Map-Making and Analysis of the Main International Initiatives on Developing Indicators on Democracy and Good Governance

Dr. Todd Landman
Deputy Director
Human Rights Centre
University of Essex
Wivenhoe Park
Colchester, Essex CO4 3SQ
United Kingdom
todd@essex.ac.uk

Rights and Humanity

Mrs. Julia Häusermann
President
Rights and Humanity
2 St. Peter’s Street
Ipswich, Suffolk IP1 1XB
United Kingdom
jhausermann@pop3.poptel.org.uk

Thursday, 24 July 2003
Preface

PART 1 – OVERVIEW

1.1 Introduction

1.2 Conceptual Framework

1.3 Synopsis of Methodologies and Classification of Indicators Used in Main Initiatives

PART 2 - ANALYSIS & EVALUATION OF THE DIFFERENT APPROACHES AND METHODOLOGICAL OPTIONS

2.1 Democracy

   Analysis of Main Initiatives for Measuring Democracy
   Evaluation of Main Initiatives for Measuring Democracy

2.2 Human Rights

   Analysis of Main Initiatives for Measuring Human Rights
   Evaluation of Main Initiatives for Measuring Human Rights

2.3 Good Governance

   Analysis of Main Initiatives for Measuring Good Governance
   Evaluation of Main Initiatives for Measuring Good Governance

PART 3 - CONCLUSIONS AND RECOMMENDATIONS

3.1 Conclusions

3.2 Recommendations

   General Recommendations
   Specific recommendations to the European Commission

ANNEX I Inventory of Initiatives of that measure democracy, human rights and good governance

ANNEX II References cited in the Final Report

ANNEX III Conceptual Issues

ANNEX IV Guide to the Database of Initiatives
Preface

The project entitled “Map-Making and Analysis of the Main International Initiatives on Developing Indicators on Democracy and Good Governance” was commissioned by the Statistical Office of the Commission of the European Communities (EUROSTAT). The overall objectives of the project were:

- to provide a synopsis of the different approaches and methodological options available for measuring Democracy and Good Governance to those interested in this particular field;
- to supplement and enrich the documentation related to the development of indicators on Democracy and Good Governance, subject of the EUROSTAT seminar on Human Rights in 2002; and
- to serve for increased efficiency in the development of indicators related to Democracy and Good Governance aimed at monitoring governmental actions.

The main activities consisted of: identifying, collating and analysing the main initiatives to develop indicators for measuring democracy, human rights and good governance by academics, inter-governmental organizations (IGOs), and non-governmental organizations (NGOs); evaluating the strengths and weaknesses of the main initiatives; assessment of the links (or missing links) of the different initiatives with on-going academic research on indicators; and the development of general recommendations on priority setting and basic orientations for developing related indicators.

The project was carried out by a team comprising Dr. Todd Landman (Project Leader and Deputy Director of the Human Rights Centre at the University of Essex), Mrs. Julia Häusermann (Primary External Consultant and President of Rights and Humanity), Mr. Sebastian Dellepiane (University of Essex Senior Research Officer), Ms. Olivia Wills (Research Officer at Rights and Humanity), and Mr. Matthew Sudders (Consultant to Rights and Humanity).

The six-month project commenced on 2nd December 2002. An interim report was submitted to EUROSTAT on 2nd March detailing the preparatory work undertaken (including the pre-contract participation of the Project Leader and Primary External Consultant in the EUROSTAT seminar on Human Rights and Statistics in November 2002), and setting out the preliminary findings.

This Final Report, submitted on 2nd June, 2003, is structured in three parts:

- Part One provides an overview of the conceptual framework and terminology used and a synopsis of the methodologies and categories of indicators used in the main initiatives.
- Part Two provides an analysis and evaluation of the main international initiatives in this regard, together with an evaluation of the different approaches and methodological options.
- Part Three comprises conclusions and recommendations.

This Final Report is accompanied by a number of annexes including a comprehensive inventory of reviewed initiatives.
PART 1 – OVERVIEW

1.1 INTRODUCTION

1. The project team has identified, and where possible made contact with, over 550 initiatives or written texts on democracy, human rights and good governance. These originate from academics and academic centres, inter-governmental organizations (IGOs), and non-governmental organizations (NGOs). Of these, over 170 initiatives were identified as having served as seminal efforts to measure democracy, human rights, and good governance. Of all of the reviewed initiatives, 45 main initiatives have developed methodologies or indicators that have stood the test of time, are used frequently in empirical studies and policy documents, are updated regularly or are cited as examples of best practice (see Excel version of the database-Main Initiatives). Moreover, the team sought to concentrate on those initiatives that have wide geographical and temporal coverage. These main initiatives are analysed and evaluated in Part II below.

2. Annex I provides an inventory of the initiatives, conceptual and methodological discussions, and secondary analyses; Annex II lists the references cited in the final report; Annex III discusses further conceptual issues. The final report is accompanied by Access and Excel versions of the database of 178 total initiatives in which the 45 key initiatives are flagged with a separate field (Annex IV is a guide to both).

3. The team has sought to concentrate on those sources that represent the genesis of measures for democracy, human rights and good governance. However, it has also been important to review initiatives that use secondary sources as a basis for their work to map the concepts globally, test empirical relationships, and draw larger comparative inferences (see Annex I and Literature Review in the Excel Version of the Database). Such secondary sources can provide important advocacy tools for the promotion of democracy, human rights and good governance and uncover important methodological problems, including informational source bias, variance truncation, conceptual confusion, problems of aggregation, and other biases (e.g. gender bias in some democracy measures, ideological biases in some freedom measures and cultural biases in some good governance measures).

4. It is important to note that the views, comments, discussions, and recommendations contained within this Final Report are those of the principal authors and not of the European Commission.

1.2 CONCEPTUAL FRAMEWORK

5. Both democracy and good governance remain ‘essentially contested concepts’ (Gallie 1956), since there is not now, nor will there likely be, a final consensus on their definition or content. It is not surprising, therefore, that the European Union avoids defining the term ‘democracy’. For instance, in the revised fourth Lomé Convention it opted instead for the phrase ‘democratic principles’ (Article 5, revised fourth Lomé Convention). It did so in order to emphasise ‘the universally recognised principles that must underpin the organisation of the state and guarantee the enjoyment of rights and fundamental freedoms, while leaving each country and society free to choose and develop its own model’ (European Commission 1998). The European Commission
considers that the principles can be defined in terms of three fundamental characteristics: legitimacy, legality and effective application.

6. There is much greater clarity concerning human rights. These have now been codified in a wide range of UN and regional texts. The UN legal framework comprises the UN Charter, the Universal Declaration of Human Rights and six other core treaties (see Annex III) and covers civil, cultural, economic, political and social rights. Human rights have been recognised by the world community as being universal - every human being is entitled to these rights simply by reason of being human.

