustrating that researchers often can improve their consciousness about the differences in 'distance' to an overarching concept is provided by the EIU index since basic security on the one hand and political restrictions on access to the internet are placed on the same aggregation level. Moreover, the civil rights component of the BTI measure is rather unspecified and lumps together many different elements. The component covers human rights, rights to justice, bans on discrimination, freedom of religion, and the extent to which citizens can seek redress for violations of these freedoms. But it is not clear from the guidelines what is meant by human rights, in particular because three obvious human rights aspects are mentioned separately in addition to the underspecified human rights category.

Somewhat disputable, but definitely not without significance, it is questionable whether the combination of the component political participation with attributes reflecting civil liberty is advantageous as exemplified by the CIRI Empowerment Index, the DIHR index as well as the Contestation and the Political Participation Scales so they reflect more or less 'thick' conceptions of democracy. A more plausible alternative could be to treat it as a separate concept just as some of the latest (as well as older) attempts to conceptualize political participation rights do (e.g. Elklit & Reynolds, 2005; Munck & Verkuilen, 2003; Gurr *et al.*, 1990; Alverez *et al.*, 1996). To some extent, the same reasoning is valid for various legal rights (e.g., independent courts, due process), of which different elements are incorporated in some of the indices.

Drawing the attention back to the outright personal integrity indices, James McCormick and Niel Mitchell (1997: 511) have argued that the treatment of the included rights as one dimension both conceptually and empirically confounds the use of political imprisonment on one side and the use of torture and killing on the other – and produces a measure closer to the latter. The critique says that political imprisonment and torture are qualitatively different activities. The use of torture and killing probably carries higher external costs than imprisonment, and governments implicitly recognize this when they opt for means of control. In short, governments choose different mixes of qualitatively different methods, and this variation is veiled if the personal integrity rights mentioned are treated as a one-dimensional phenomenon (1997: 514). This caveat is of great relevance, especially if the lack of coherence is supported by the data.

IV. MEASUREMENT

After clarifying the basic conceptual landscape, the time has come to evaluate the construction of empirical data reflecting the constitutive attributes and eventually the

[9] A very similar point can be made with regard to the two last components of the Contestation Scale and the five first components of the EIU index.

core concept in question. In addressing the formation of measures, it is an indispensable requirement that analysts record and make public the rules and choices guiding the coding process to increase the transparency and facilitate a high degree of replicability. For instance, an account for the sources providing background information for coding is essential. A practical difficulty related to this task is that the existing sources suffer from a number of well-known inadequacies. The availability of annual, worldwide information on civil liberties is generally rather limited, and the quantity and quality of this information is often inconsistent across nations, over time, and vis-à-vis different aspects (Lopez & Stohl, 1992: 217; Bollen, 1992: 189). Accessible information can be ordered according to its value. The best situation is of course characterized by the availability of all relevant raw data. The ordering of informational value continues as follows: recorded, accessible, locally reported, internationally reported, and nationally (often US based) reported. Bollen (1992: 198-199; cf. Milner *et al.*, 1999: 20), the originator of this ranking, also argues that movement from the first to the latter resembles a filtering process where some information passes through and some does not. This probably causes some kind of bias as filters often tend to be selective in some way or another.

Considerations of expediency, however, mean that many of the civil liberty measures are mainly based on reference materials provided by organizations, media etc. situated in western countries. These sources are often the most complete and certainly the easiest available, but a more or less exclusive use will undoubtedly introduce a systematic bias (Foweraker & Krznaric, 2000: 766). As shown in table 3 below, six out of ten measures do not use more than three sources for their data. Some of them, the CIRI Empowerment Index, the Contestation Scale,[10] and the political terror scales, actually just use one each. The sources used as informational background in the assessments exhibit a flagrant congruence. Five measures explicitly use the *Country Reports on Human Rights Practices* compiled by the US State Department, and three use the Amnesty International *Annual Reports*.[11] The CIRI Physical Integrity Index and the DIHR index are based on both, and the latter also draws on information provided by Human Rights Watch in its *World Reports*. The overlap is no coincidence because the material from the two worldwide human rights organizations and, especially, the State Department is widely known to provide the most consistent and valuable cataloguing of information on various human rights issues besides touching on most countries across several years.