7. It is also recognised that all human rights, be they civil and political or economic, social and cultural, are indivisible and inter-dependent (World Conference on Human Rights, Vienna, 1993). Neither branch of human rights should be given priority over the other, and states have the primary obligation to respect, protect and ensure the universal enjoyment of all human rights. Governments have the obligation to ensure enjoyment of some human rights immediately, whilst others, predominantly economic, social and cultural rights, are to be realised progressively. These are important distinctions when it comes to measuring government performance in the field of human rights. It is also important to distinguish between government obligations on the one hand, and enjoyment of human rights by individuals and groups on the other, in order that appropriate measurement tools might be developed for each of these aspects.

8. The term ‘good governance’ emerged in the late 1980s and early 1990s primarily in the World Bank, which was concerned about the ways in which governance influenced economic performance (see World Bank 1992). The economic dimension of good governance has variously included public sector management, organisational accountability, the rule of law, transparency of decision-making, and access to information. This idea was taken on board by the OECD and EU and integrated into its requirements for development assistance. It was later expanded by the United Nations Development Programme (UNDP) to incorporate a political dimension that includes government legitimacy, government accountability, government competence, and the protection of human rights through the rule of law.

9. The European Commission has defined good governance as ‘the transparent and accountable management of all a country’s resources for its equitable and sustainable economic and social development’. It lists a number of aspects of good governance, such as equity and the primacy of law in the management and allocation of resources, an independent and accessible judicial system and transparency, and recognises that corruption is the main obstacle to good governance (European Commission 1998).

10. More recently, the European Commission has regarded the term as comprising six components: human rights, democratisation, the rule of law, the enhancement of civil society and public administration reform (including decentralisation) (Draft EC Good Governance Manual, version created 04/02/2003). In other words, it regards democratisation and respect for human rights as being essential ingredients of good governance. As we have seen above, the EC also regards democratic principles as underpinning the guarantee of the enjoyment of rights and fundamental freedoms, and thus regards all three categories as being interlinked. (The database included with this report contains a field that matches the initiatives to the clusters. In some cases, the initiatives may map onto several clusters).
11. Indeed, the most popular definitions of democracy and good governance now include reference to the protection of certain categories of human rights, especially civil and political rights. But they also make reference to some economic and cultural rights, such as property rights and the rights of minorities (see Foweraker and Krznaric 2000). Similarly, definitions of human rights, drawn from the long history of their international legal evolution make reference to the right to participate in public affairs and democratic decision-making, and make explicit reference to a right of everyone to take part in the government of his country, directly or through freely chosen representatives (e.g. Article 21(1) of 1948 Universal Declaration of Human Rights). Moreover, many consider democracy to be ‘hollow’ without the protection of civil and political rights (Diamond 1999), while governance is considered to be ‘bad’ without the rule of law and the protection of human rights.

12. Despite their inextricably linked components, the concepts of democracy, human rights and good governance should not be seen as equivalent concepts since each has important exclusive characteristics as well as shared elements. In the present study, therefore, the measurement of each of these concepts was reviewed separately.

13. In the absence of internationally agreed definitions of the terms democracy and good governance, it is not surprising that the initiatives mapped indicate that much conceptual confusion remains (see Annex III for a more detailed discussion on conceptual issues).

14. A key finding of the map-making project is that measurement is inextricably linked to conceptual definition. The existence of different definitions of these terms necessarily means the existence of different measures of these concepts. Yet, there has been a tendency to elide or conflate these three ideas as well as to form a ‘chain of equivalence’ (Howarth 2000) among them that permeates international foreign policy documents.

15. These differences in the use of terms and in intellectual understanding of concepts compound a variety of methodological problems, including source bias, coding bias, non-transparency of coding rules, reliability, and validity. A prerequisite for a consensus on measurement tools and indicators is the identification of a clear conceptual framework that disentangles these ideas as much as possible in order to allow the development of agreed meaningful cross-national and time-series indicators.

1.3 SYNOPSIS OF METHODOLOGIES AND CLASSIFICATION OF INDICATORS USED IN THE MAIN INITIATIVES

16. Different methodologies and tools have been developed depending on whether the purpose of measurement is, for example, to analyse the overall situation or general trends within a country, to determine the degree to which a specific government is respecting its obligations in principle, or to determine the practical situation on the ground, for instance the enjoyment by individuals of their human rights.

17. Initiatives developed to measure democracy, human rights and good governance face the difficulty that there are few uncontroversial tools of measurement (UNDP, Human
Development Report 2002, page 36). An analysis of the main initiatives shows that across the three concepts, a number of methods of primary data collection have been utilised. These include: national census, household surveys, different approaches for ascertaining public perceptions (such as international panels, public opinion surveys, focus groups, and expert opinion), in-house expert opinion within the institution undertaking the initiative, official statistics collected by governments and inter-governmental agencies, and monitoring of government commitments and of individual events which promote or impede democracy, human rights or good governance. In addition, a number of secondary methods of data collection have been developed, primarily by academics who use a variety primary source material to produce comparable measures of democracy, human rights, and good governance.

**Categories of data collection**

18. These different methodologies give rise to differing categories of data. **Survey-based data** identify the level of perceptions of democracy, human rights, and good governance. Using, for instance, standardised household survey instruments, such measures ask a battery of questions related to one or more of the three concepts to determine popular perceptions. Most surveys uses random samples of the target population, making the use of inferential statistical techniques possible; however, the measures of governance and good governance by and large do not use random samples, but small samples of ‘important people’ (20 < N < 50) within each country such as those from business, government, and industry (see United Nations University World Governance Survey; Transparency International Corruption Perceptions Index). Examples of initiatives collecting survey-based data on democracy that use random samples include the Global Barometer Surveys, currently comprising the New Europe Barometer (since 1991), Latinobarometer (since 1995) Afrobarometer (since 1999) and East Asia Barometer (since 2001). Those surveys using random samples on human rights include the World Values Survey (Inglehart 1997), which has individual level data on 43 countries, and the Gallup Millennium Survey, which produced individual level data on 61 countries from a one-off survey at the turn of the century.

19. **Standards-based data** use as a framework ideals derived from the concepts of democracy, human rights or good governance. This framework is then used to judge the degree to which the ideals have been realised (or protected) based on a variety of source information. Such measures use a standardised scale that is ordinal, interval or nominal to award a country a score for its performance. These scales are normally aggregated into a single performance index. For example, the ‘Polity’ series of data sets on democracy provides standards-based scales on components such as executive constraint and the competitiveness of the nomination process, which are then aggregated into a single democracy score (Jaggers and Gurr 1995). The ‘political terror scale’ specifies a 1 to 5 scale that measures the degree to which a country violates so-called personal integrity rights, including torture, political imprisonment, extra-judicial killing, and disappearance (see Poe and Tate 1994). Initial (and some current) research on good governance uses the Freedom House civil and political liberties scores, which are standards-based ordinal scales that range from 1 to 7 (see below and Knack 2002).

---

1 The ‘old’ Eurobarometer has been collecting standardised survey data on democracy and other issues affecting the European Union since the 1960s (see Kasse and Newton 1995).
20. **Events-based data** result from counting specific events that promote or impede democracy, human rights and good governance. Sources for events-based data include newspapers, NGO informational networks, human rights victim testimonies, and business monitoring of government performance. The events used to collect such data are usually divided between the occurrence of *positive* events (e.g. a free and fair election) or *negative* events (e.g. a discriminatory act, a corrupt practice, or violation of human rights, such as torture, mistreatment, harassment, extra-judicial killing, or disappearance). The units of analysis for positive events include individual and collective occurrences that in some way reflect improvements or achievements in democracy, human rights or good governance. For example, Vanhanen’s (1997) measure of democracy includes voter turnout, which is meant to capture the degree of participation in a country. In similar fashion, the units of analysis for negative events include *individual* violations (e.g. denial of the right to vote, instances of torture or extra-judicial killing, or acts of corruption) or *collective* violations (e.g. awarding government contracts to favoured parties, or denial of access of members of a particular group to a hospital or polling booth).

21. It is important to note that recent work on monitoring and counting human rights violations has cast serious doubt on the reliability of newspaper sources. Long the preferred and mainstay source for behaviouralists working in political science research on political conflict (see Taylor and Hudson 1972; Taylor and Jodice 1983; Banks 1971, 1979, 1994), newspapers are increasingly seen as problematic for coding events in any meaningful way (see Foweraker and Krznaric 2000). Indeed, by comparing NGO-collated data on gross human rights violations in Guatemala and newspaper reports, one human rights analyst has demonstrated that there is an inverse relationship between the occurrence of human rights violations and the newspaper reporting of such violations (see Ball 2003).

**Classification of Indicators**

22. The study reveals that indicators developed to measure democracy, human rights and good governance are categorised in various ways. For instance, UNDP refers to *objective* indicators, such as voter turnout, socio-economic performance, or the ratification of human rights treaties, and *subjective* indicators based on expert or mass opinion surveys on such issues as the level of freedom of expression in a given country or perceptions of corruption among public officials (UNDP 2002, page 36).

23. Indicators can also be classified according to whether they are aimed at measuring the performance of an obligation bearer (‘input indicators’), the process of policy implementation (‘process indicators’) or the level of progress achieved (‘outcome indicators’). To take examples from human rights measurement: input indicators include the ratification of international human rights texts and the protection of human rights in national law (*de jure* protection of human rights) or the allocation of the necessary resources for the provision of public services; process indicators measure how the policy is implemented, for example the level of accountability and participation of different sectors of society; and outcome indicators include the level of enjoyment of human rights by individuals and groups (*de facto* enjoyment of human rights) (Rights and Humanity 2001; Häusermann, 2002). However, this categorisation has also been employed differently. For example, Malhotra regards the term process indicators to mean those ‘indicators that reflect the progressive realisation of the rights,
milestones on the path towards realising the desired outcomes’ (Malhotra, 2002), whereas the present authors consider that changes in the progressive realisation would be revealed by the data from outcome indicators collected regularly over a period of time. Nevertheless, in a rare attempt to collate indicators to measure a rights-based approach to development, Malhotra correctly lists as process such state policy indicators for improving food availability (legal rights of women to agricultural land, public investment in irrigation, food subsidies), the population covered by the public distribution system (proportion of household food consumption), early warning indicators on crop failure and food security, and the capacity for governments to provide relief in the case of food security problems, crop failure, and famine. Governments possess such information and produce official statistical reports on them. National statistical offices could be better informed about how the indicators they collect can serve as indicators for the different categories of human rights.

24. In addition, as noted above, human development indicators are sometimes used as proxy outcome indicators of the general trend in the enjoyment of human rights. For instance, improvements in national literacy rates may be cited as evidence of progress in the realisation of the right to education. However, national literacy rates may not reveal denial of enjoyment of the right to education by members of a minority group or others suffering discrimination. Equally, it is possible to provide indicators such as hospital beds per 100,000 people, percentage of governmental expenditure on the national healthcare system, and number of doctors per hospital as measures of provision of healthcare resources, but none of these indicators actually measures the denial of access to healthcare services. Thus, there is still a lack of complementary indicators that reflect a human rights perspective (see Conclusions and Recommendations).

25. In general, it must be stressed that there are always trade-offs between the different types of indicators. Those that achieve global coverage tend to have a higher level of abstraction and may not provide the kind of differentiation required for policy analysis or policy decision-making. Those indicators that provide highly detailed event counts are difficult to produce across a large global sample of countries. Moreover, proxy indicators for the three concepts may be readily available but are less valid since there is a great ‘distance’ between the concept that is being measured and the indicators that is being used to measure it. The development of indicators has thus had to confront these inevitable trade-offs.
PART 2 – ANALYSIS AND EVALUATION OF THE DIFFERENT APPROACHES AND METHODOLOGICAL OPTIONS

26. This section of the report combines a descriptive account with evaluative statements about the initiatives with respect to their methodological strengths and weaknesses. The critical evaluation of indicators by concept contained in this section of the Final Report focuses on the following eight methodological problems: (1) validity (i.e. does the indicator measure what it purports to measure?),\(^2\) (2) reliability (i.e. can the indicator be produced by different people using the same coding rules and source material?), (3) measurement bias (i.e. are there problems with systematic measurement error?), (4) lack of transparency in the production of the indicator, (5) representativeness (i.e. for survey data, what is the nature of the sample of individuals?), (6) variance truncation (i.e. the degree to which scales force observations into indistinguishable groupings), (7) information bias (i.e. what kinds of sources of information are being used?), and (8) aggregation problems (i.e. for combined scores, to what degree are aggregation rules logically inconsistent or overcomplicated).

---

\(^2\) It is interesting to note that the Freedom House scales have been used as measures of democracy, good governance, and human rights.
2.1 DEMOCRACY

Analysis of Main Initiatives for Measuring Democracy

27. Of the three concepts, the longest academic, IGO, and NGO tradition has been in the measurement of democracy. Modern social science attempts at measuring democracy began in the post World War II period, where initial attempts concentrated on the protection of political and civil rights, as well as institutional mechanisms that upheld the values of democracy. Between 1954 and 1965, there were 2,080 different indices of democracy of which only 28 percent had been used more than once (Barsh 1993, page 91). Since that time, five major traditions in the measurement of democracy have developed.

28. The first tradition establishes standards-based scales of different dimensions of democracy. The major influence in this tradition came from Robert Dahl, who provided measures of ‘polyarchy’ for 114 countries circa 1969. Such initiatives have continued and have inspired four of the most dominant and enduring indicators of democracy that use scales: (1) the Raymond D. Gastil and Freedom House 7-point scales of political and civil liberties, which have been produced on an annual basis since 1972 and cover all the independent nation states in the world (www.freedomhouse.org; see also Coppedge and Reinicke 1988, 1990, 1991); (2) the ‘Polity’ data series (Polity I, II, III, and IV) that contain 11-point scales of autocracy and democracy (0-10) for all the independent nation states in the world since the 1850s (see Jaggers and Gurr 1995; Marshall and Jaggers 2000); (3) Banks’s (1994; 1997) institutional scales of democracy for 115 countries between 1850 and 1997 (see also Fowleraker and Landman 1997, Appendix B, pages 251-252), and (4) Bollen’s (1998) global index of liberal democracy for 1950-1990. This tradition of standards-based scales has also extended to the measurement of human rights (see the discussion below).