Regarding all the examined datasets, the information on civil liberty conditions found in these reports and elsewhere is subsequently classified according to a predetermined set of coding standards. Such measurement of respect for civil liberty is bound to be disputed. Some researchers have strongly disavowed subjective measures in general and particularly in regard to human rights issues (e.g. Raworth, 2001). The main reason is that the coding can give rise to substantial reliability problems due to possible random and systematic measurement errors

[10] The 1985 data was coded on the background of multiple sources by either Michael Coppedge or Wolfgang Reinicke; both of them if superficial disagreement among the sources existed.

[11] Besides, all sources explicitly mentioned are form part of the list of sources used by Freedom House.

Table 3: Sources and Coders of the Measures

Name of Index/Scale	Sources	Coders
Physical Integrity Index (CIRI)	Amnesty International: Annual Reports State Department: Country Reports on Human Rights Practices	Trained students - at least two
Empowerment Index (CIRI)	State Department: Country Reports on Human Rights Practices	Trained students - at least two
Civil and Political Rights Index (DIHR)	Amnesty International: Annual Reports Human Rights Watch: World Reports State Department: Country Reports on Human Rights Practices	Trained student
Civil Liberties Rating (Freedom House)	Broad range of information, including news reports, academic analyses, organizations, professional contacts, and visits	Researchers and consultant writers. Reviewed by academic advisors
Civil Liberties Index (EIU)	Broad range of information, including surveys	Country expert. Review at the regional and global level.
Political Participation Scale (Bertelsmann)	Broad range of information	A domestic and a foreign expert. Adjustment/ review by regional coordinator and academic board
Rule of Law Scale (Bertelsmann)	Broad range of information	A domestic and a foreign expert. Adjustment/ review by regional coordinator and academic board
Contestation Scale (Coppedge)	State Department: Country Reports on Human Rights Practices	Trained students - three to five
Political Terror Scale (AI) (PTS)	Amnesty International: Annual Reports	Trained students - at least two. Three if disagreement
Political Terror Scale (SD) (PTS)	State Department: Country Reports on Human Rights Practices	Trained students - at least two. Three if disagreement

introduced by the raters who interpret the sources differently. On the other hand, some human rights aspects are difficult to illuminate adequately without using subjective data, and the validity of measures should have the highest priority (Munck & Verkuilen, 2002: 18; Hadenius & Teorell, 2004: 13; Bollen, 1993: 1210).

The major advantages of subjective measures are their ability to estimate key traits of civil liberty and to incorporate many different factors undetectable by objective measures.

A comprehensive coding manual may answer some of the critique facing subjective data because it can decrease coding inconsistencies. The coding of the CIRI dataset is based on a detailed coding manual of 46 pages (Cingranelli & Richards, 2004), which, apart from presenting definitions of the indicators, provides guidelines on how to handle disagreements among the sources as well as illustrative examples to help code instances of doubt. The coding guidelines connected to the remaining measures are less detailed and supportive (cf. Gibney & Dalton, 1996; Bertelsmann, 2005a; Freedom House, 2004; Sano & Lindholt, 2000; Coppedge & Reinicke, 1990: 53-54[12]), and in some instances they consist of little more than a checklist defining the parameters of the indicators combined with bare standard descriptions of how to interpret the numerical scores assigned to the cases. The guidelines outlined in the Bertelsmann codebook are merely meant as suggestions that may prepare the way for differentiated analyses of real situations (Bertelsmann, 2005b: 102), whereas separate codebooks for the EIU and DIHR[13] indices are virtually non-existent. Concerning the Freedom House rating, the checklist used in the data construction has undergone changes during the years implying that the diachronic, internal consistency of the scores is questionable (Munck & Verkuilen, 2002: 21). Then again, the Freedom House rating and the Bertelsmann measures deviate positively from the rest as they are accompanied by narrative country reports describing the circumstances (events, etc.) that have influenced the assessment.