29. The second tradition in the measurement of democracy draws on Aristotle’s analysis of good and corrupt forms of rule which categorised regime types using the city-states of the time as the empirical base (Landman 2003, pages 7-8; see also Finer 1997). In 1959, Seymour Martin Lipset categorised countries in Europe and Latin America into stable democracies, unstable democracies, stable dictatorships, and unstable dictatorships. The tradition of regime classification has continued until today and is best represented by (1) Gasiorowski’s (1996) political regime change data set, (2) the global study of Przeworski, Alvarez, Cheibub, and Limongi (2000), which has categories for democracies and non-democracies for the period 1950 to 1990 and (3) Dorenspleet’s (2000, 2001) work on the ‘waves’ of democratisation, which extends the categorisation found in Przeworski et al. (2000) to 1994.

30. The third tradition in measuring democracy attempts to establish objective indicators that measure Dahl’s two dimensions of a polyarchy: contestation and participation. The sole source of such objective indicators has been Tatu Vanhanen. The contestation indicator is the smallest parties’ share of the vote (i.e. 100 minus the largest party share) and the participation indicator is voter turnout (i.e. percentage of voters that vote in an election). Vanhanen multiplies the two indicators together and divides the product by 100 to produce an ‘index of democratization’. His first global data set was made available in 1984 and covered 116 states from 1850 to 1979. It has been updated to
cover new countries and time, and now includes measures of democracy for 187 countries from 1810 to 1998 (Vanhanen 2000). The data set is currently available at: www.sv.ntnu.no/iss/data/vanhanen/.

31. The fourth tradition, begun most notably in the Civic Culture by Almond and Verba (1963), uses survey-based indicators of mass public perception of democracy and the quality of democratic institutions. Two sources produce regularly updated mass data on a variety of countries: (1) the Global Barometer Surveys (as above and see www.globalbarometer.org) and (2) the World Values Surveys, which now contain data on 43 countries (see Inglehart 1977; 1990; 1997).

32. The fifth and final tradition (and related to the fourth) creates so-called ‘image indices’, which poll expert opinion on the quality of democracy at a given time and place. For example, rather than poll mass publics, Fitzgibbon and Johnson sought to measure the quality of democracy in Latin America using a systematic survey instrument that probed the views of country specialists on a series of social and political scales ranging from 1 to 5 that they felt represented both the preconditions and manifestations of democracy (see Fitzgibbon 1967, page 135). Their index has been produced every five years from 1945 to 1985. This tradition in polling expert opinion continues today and has been used for early human rights measurement (see Schoultz 1981), but such measures are less prevalent for democracy and human rights, and are now being used mostly for measures of good governance (see below).

33. Of the five traditions in the measurement of democracy, the first four have stood the test of time and have been updated on a regular basis. Table 1 provides descriptive information on these main initiatives in the measurement of democracy, including the source of the initiative, the type of measurement, its geographical coverage, and its temporal range. These initiatives are flagged in the accompanying database as ‘key’ initiatives.

**Evaluation of Main Initiatives for Measuring Democracy**

34. The previous section showed that the predominant initiatives in measuring democracy include standards-based scales and their categorical variants, survey-based indicators, and the unique ‘objective’ indicator of democracy produced by Tatu Vanhanen. Their various methodological problems are discussed in turn.

35. Standards-based scales translate qualitative information into quantitative statements that range across a discrete set of values (e.g. 1-10). The key initiatives such as Freedom House, the Polity series, and Bollen all have wide geographical and long temporal coverage. Yet, many of the standards-based scales of democracy are not entirely explicit if a move from 1 to 2 is the same as a move from 3 to 4; or how qualitative information yields the score that is assigned. Correlation with other existing democracy scores is not a foolproof method for verification since they all may suffer from the same systematic biases (usually ethnocentric). Of all the standards-based scales, Freedom House suffers from the largest number of methodological problems (see Munck 2002), including validity (it is technically a measure of freedom not democracy), reliability (it has never been clear how its checklists translate into the 7-point scores), ideological biases against former communist states and Islamic states, variance truncation (Freedom House cannot differentiate among the advanced
industrial democracies), and aggregation problems (e.g. the sum of a civil liberty score of 4 and a political liberty score of 2 is the same as the sum of a civil liberty score of 2 and a political liberty score of 4 even though the substantive interpretation of these different combinations is different).

36. Dichotomous classification of regimes into either democracies or non-democracies rests on the conceptual assumption that democracy is an ‘all or nothing affair’. Those initiatives that code countries in this way have a minimum threshold for classifying a country as a democracy. For example, Przeworski et al. (2000: pages 18-29) apply a simple set of decision rules to classify a country as a democracy: (1) the chief executive must be elected, (2) the legislature must be elected, (3) there must be more than one party, and if all these rules are met (4) there has been significant alternation in power. Even applying the strict criteria for the dichotomous categories has yielded exceptions, which may appear inconsistent since Przeworski et al (2000: pages 23-25) want to avoid making a type-II error, i.e. classifying a non-democratic regime as democratic (the absence of significant alternation of power in Botswana is illustrative). They thus cautiously *undercount* the number of democracies. Moreover, secondary analysis of the Polity democracy score carried out by Doorenspleet (2001) shows a largely bi-modal distribution, suggesting that dichotomous categorisations such as those employed by Przeworski et al. (2000) may capture most of the global variation in democracy anyway. The key question is where specific countries fall within that distribution.

37. In their critique of extant democracy measures, Foweraker and Krznaric (2001) argue that there is still a large institutional bias toward measuring electoral and procedural democracy but not liberal democracy, which takes into account the protection of civil rights, property rights, and minority rights. To date, the key initiatives all suffer from this narrow institutional focus. Indeed, the four rules from Przeworski et al. (2000) outlined above say nothing about the protection of rights.

38. Vanhanen's 'index of democratization' is ostensibly the only objective measure of democracy since it relies on indicators of Dahl's dimensions of participation and contestation that are the directly measurable result of an election: voter turnout and party share. The index is easy to comprehend and the data are readily available. However, the index is not without its problems. First, it further reduces Dahl's concept of democracy in its operationalisation, leading many to doubt whether turnout is a valid measure of participation. Voter turnout is normally calculated by using the registered number of voters who voted, but in many developing countries, the key problem is that many voters never get registered. Turnout figures are suspect since for many areas the count of voters is often fraught with difficulties. These problems are not isolated to developing countries as the 2000 US presidential election vote count in Florida aptly demonstrates. In addition, the party share indicator does not reflect the nature of the party system or the electoral system. Single-member and proportional electoral systems tend to be associated with two-party and multiple party systems, respectively. Thus subtracting the largest party share from 100 as Vanhanen does is a crude measure of contestation. Third, the simple multiplication procedure for aggregating the two components into the index has been criticised, since there is no clarification on

---

3 There is also the conceptual problem of defining what contestation and representation actually mean. There is a large literature in political science on counting parties and whether party systems are representative or not (see Lijphart 1994).
weighting either component and since a low score on one component cannot be compensated by a high score on the other component (Munck and Verkuilen 2000, page 36.)