Among the guidelines found in the codebooks we find a presentation of the range and graduation of numerical scores applied in the coding process. Generally, the measurement levels of the examined measures are neither theoretically justified nor tied to explicit discussions directed towards maximizing homogeneity within classes using a minimal number of distinctions (cf. Munck & Verkuilen, 2002: 17-18). An illustrative example of a problematic measurement level is the CIRI components freedom of religion and freedom of movement and the EUI component torture. The degree of restrictions of these rights is only assessed through the use of the very roughest form of graduation, that is, presence or absence (Cingranelli & Richards, 2004: 17-20), even though such dichotomous treatment is inadequate to distinguish different levels of government repression. The remaining CIRI components as well as all the DIHR and EUI components on the lowest level of aggregation are scored using three-point scales and thereby also express rather low discriminatory power. Freedom House applies a five-point scale and Bertelsmann (2005) even a ten-point

12 Coding criteria were amended slightly during the coding process (Coppedge, 2005).

13 A few guidelines are listed in a methodological outline, but some instructions are inopportune. For example, if no information is available, the score is 0. In addition, "In cases where only one source indicates a violation and two others provide no information, the case is [scored 0=no violation] … unless the wording of the single source is sufficiently strong to indicate confirmed and prevailing cases of violations" (Sano & Lindholt, 2000: 75). This suggestion is rather inadequate as two of the three sources just rarely comment on anything else than violations of personal integrity rights.

scale, which diminishes variance truncation but, conversely, makes it more demanding to define criteria for each score and more difficult for the coders to employ consistently.

As to the use of coders, all but two datasets, that is, EIU and DIHR, are based on the work of more than one coder in order to improve the intercoder reliability.[14] Furthermore, four out of ten measures stand out as they are based on assessments made by experts instead of students and because the data are scrutinized through review processes. Proper interrater reliability tests, though, is just undertaken for the CIRI data (Krippendorff's r-bar for each variable) and only for the year 2004. Bertelsmann and Coppedge merely report the overall percentage of agreement (or near-agreement) for their first and last year of coverage, respectively (Bertelsmann, 2005b: 103; Coppedge, 2005), which is a rather crude and inadequate appraisal. Neither Freedom House nor PTS provide any information on this issue, which further supports the call for increased conscience about measurement issues.

V. AGGREGATION

Researchers often feel compelled to combine the disaggregated scores collected in the measurement process into an overall index score in order to focus on the overarching core concept such as civil liberty. But the use of standard procedures such as simple addition rather than theoretical and empirical justification seems more to be the rule than not. In line with this general critique, only the creators of the CIRI indices and the Contestation Scale have used statistical tools to test the empirical dimensionality of their measures (Cingranelli & Richards, 1999: 410; Richards *et al.*, 2001: 226; Coppedge, 2005). They all show strong uni-dimensionality as their respective items are highly correlated. In contrast, a similar examination of the DIHR components (results not reported) indicates that this measure covers at least two dimensions. As a consequence, one could consider disintegrating it into at least two indices if also conceptual and theoretical considerations suggest this procedure-as the above discussion of conceptualization issues showed they do.

Unfortunately, the generators of the remaining measures do not provide data on a disaggregated level for all years of measurement in spite of definitions disaggregating the main concepts. Freedom House and EIU have not made their lower-level data publicly available, whereas for some of the Bertelsmann scores and for all the PTS scores, disaggregated data simply do not exist despite the rather low extra effort required to provide such. In this way, these measures only allow researchers to discriminate between countries according to their overall score on the five point scales. As from 2005, however, Freedom House and Bertelsmann provide disaggregated scores.[15] This change facilitates statistical dimensionality tests of their indices (not reported) demonstrating that the respective measures (individually) reflect only one empirical dimension.

[14] From 1977 to 1989 a single coder, Raymond Gastil, did the coding of the Freedom House scores himself. However, new coders besides adjusted coding rules introduced in 1993 do not seem to have had any significant impact (Hadenius & Teorell, 2004: 22-23).

[15] Freedom House still not on the lowest level of measurement, though.

A subsequent aggregation of lower-level data into indices is an explicit choice, which has to be justified in the light of the need to balance the desire for parsimony with a concern for underlying dimensionality and differentiation. Accordingly, aggregation rules needs foundation in an explicit theory concerning the relationship between the attributes (Munck & Verkuilen, 2002: 22-24). Yet none of the indices are accompanied by a substantial discussion establishing such a link, and the dominant aggregation rule is, without exception, simple addition at the level of components.[16] There has apparently been no effort to try out alternative aggregation rules such as multiplication, different weighting of the items or considering some of the components as either necessary or sufficient in order to test the robustness of the indices.