39. Finally, survey-based measures of democracy are not measures of democracy *per se*, but merely reflections of mass or expert opinion on the quality of democracy and the quality of democratic institutions. In many instances, the survey questions merely ask whether the respondent is happy with democracy, or happy with the kind of democracy that exists in his or her country. Thus, secondary analysis of *Eurobarometer* surveys shows a steady 60% support for democracy in the post war period, but says very little about the quality of democracy itself (see Kaase and Newton 1995). In similar fashion, all 21 republics in Latin America have in place at least electoral democracies, where since 1989 every leader has been elected through a popular mandate. Yet, the *Latinobarometro* surveys register generally a great dissatisfaction with democratic institutions in the region, often claiming more support for the church and the military (see Lagos 1997). Thus, there may not be a correspondence between the formal type of regime that exists within a country and public perceptions of it. Despite these negative observations, the key strength of mass surveys is their representativeness, since in general they tend to use random sampling techniques.

40. Table 1 provides a summary of the description and evaluative statements about the key initiatives in measuring democracy including the initiative, its strengths, and its weaknesses.
<table>
<thead>
<tr>
<th>Initiative</th>
<th>Type of measure</th>
<th>Geographical Coverage</th>
<th>Temporal range</th>
<th>Strengths</th>
<th>Weaknesses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freedom House</td>
<td>Scale-based - 7-point scales of political and civil liberties</td>
<td>Global</td>
<td>1972-present, updated annually</td>
<td>Global coverage; long time-series (since 1972); widely used; updated annually</td>
<td>Lack of transparency in coding; unknown source material mixes economic assessments with political assessments; ideological biases; aggregation problems</td>
</tr>
<tr>
<td>Polity I, II, III, IV</td>
<td>Scale-based – 11-point scales of autocracy and democracy</td>
<td>Global</td>
<td>1800-2000, updated periodically</td>
<td>Global coverage; long time-series (since 1850s); widely used</td>
<td>Narrow institutional focus; largely bi-modal distribution of states</td>
</tr>
<tr>
<td>Banks</td>
<td>Scale-based – 17 point scales of democratic institutions</td>
<td>Global</td>
<td>1850-1997, updated periodically</td>
<td>Global coverage; long time-series (since 1950)</td>
<td>Narrow institutional focus</td>
</tr>
<tr>
<td>Przeworski, et al. 2000</td>
<td>Categories</td>
<td>Global</td>
<td>1950-1990</td>
<td>Global coverage; long time series (since 1950);</td>
<td>Dichotomous categories of democracy and non-democracy; minimum threshold coding difficult to keep consistent.</td>
</tr>
<tr>
<td>Vanhanen (1984, 1990, 1997, 2000)</td>
<td>Objective indicators; Party share &amp; turnout</td>
<td>Global</td>
<td>1810-1998, updated by election year</td>
<td>Global coverage; long time-series (since 1810); objective indicators</td>
<td>Party share does not reflect electoral system or party system; turnout inherently problematic to estimate; data unevenly available by election year; reductionist definition of democracy</td>
</tr>
<tr>
<td>Eurobarometer</td>
<td>Mass public opinion</td>
<td>European Union</td>
<td>1960 – present</td>
<td>Representative random samples;</td>
<td>Limited geographical coverage;</td>
</tr>
<tr>
<td>Survey Name</td>
<td>Type of Opinion</td>
<td>Area Covered</td>
<td>Time Frame</td>
<td>Sampling Method</td>
<td>Data Limitations</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>-----------------</td>
<td>-----------------------------</td>
<td>-----------------------------</td>
<td>-------------------------------------------</td>
<td>----------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>New Europe Barometer</td>
<td>Mass public opinion</td>
<td>Europe (16 post-Communist countries to date)</td>
<td>1991 to present – updated periodically</td>
<td>Representative random samples; battery of standard questions</td>
<td>Limited geographical coverage; public opinion data not measure democracy per se; 'doorstep' opinion; value bias</td>
</tr>
<tr>
<td>Latinobarometer</td>
<td>Mass public opinion</td>
<td>Latin America (17 countries to date)</td>
<td>1995 to present – updated annually</td>
<td>Representative random samples; battery of standard questions</td>
<td>Limited geographical coverage; public opinion data not measure democracy per se; 'doorstep' opinion; value bias</td>
</tr>
<tr>
<td>Afrobarometer</td>
<td>Mass public opinion</td>
<td>Africa (15 countries to date)</td>
<td>1999 to present</td>
<td>Representative random samples; battery of standard questions</td>
<td>Limited geographical coverage; public opinion data not measure democracy per se; 'doorstep' opinion; value bias</td>
</tr>
<tr>
<td>East Asia Barometer</td>
<td>Mass public opinion</td>
<td>East Asia (8 countries to date)</td>
<td>2001 to present</td>
<td>Representative random samples; battery of standard questions</td>
<td>Limited geographical coverage; public opinion data not measure democracy per se; 'doorstep' opinion; value bias</td>
</tr>
<tr>
<td>World Values Surveys</td>
<td>Mass public opinion</td>
<td>20 Countries 43 Countries</td>
<td>1990 1997 updated periodically</td>
<td>Increasing global coverage; representatives samples; battery of standard questions</td>
<td>Limited geographical coverage; public opinion data not measure democracy per se; 'doorstep' opinion; value bias</td>
</tr>
</tbody>
</table>
2.2 HUMAN RIGHTS

Analysis of Main Initiatives for Measuring Human Rights

41. The measurement of human rights can be considered within three streams: human rights monitoring, the contribution of political science analysis and the more recent contribution of development analysis. By and large these initiatives have had separate developments whilst at the same time utilising common information sources. However, in recent years there has been a convergence of these analyses to some extent in both the conceptual and practical domains.

Human Rights Monitoring

42. Human rights monitoring measures both states’ compliance with their human rights obligations (de jure protection of human rights) and violations or enjoyment of human rights in practice (de facto realisation of human rights). Such monitoring is the very rationale of human rights work with an emphasis on holding states accountable and providing redress. Whilst such monitoring is of considerable importance in the measurement of human rights, it does not usually lead to the coding or ranking of countries, and the human rights community has only recently taken up the challenge of developing indicators for specific rights.