Actually, the definitions behind all of the measures implicitly or explicitly connect the different attributes with (logical) *and* rather than *or*. This indicates that a high score on one attribute cannot compensate a low score on another. Hence, the appropriate aggregation method would be multiplication if the defining attributes are mutually dependent, i.e., a high score on one attribute should be dragged down by a low score on another. If interaction between the attributes is not the case, one should instead take the minimum score of the attributes. Using the average is only suitable when the relationship between the attributes is partially compensatory and not interactive. Anyhow, the choice of aggregation ruled should always be based on theoretical arguments whereas statistical scalability is secondary (Munck, 2009: 48-51, 70-72; cf. Goertz, 2006: ch. 5).

Hitherto, the most common comparison of 'rival' indices has merely been simple correlation tests on aggregate data to see whether they tap into the same latent phenomenon (cf. Bertelsmann, 2005b: 107-108; Pickel, 2000: 254; Gaber, 2000: 122). This kind of test is replicated on the evaluated indices/scales, and the associated correlations are shown in table 4.

Despite the differences in conceptualization, sources, coding, etc., the measures are highly correlated. All the signs display the expected direction and a closer look at the correlations support one of the findings from the conceptual analysis, namely that the measures seem to represent two principal dimensions. A closer look at the datasets also supports this interpretation as it is plausible that countries such as Colombia, Brazil, and India get fairly high scores on personal exertion rights but low on personal integrity concerning recent years, whereas countries such as Singapore, Bahrain, Benin, Brunei and Oman have been characterized by the opposite pattern. Furthermore, the two-dimensional pattern a rotated factor analyses including the measures with the longest duration and greatest overlap (cf. Munck & Verkuilen, 2002: 30). As shown in table 4, the results highlight two significant principal factors, which account for 67% and 20% of the variation, respectively. The first latent component tends to reflect respect for personal integrity whereas the second primarily covers respect for personal exertion rights.

However, the high correlations do not necessarily mean that similar results would emerge from using them interchangeably. First, many of the bivariate correlations

[16] No aggregation rule is applied in the construction of the Bertelsmann scales for 2003 and the political terror scales as the cases are not coded at a disaggregate level.

Table 4: Correlations between Civil Liberty Scales/Indices

	Empowerment Index	Civil Liberties Rating	Physical Integrity Index	PTS (SD)	PTS (AI)	Political Participation Scale
Empowerment Index	1.00					
	(3986)					
Civil Liberties Rating	–0.82	1.00				
	(3977)	(6105)				
Physical Integrity Index	0.52	–0.60	1.00			
	(3953)	(3947)	(3956)			
Political Terror Scale (SD)	–0.49	0.58	–0.82	1.00		
	(3824)	(4889)	(3809)	(4897)		
Political Terror Scale (AI)	–0.38	0.49	–0.76	0.79	1.00	
	(3277)	(4026)	(3265)	(3892)	(4036)	
Political Participation Scale	0.80	–0.91	0.55	–0.53	–0.43	1.00
	(233)	(235)	(232)	(234)	(222)	(235)
Rule of Law Scale	0.73	–0.86	0.56	–0.60	–0.48	0.89
	(233)	(235)	(232)	(234)	(222)	(235)
Contestation	0.83	–0.90	0.41	–0.50	–0.30	–
	(288)	(357)	(287)	(330)	(266)	
Civil and Political Rights Index	–0.59	0.72	–0.66	0.65	0.53	–
	(179)	(182)	(179)	(182)	(169)	
Civil Liberties Index	0.81	–0.89	0.52	–0.50	–0.50	–
	(164)	(164)	(163)	(161)	(144)	
Factor Loadings (component 1)	0.07	0.11	–0.88	0.92	0.96	–
Factor Loadings (component 2)	0.99	–0.89	0.07	–0.03	0.07	–

Note: Results refer to bivariate Pearson's r correlations (n in parentheses) and a principal component factor analysis using Oblique rotation.

are not extremely high (below 0.75) –especially between the measures tapping into different dimensions of civil liberty– which leaves room for significant discrepancies. Second, one cannot be certain that even very high correlations lead to similar results as shown by a number of studies that have examined the impact of using different measures of democracy (Bollen and Paxton 2000; Casper and Tufis 2003; Hadenius and Teorell 2005). Third, once again focusing on the five measures with the broadest scope, I have explored the number of outspoken differences between the indices most highly correlated. More particularly, I have recalibrated the measures to have the same range (0-100) and direction and then counted the number of deviations of more than a quarter of the range, i.e., 25 points.

This analysis reveals that such large deviations make up no less than 20 percent (798/3977) of the common country-years when looking at the Civil Liberties Rating vis-à-vis the Empowerment Index. The percentage of large deviations is somewhat lower –between three and ten– for the indices measuring personal integrity rights but still noteworthy. Digging a bit into the cases behind the numbers lends more support against the Empowerment Index than the Civil Liberties Rating. Illustrating this point, the Empowerment Index's scores show a number of surprising and abrupt decreases. For instance, from 2002 to 2003 Germany, Greece, and Chile drop from very high levels to midrange or even low levels without obvious, drastic changes in respect for civil liberty in these countries. Moreover, in 2004 Germany, Lithuania, and Panama achieve a lower (and Romania a much lower) index score than Macedonia, Albania, Mali, and Paraguay and the same score as Kyrgyzstan, Haiti, Burkina Faso, and Kenya. Many similar examples could be added but suffice to say here that even though the Civil Liberty Rating is certainly not free from highly problematic country assessments, they tend to be less pronounced. A similar pattern is not identifiable with regard to the personal integrity measures which is not very surprising given their common sources and/or methodology.

VI. CONCLUSION

The goal of this article was to shed light on and assess up-to-date, standards-based civil liberty measures. The evaluation carried out provides basis for three general conclusions. First, the various scales/indices diverge greatly on all the examined parameters, that is, scope, conceptualization, measurement, and aggregation. These differences, furthermore, were not only a question of form but, indeed, also of appropriateness. Second, both theoretically and empirically the measures tapped into two related but fairly distinct dimensions of civil liberty, namely personal integrity rights and personal exertion rights. Third, all measures turned out to be infected with several weaknesses although not to the same degree.

The utility of the indices compiled by Bertelsmann, EIU, Coppedge, and DIHR is very restricted as they do not offer data on more than a year or two. Among the five remaining measures, the data on personal integrity rights provided by CIRI have the strongest standing as scores are available on a disaggregate level and the coding is based on a comprehensive codebook. The picture is not equally clear concerning the two measures with an extensive scope that (primarily) cover the respect for personal exertion rights. The CIRI data is superior in some respects as it offers data on a disaggregated level based on explicit and detailed guidelines. However, since Freedom House took over the responsibility from Gastil, the Civil Liberty Rating has been based on expert assessments and it tends to produce fewer scores that disagree strongly with empirical realities. Thus, a choice between these measures has to be directed more firmly by the needs of users.

Future research can benefit from the explications and critical points put forward in this review in two main ways. First of all, the attention has been directed toward a number of shortcomings of potential relevance for scholars who plan to apply the measures in

their studies. In addition, it can help guide the construction of new scales and indices in general and within the field of human rights in particular. Even though all of the civil liberty measures embody shortcomings, it is not the intention of the evaluation, however, to discard the use of them and even less so to reject the formation of alternatives (cf. Bollen, 1993: 1226; Munck & Verkuilen, 2002: 31). The purpose has rather been to increase awareness and caution in relation to the construction and selection of measures in general, and civil liberty indices in particular.

As emphasized by Munck and Verkuilen (2002: 31): "it is important to recognize the importance the independent value of evaluations of existing data sets" but they also insist that "Ultimately, the value of analyses of measures has to be assessed in terms of the ability to generate better data and not only evaluate existing data". Concerning the theme of this article, the claim can be interpreted as a call for an improvement of the compilation and use of already existing datasets as well as the informed construction of new ones.[17]

REFERENCES

Álvarez, Michael; Cheibub, José A.; Fernando Limongi & Adam Przeworski, 1996. 'Classifying Political Regimes', *Studies in Comparative International Development* 31(2): 3-36.

Baker, Peter J. and Koesel, Karrie J. (2001). *Measuring "Polyarchy Plus": Tracking the Quality of Democratization in Eastern Europe, 1992-2000.* Paper Presented at the APSA Annual Meeting, San Francisco, August 30 - September 2, 2001.