43. International human rights law provides the standards against which such monitoring is conducted. When a state accedes to a UN or regional human rights treaty (such as the European Convention on Human Rights) it is obliged to file notification with the appropriate official international bodies. The latter provide regularly updated listings of accessions, ratifications and reservations entered. These lists provide primary source material for initiatives assessing state commitment to international human rights norms (Office of the High Commissioner for Human Rights (OHCHR) www.unhchr.ch), such as those found in the UNDP human development reports, academic analyses (e.g. Keith 1999; Hathaway 2002; Landman 2004), and other sources (www.bayefsky.com).

44. The UN has established a number of monitoring procedures and mechanisms. These include the Commission on Human Rights and its Sub-Commission, a series of bodies established to monitor state compliance with the six core human rights treaties and individual complaints mechanisms. Over the years, the United Nations has also developed an independent and ad hoc system of fact-finding outside the treaty framework, which is referred to as extra-conventional mechanisms or ‘special procedures’ categorised either by country or by theme. At the regional level, the Council of Europe, the Organisation of African States and the Organisation of American States have all developed courts for the determination of allegations of human rights violations. Whilst focussing primarily on individual cases, these mechanisms provide a body of juridical information concerning violations, which provides primary source materials for measuring human rights performance.

45. In recent years the UN has mounted human rights field operations in a number of countries suffering gross violations. These operations have combined fact-finding with action to address the human rights problems. Fact-finding entails widespread information-gathering in order to establish and verify the facts surrounding an alleged
human rights violation. The UN is pursuing reliability through the use of generally accepted objective human rights data-collection procedures, such as interviews with concerned parties, observing events (elections, trials, demonstrations etc.), visiting sites such as places of detention and refugee camps, and verification through the requirement of corroborating evidence. A *Training Manual on Human Rights Monitoring*, developed by Minnesota Law School, builds on the experience to date and sets out the methodology to be used by UN human rights officers. (www1.umn.edu/humanrts/monitoring).

46. The Committee on Economic Social and Cultural Rights, which monitors states’ compliance with their obligations under the International Covenant on Economic, Social and Cultural Rights, has for some years been encouraging the development of appropriate indicators for measuring this branch of rights. In 2001, the Committee requested the Office of the High Commissioner for Human Rights (OHCHR) to develop guidelines for the integration of human rights into poverty reduction strategies. The draft guidelines analyse the human rights relevant in poverty reduction and propose indicators for measuring specific rights (OHCHR 2002). In a recent draft paper, Malhotra provides a listing of potential indicators for measuring selected economic, social and cultural rights, relying heavily on existing development indicators. He categorises these as input, process or outcome indicators (see discussion below and Malhotra 2002).

47. A number of other UN agencies are also engaged in aspects of human rights measurement, for example, the ILO (labour rights), UNESCO (education rights) UNICEF (children’s rights) and UNFPA (reproductive health rights). The ILO has produced a draft working paper proposing the development of indicators for measuring numerically the gap between the real and the ideal world of basic workers’ rights with the help of the ratification, reporting, supervisory, complaints and financial information at the disposal of the ILO (Böhning 2003).

48. Monitoring human rights is also undertaken at the national level. Many countries have established human rights commissions, specialised commissions aimed at protecting the rights of vulnerable groups or ombudsmen/women and national human rights institutions, which aim to investigate and/or document human rights cases and sometimes also trends. In addition, other national bodies may measure some aspects of human rights such as allegations of racial harassment or police brutality.

49. Worldwide, many documentation centres have been established comprising the collection of published documents and organising these following established techniques in librarianship such as cataloguing. In the 1980s, HURIDOCS developed the concept of ‘controlled vocabularies’ and ‘standard format’ tools for documenting human rights that enable data to be recorded in a systematic way. For example, the Events Standard Formats assist in the standardisation of the investigation and recording of human rights violations (HURIDOCS 1985, 1993). Such standard formats enable the compilation of comprehensive data that can record in minute detail both individual violations and patterns of gross violations within a country. This system is now widely used by the Council of Europe, UN agencies (e.g. the UN High Commissioner for Refugees) and NGOs (e.g. Amnesty International).
50. Collaboration between HURIDOCS and the Science and Human Rights Program of the American Association for the Advancement of Science (AAAS) has led to the further refinement of the ‘events’ methodology and the publication in 1994 of a *Definition of Database Design Standards for Human Rights Agencies* (Ball 1994). Together AAAS and HURIDOCS have also been developing since 1996 a ‘violations approach’ to monitoring economic, social and cultural rights (AAAS http://shr.aaas.org/escr) which complements the ‘progressive realisation approach’ advocated by organisations such as Rights and Humanity since its establishment in 1986.

51. There are numerous NGOs engaged in the monitoring of human rights, the majority of which are using events-based methodologies. A review of the initiatives studied indicates that this is the preferred methodology for monitoring violations of civil and political rights. It involves identifying the various acts of commission and omission that constitute or lead to human rights violations, such as extra-judicial killings, arbitrary arrest or torture. Events and violations-based measures count the number of violations per country over time. Such data tend to be disaggregated to the level of the violation itself, which may have related data units such as the perpetrator, the victim, and the witness (Ball, Spirer, and Spirer 2000). On the other hand, the methodology developed by AAAS/HURIDOCS to measure violations of economic, social and cultural rights relies less on events-based monitoring than on measures designed to identify whether a state has complied with its duty to provide the core minimum content of the rights (Chapman 1996, 1998; AAAS http://shr.aaas.org/thesaurus/ June 2002).

52. Human rights event-based monitoring tends to focus on violations of particular rights - such as freedom of expression (Article 19 www.article19.org); press freedom (Index on Censorship www.indexoncensorship.org); violations of housing rights such as forced evictions (Habitat International Coalition http://home.mweb.co.za/hic/index.html) - or on violations suffered by particular groups - such as minorities (Minorities At Risk Project, Gurr 1993, 2000); children (Defence for Children International www.defence-for-children.org). Some organisations focus on both themes and groups (Amnesty International - torture and the death penalty and, more recently, women’s rights www.amnesty.org).

53. There has been an additional impetus in the collection of data on violations within the context of international criminal courts and truth and reconciliation processes. Patrick Ball, and Herbert and Louise Spirer, have developed a technique of cross-checking original information on gross violations by using multiple sources of information, including governmental sources (such as border controls), UN agencies, the Red Cross and grass roots NGOs. Their data have been collected and analysed primarily for purposes of providing testimony to truth and reconciliation commissions and international tribunals, such as those for the Former Yugoslavia and Rwanda. The data are in-depth and highly disaggregated in terms of the units of analysis (the individual violation) and the time periods (days) (see shr.aaas.org/hrdag.idea/msm/index.html).

54. An example of a new initiative, piloted in 2002, combines several methodologies. The American Bar Association’s Central European and Eurasian Law Initiative (ABA/CEELI) seeks to monitor women’s human rights protected by the Convention on the Elimination of All Forms of Discrimination Against Women by assessing, article by article, both *de jure* and *de facto* state compliance. The *de jure* analysis is
undertaken by a national assessment team including lawyers and NGO representatives, whereas the *de facto* analysis combines interviews, focus groups and other fact-finding techniques, such as expert local opinion (www.abanet.org/ceeli/home.html).