Barsh, Russel L. (1993). "Measuring Human Rights: Problems of Methodology and Purpose," *Human Rights Quarterly*, 15(1): 87-121.

Beetham, David (2005). "Freedom as the Foundation," in Larry Diamond and Leonardo Morlino (eds.), *Assessing the Quality of Democracy*. Baltimore: Johns Hopkins University Press.

Berlin, Isaiah (1997). "Two Concepts of Liberty", in Isaiah Berlin, *The Proper Study of Mankind: An Anthology of Essays*, London: Chatto & Windus.

Bertelsmann Stiftung (2005a). "Codebook," in *Bertelsmann Transformation Index 2003*, Gütersloh: Verlag Bertelsmann Stiftung (on CD-room in book).

Bertelsmann Stiftung (ed.) (2005b). *Bertelsmann Transformation Index 2003: Towards Democracy and Market Economy*, Gütersloh: Verlag Bertelsmann Stiftung.

Bertelsmann Stiftung (ed.) (2006). *Bertelsmann Transformation Index 2006: Towards Democracy and Market Economy*, Gütersloh: Verlag Bertelsmann Stiftung.

Bollen, Kenneth A. (1992). "Political Rights and Political Liberties," in Thomas B. Jabine and Richard P. Claude (eds.). *Human Rights and Statistics: Getting the Record Straight*, Philadelphia: University of Pennsylvania Press.

Bollen, Kenneth A. (1993). "Liberal Democracy: Validity and Method Factors in Cross-National Measures," *American Journal of Political Science*, 37(4): 1207-1230.

Bollen, Kenneth A. and Paxton, Pamela (2000). "Subjective Measures of Political Democracy," *Comparative Political Studies*, 33(1): 58-86.

[17] For a detailed description of the theoretical background and construction of a new civil liberty covering respect for personal exertion rights in 20 Latin American and 28 post-communist countries from 1977 and onwards, which is based on the framework and guidelines presented in this article, see Skaaning (2007 wwwdemocracyassesment.dk

Casper, Gretchen, and Claudio Tufis. 2003. "Correlation versus Interchangeability: The Limited Robustness of Empirical Findings on Democracy using Highly Correlated Datasets," *Political Analysis*, 11(2): 196-203.

Cingranelli, David L. and Richards, David L. (1999). "Measuring the Level, Pattern, and Sequence of Government Respect for Physical Integrity Rights," *International Studies Quarterly*, 43(2): 407-417.

Cingranelli, David L. and David L. Richards (2004). *Cingranelli-Richards (CIRI) Human Rights Database Coder Manual*, retrieved February 15, 2006, from http//:ciri.binghamton.edu/

Cingranelli, David L., and Richards, David L. (2005). *Measuring Economic and Social Human Rights: Government Effort and Achievement*. Paper Presented at the Conference on Economic Rights: Conceptual, Measurement, and Policy Issues, University of Connecticut, October 27-29, 2005.

Collier, David & James E. Mahon (1993). "Conceptual 'Stretching' Revisited: Adapting Categories in Comparative Analysis," *American Political Science Review*, Vol. 87, No. 4, pp. 845-855.

Constant, Benjamin (1988). "The Liberty of the Ancients Compared with that of the Moderns", in Benjamin Constant, *Political Writings*, Cambridge: Cambridge University Press.

Coppedge, Michael (2005). *Data used in Publications on Democratization: Excel Spreadsheet with Full Documentation*, retrieved April 24, 2006, from http//:www.nd.edu/~mcoppedg/

Coppedge, Michael and Reinicke, Wolfgang (1990). "Measuring Polyarchy," *Studies in Comparative International Development*, 25(1): 51-72.

Diamond, Larry and Morlino, Leonardo (2005). "Introduction," in Larry Diamond and Leonardo Morlino (eds.), *Assessing the Quality of Democracy*, Baltimore: The Johns Hopkins University Press.

Economist Intelligence Unit (2007). The Economist Intelligence Unit's Democracy Index, retrieved January 30, 2007, from http://www.economist.com/media/pdf/DEMOCRACY_INDEX_2007_v3.pdf

Elklit, Jørgen and Reynolds, Andrew (2005). "A Framework for the Systematic Study of Election Quality," *Democratization*, 12(2): 147-162.

Foweraker, Joe and Krznaric, Roman (2000). "Measuring Liberal Democratic Performance: an Empirical and Conceptual Critique," *Political Studies*, 48(4): 759-787.

Foweraker, Joe and Krznaric, Roman (2001). "How to Construct a Database of Liberal Democratic Performance," *Democratization*, 8(3): 1-25.

Freedom House (2006a). *Nations in Transit 2006: Democratization in East Central Europe and Eurasia*, Lanham: Rowman and Littlefield Publishers.

Freedom House (2006c). *Freedom of the Press 2006: A Global Survey of Media Independence*, Lanham: Rowman and Littlefield Publishers.

Freedom House (2004). *Methodology*, retrieved December 12, 2004, from http//:www.freedomhouse.org/template.cfm?page=35&year=2004.

Gaber, Rusanna (2000). "Demokratie in quantitativen Indizes: Ein mehr- oder eindimensionales Phänomen?" in Hans-Joachim Lauth, Gert Pickel, and Christian Welzel (eds.), *Demokratiemessung: Konzepte und Befunde im internationalen Vergleich*, Opladen: Westdeutscher Verlag.

Gasiorowski, Mark J. (1996). "An Overview of the Political Regime Change Dataset," *Comparative Political Studies*, 29(4): 469-483.

Gibney, Marc and Dalton, Matthew (1996). "The Political Terror Scale," in David L. Cingranelli (ed.), *Human Rights and Developing Countries*, Greenwich: JAI Press.

Goertz, Gary (2006). *Social Science Concepts: A User's Guide*, Princeton: Princeton University Press.

Goldstein, Robert J. (1992). "The Limitations of Using Quantitative Data in Studying Human Rights Abuses," in Thomas B. Jabine and Richard P. Claude (eds.), *Human Rights and Statistics: Getting the Record Straight*, Philadelphia: University of Pennsylvania Press.

Gurr, Ted R., Jaggers, Keith, & Moore, Will H. (1990). "The Transformation of the Western State: The Growth of Democracy, Autocracy, and State Power Since 1800," *Studies in Comparative International Development*, 25(1): 73-108.

Hadenius, Axel (1992). *Democracy and Development*, Cambridge: Cambridge University Press.

Hadenius, Axel & Teorell, Jan (2004). *Same, Same - But Different: Assessing Alternative Indices of Democracy*. Paper Presented at the APSA Annual Meeting, Chicago, September 2-5, 2004.

Hathaway, Oona (2002). "Do Human Rights Treatises Make a Difference?" *Yale Law Journal*, 111(8): 1935-2042.

Hathaway, Oona and Ho, Daniel E. (2004). *Characterizing Measurement Error in Human Rights*, Paper Presented at the APSA Annual Meeting, Chicago, September 2-5, 2004.

Kray, Aart C., Kaufmann, Daniel and Mastruzzi, Massimo (2005). *Governance Matters IV: Governance Indicators for 1996-2004*. World Bank Policy Research Working Paper, WPS3630.

Landman, Todd (2004). "Measuring Human Rights: Principle, Practice and Policy," *Human Rights Quarterly*, 26(4): 906-931.

Lauth, Hans-Joachim (2004). *Demokratie und Demokratiemessung: Eine konzeptionelle Grundlegung für den interkulturellen Vergleich*, Wiesbaden: Verlag für Sozialwissenschaften.

Lauth, Hans-Joachim, Pickel, Gert and Welzel, Christian eds. (2000). *Demokratiemessung: Konzepte und Befunde im internationalen Vergleich*. Opladen: Westdeutscher Verlag.

Lopez, George A. and Stohl, Michael (1992). "Problems of Concept and Measurement in the Study of Human Rights," in Thomas B. Jabine and Richard P. Claude (eds.), *Human Rights and Statistics: Getting the Record Straight*, Philadelphia: University of Pennsylvania Press.

Mainwaring, Scott, Brinks, Daniel and Pérez-Liñán, Aníbal (2001). "Classifying Political Regimes in Latin America, 1945-1999," *Studies in Comparative International Development*, 36(1): 37-65.