**The Contribution of Political Science Analysis**

55. Within the political science field, there have been a number of initiatives to code this primary data into standards-based scales. In addition, aspects of civil and political rights were included in some early attempts to measure democracy, not as human rights *per se*, but as crucial components of democracy. Similarly, early attempts to measure good governance used rights measures as proxies for the rule of law, which gave way to measures of the bureaucratic burden and political risks for doing business, but then returned to including the protection of human rights as part of combined indices. Outside these efforts, the political stability indicators used in the work on good governance (see below) have featured in human rights measurement, with the added indicators on the use of state coercion and ‘state-sponsored terrorism’. These two indicators serve as proxy indicators for civil and political rights measurement.

56. Raymond Gastil first began providing political and civil liberties measures in 1972, the coding and production of which was taken over by Freedom House in 1989. In 1979, Gastil also devised an early system for such measurement, which was taken up by Michael Stohl in 1983, who created the ‘political terror scale’, currently one of the most enduring measures in academic research.

57. The political science material tends to rely on two types of standard-based scales to measure human rights. Initiatives within the first type measure *de jure* state compliance with human rights obligations through counting the ratification of international or regional human rights legal texts, and/or the incorporation of human rights into the national constitution or other laws. The first large-scale academic effort to code national constitutions for their rights provisions was carried out by van Maarseveen and van der Tang (1978), an effort that was followed more than a decade later by Suksi (1993). More recently, Poe and Keith (2002) have coded national constitutions for ‘states of emergency’ clauses that when enacted abrogate their human rights commitments. At the international level, new efforts have emerged to code countries for signing and ratifying international human rights treaties in order to test compliance (Keith 1999; Hathaway 2002; Landman 2001, 2002). In some initiatives weighting is added for reservations entered by states upon ratification of international treaties (Sano and Lindholt 2000, Landman 2004).

58. The second type of standard-based scales is the coding of the *de facto* situation within a country, either generally, or with respect to a particular human right. Within civil and political rights, one of the most used and dominant of this type of measurement is the ‘political terror scale’ or the measure of ‘personal integrity rights’ protection (see above and Poe and Tate 1994). This is a five-point scale that codes the degree to which rights are protected in countries using the descriptive human rights information contained in the Amnesty International *Annual Reports* and the US State Department *Country Reports*. To date, the coverage is global for the period 1976-1996. The human rights violations included in the political terror scale include political imprisonment, torture, extra-judicial killing, and disappearances.
The methodology for the political terror scale has been replicated to measure the situation of women’s rights, minority rights, torture, and worker’s rights. The women’s rights scales provide two separate scores for women’s economic and political equality for 1994 only, but could be updated (Poe, Wendl-Blunt, and Ho 1997). The ‘minorities at risk’ project provides scales of discrimination against 285 different numerical dominant and non-dominant minorities (Gurr 1993, 2000). The torture scale measures the degree to which torture is regularly practiced (Hathaway 2002). The workers’ rights scale measures the degree to which governments protect the right to association, collective bargaining, the elimination of forced or compulsory labour, the effective abolition of child labour, and acceptable conditions of work with respect to wages, hours, and occupational health and safety (Cingranelli and Tsai 2002: 24). While these scales are in the tradition of the Freedom House scales of political and civil liberties, their focus is narrower and their coding procedures and reliability are more robust (see below).

Finally, some initiatives use survey-based data to measure popular perceptions of human rights protection. The mass surveys that measure the quality of democracy and democratic institutions contain specific questions about rights protection. Thus, the World Values Surveys, the Global Barometer Surveys and the Gallup Millennium Survey all contain questions relevant to human rights protection. In addition, there have been sporadic efforts to produce expert perception indices on human rights protection.

Contribution of Development Analysis

The last decade has seen a convergence of the human rights and development agenda (Häusermann 1997, 1998). The adoption of the human rights approach to development by the UN and a number of development donors has led to a useful cross-fertilisation of ideas and techniques. There is now an attempt to integrate human rights concerns into development indicators, whilst at the same time the human rights community is utilising human development indicators to assist in the measurement of human rights, particularly economic, social and cultural rights.

Probably one of the most significant initiatives is the introduction of a human rights component into the UN Common Country Assessment/Development Assistance Framework Guidelines, first introduced in 1997 and revised in May 2002. The CCA is a country-based process for reviewing and analysing the national development situation and is jointly undertaken by UN agencies in cooperation with government and other key partners. As the revised guidelines recognise, approaching development from the perspective of human rights creates particular demands for data that are not satisfied by traditional socio-economic indicators alone. For example, it requires that a rights element is integrated into existing indicators by identifying (i) explicit standards and benchmarks against which to measure performance, (ii) specific actors or institutions responsible for performance, (iii) rights-holders to whom responsibility is owed, and (iv) mechanisms for delivery, accountability, and redress. In addition, it envisages the measurement of certain subjective elements, such as levels of public confidence in institutions of governance, including among vulnerable or marginalized groups (CCA/UNDAF Integrated Guidelines 2002).

The CCA/UNDAF guidelines require that all relevant indicators should be disaggregated, to the greatest extent possible and where appropriate, by race, colour,
sex, language, religion, nationality, ethnic or social origin, property, disability or other status (such as being a woman or child head of a household). The CCA utilises an indicator framework comprising indicators relating to UN conference goals, such as the Millennium Development Goals (MDGs); conference and convention indicators relating to governance, democracy, justice administration and security of person; basic contextual indicators relating to the demographic and economic conditions of the country; indicators for monitoring implementation of the MDG requiring the development of a global partnership for development; and thematic indicators to provide further insights into issues of major concern for development (CCA/UNDAF Integrated Guidelines 2002).

64. At the same time, development indicators are increasingly being employed as proxy indicators for measuring the progressive realisation of economic, social and cultural rights (Malhotra 2002 cited above). Article 2 of the International Covenant on Economic, Social and Cultural Rights requires states to take steps, to the maximum of their available resources, towards the progressive realisation of these rights. This requires states to set goals, targets and timeframes for national plans to implement these rights. Development indicators are thus seen as suitable proxy measures to capture the degree to which states are implementing these obligations. For example, literacy rates and gender breakdown of educational attainment are seen as proxy measures of the right to education; daily per capita supply of calories and other nutritional rates are seen as proxy measures of the right to food; and under-five mortality rates and the numbers of doctors per capita are seen as proxy measures of the right to health (OHCHR 2002).

65. To date, development indicators have primarily been applied to economic and social rights, but following the work of the United States Agency for International Development (USAID), new efforts propose the use of development indicators as potential proxy measures for civil and political rights. For example, investment in prison and police reform, the processing of cases, and the funding of judiciaries are all seen as proxy measures for state commitment to upholding civil and political rights. Such information is available from national statistical offices, which with the appropriate sensitisation, could think about how to use such information to provide indicators for civil and political rights.