McCormick, James M., & Mitchell, Niel J. (1997). "Human Rights Violations, Umbrella Concepts, and Empirical Analysis," *World Politics*, 49(4): 510-525.

McHenry, Dean E. (2000). "Quantitative Measures of Democracy in Africa: An Assessment," *Democratization*, 7(2): 168-185.

Miller, David (1991). "Introduction", in David Miller (ed.), *Liberty*, Oxford: Oxford University Press.

Milner, Wesley T., Poe, Steven C. and Leblang, David (1999). "Security Rights, Subsistence Rights, and Liberties: A Theoretical Survey of the Empirical Landscape," *Human Rights Quarterly*, 21(2): 403-443.

Munck, Gerardo (2009). *Measuring Democracy: A Bridge between Scholarship and Politics*, Baltimore: The Johns Hopkins University Press.

Munck, Gerardo and Verkuilen, Jay (2002). "Conceptualizing and Measuring Democracy: Evaluating Alternative Indices," *Comparative Political Studies*, 35(1): 5-34.

Munck, Gerardo and Verkuilen, Jay (2003). *Bringing Measurement Back In: Methodological Foundations of the Electoral Democracy Index*, Paper Presented at the APSA Annual Meeting, Philadelphia, August 28-31, 2003.

Pickel, Gert (2000). "Subjektive und objektive Indikatoren der Demokratiemessung im Vergleich - Grundlegende Unterschiede oder gleiche Ergebnisse?" in Hans-Joachim Lauth, Gert Pickel, and Christian Welzel (eds.), *Demokratiemessung: Konzepte und Befunde im internationalen Vergleich*, Opladen: Westdeutscher Verlag.

Poe, Steven C., Carey, Sabine C., & Vazquez, Tanya C. (2001). "How are These Pictures Different? A Quantitative Comparison of the US State Department and Amnesty International Human Rights Reports, 1976-1995," *Human Rights Quarterly*, 23(3): 650-677.

Raworth, Kate (2001). "Measuring Human Rights," *Ethics and International Affairs*, 15(1): 111-131.

Reich, Gary (2002). "Categorizing Political Regimes: New Data for Old Problems," *Democratization*, 9(4): 1-24.

Reporters Without Borders (2004). *Third Annual Worldwide Press Freedom Index*, retrieved December 12, 2004, from http://www.rsf.org/.

Richards, David L., Gelleny, Ronald D. and Sacko, David H. (2001). "Money with a Mean Streak? Foreign Economic Penetration and Government Respect for Human Rights in Developing Countries," *International Studies Quarterly*, 45(2): 219-239.

Sano, Hans-Otto & Lindholt, Lone (2001). *Human Rights Indicators: Country Data and Methodology 2000*, retrieved June 13, 2005, from http//:www.humanrights.dk/.

Sartori, Giovanni (1975). "Guidelines for Concept Analysis," in Giovanni Sartori (ed.), *Social Science Concepts: a Systematic Analysis*, Beverly Hills: Sage.

Saward, Michael (1994). "Democratic Theory and Indices of Democratization," in David Beetham (ed.), *Defining and Measuring Democracy*, London: Sage.

Shue, Henry (1980). *Basic Rights: Subsistence, Affluence, and U.S. Foreign Policy*, Princeton: Princeton University Press.

Skaaning, Svend-Erik (2006). Defining and Founding Civil Liberty. Working Paper Nº 56. Center on Democracy, Development, and the Rule of law, Stanford University.

Skaaning, Svend-Erik (2007). The Civil Liberty Data set: Conceptualization and mea surement Zeitschreft für vergleichende Politikwissenschafty 2(1), pp. 29-51.

State Failure Task Force (2000). *State Failure*, retrieved February 15, 2006, from http//:globalpolicy.gmu.edu/pitf/

Vanhanen, Tatu (2000). "A New Dataset for Measuring Democracy, 1810-1998," *Journal of Peace Research*, 37(2): 251-265.

Svend-Erik Skaaning, PhD in Political Science, Assistant Professor, Department of Political Science, Aarhus University. Visiting scholar at Stanford University in 2006 and at University of North Carolina at Chapel Hill in 2009. Main research interests: democratization, conceptualization and measurement of democracy and human rights, comparative methodology. (E-mail: Skaaning@ps.au.dk)