66. The NGO Rights and Humanity is involved in the development of methodologies and indicators to measure the contribution of a human rights approach to the achievement of human development (Rights and Humanity 2001, 2002). It aims to identify a few core indicators to measure the impact of the integration into development cooperation of such human rights principles as popular participation in the development process, empowerment of poor and marginalised groups, freedom from discrimination and government accountability.

**Evaluation of Main Initiatives for Measuring Human Rights**

67. As has been shown, there is much greater academic and global consensus on the enumeration of human rights than there is on the definitions of democracy and good governance and, not surprisingly, the initiatives measuring human rights are worldwide and numerous. Although the human rights community has not traditionally used the
language of indicators, the monitoring of civil and political rights is in fact based on the use of events-based indicators. For instance, counting the number of incidents of torture is an appropriate indicator of violations of the right to freedom from torture. However, there remain a number of methodological challenges in events-based monitoring of human rights. Particularly in the early days, difficulty was sometimes experienced in recording events that carried multiple acts, for example, an event involving the arrest, torture and illegal detention of the same person. In early monitoring efforts there was a practice for many organisations to pick the gravest kind of violation and classify the event only as such a violation (Guzman 2000).

68. Conversely, many single instances of violence involve numerous victims and/or numerous violators leading to recording difficulties and subsequent problems in generating frequencies of the incidence of political violence (Cheater 2000). As AAAS identifies, such overlapping information contains extremely valuable statistical information that can be used for multiple systems estimation. Instead of deleting the overlapping information, it recommends that judgments are made about what is truly unique and what interviews describe the same events and violations. This process is called creating a ‘judgment layer’ on top of the many sources (AAAS Human Rights Data Analysis Group http://shr.aaas.org/hrdag/idea/datamodel/).

69. There has been a marked tendency for such initiatives to focus on violations of a few civil rights, such as torture and other grave breaches. However, it seems equally plausible and desirable to devise violations approaches to measure political rights (e.g. intimidation of voting). Although some of the initial methodological difficulties have been resolved, the data are not at present available for large samples of countries, nor are they available for those countries without grave human rights situations.

70. Despite the broad agreement on the fact that all human rights are indivisible and should be accorded equal priority, the initiatives studied illustrate that far greater progress has been made on developing indicators to measure civil and political rights than has been made on measuring economic, social and cultural rights. This reflects the general tendency during early human rights monitoring to favour civil and political rights at the cost of economic, social and cultural rights. But this situation is also partly due to the lack of clarity that remains concerning the precise scope of state obligation and the core content of individual economic, social and cultural rights, which in turn makes it difficult to identify events that clearly amount to violations. The AAAS/HURIDOCS initiative reviewed above is contributing to overcoming this obscurity. The efforts by ABA/CEELI to provide de facto analysis of women’s rights violations are still quite limited since they advocate interviewing a relatively small sample of respondents (N ≅ 30), making the generation of strong inferences problematic.

71. As noted above, in the absence of clear economic and social rights indicators, attention is being paid to using or adapting development indicators. The strength in using development indicators for human rights measurement lies in their regular availability, global coverage, ease of understanding and long time-series (see UNDP’s human

---

4 Stems from the ideological differences between East and West during the Cold War.

5 Some progress is being made in this regard by the Committee on Economic, Social and Cultural Rights (see, for example, its General Comment No 3).
development indicators and the World Bank’s online database which has itemised common development indicators since 1950).

72. A few development indicators have a good correlation with indicators of human rights. For instance, the contraceptive knowledge indicators developed and utilised by UNFPA are good indicators for measuring part of the reproductive health rights confirmed by the International Conference on Population and Development (Cairo, 1992). These indicators are derived from sample survey reports and estimate the proportion of women who have knowledge of a method of family planning and know a source from which contraceptives can be obtained (United Nations Population Division 1996).

73. The key weakness in using development indicators for human rights measurement is the serious question of their validity. For example, using literacy rates, educational attainment, gender breakdown in education, and investment in schools only demonstrates the state of education in a particular country not whether the right to freedom from discrimination in education is being upheld. Even among the most advanced industrial democracies of the West (which would score high on all development indicators for education listed above), there are problems with discrimination. The recent report of the Democratic Audit of the United Kingdom demonstrates ‘the persistence of de facto discrimination of some marginalized and vulnerable groups in society, especially ethnic minorities and people with disabilities, in employment, housing, education, and other areas of life…’ (Beetham, Byrne, Ngan, and Weir 2002: page 60). Such discrimination must be captured in order to be a valid indicator of the right to education. Development indicators do not always measure what some are claiming they measure.

74. The ILO initiative on workers’ rights is also considered to leave a distinction between what would actually be measured and what the initiative claims would be measured. A draft working paper (revision 1) prepared by the ILO proposes a sophisticated system of coding in order to rank state parties to the core ILO labour conventions in terms of the gap between their willingness to adhere to the rights covered and the practical implementation of workers’ rights (Böhning 2003). The benefit of this initiative is that it plans to draw on material already available to the ILO covering the period since the mid-1980s. This work would be a positive step towards greater transparency in at least part of the ILO reporting process. As such the initiative might also act as a catalyst to encourage the further development of some sort of indicator scores for reporting to the Treaty Bodies.

75. However, the data used for the coding is limited to such sources as: the listings of ratification and reservations, reports submitted to the ILO monitoring mechanism, comments made by the monitoring bodies and information indicating the sensitivity of the state to such comments. The initiative does not include any external source of information (from trade unions, companies, NGOs or other sources) on the actual situation relating to workers’ rights in the country. This can be considered a weakness, particularly in view of the fact that the original reports to the monitoring mechanisms are prepared by states. Furthermore, the degree to which additional non-governmental information was available to the monitoring body at the time it reviewed the state report is likely to have been varied between states. This might undermine the validity of ranking states.
An evaluation of other standards-based scales reviewed above indicates that they share between themselves a number of strengths and weaknesses. The shared strengths of the political terror scale and the associated scales on torture, women’s rights, and worker’s rights include global coverage, long time-series (with the exception of the women’s rights indicators), common source material for coding (Amnesty International and US State Department), transparent coding rules and robust inter-coder reliability, and well-defined sets of human rights.

The ‘minorities at risk’ project is a unique effort to study the state of minority groups around the globe. Its strength lies in its focus on minority groups and, in particular, discrimination against such groups. The methodology shows promise for developing indicators of violations approaches to the measurement of economic, social, cultural, and women’s rights since it has developed ways in which to measure discriminatory practices.

Their shared weaknesses include variance truncation (they lack sensitivity at both ends of the scales), they are not updated annually or sufficiently frequently for timely policy analysis and they cannot be used for in-depth country analysis and political dialogue. A major difficulty with using the standard-based scales developed within the discourse of political science as a basis for measuring human rights, is that such scales are not necessarily related to the ideal or standards recognised within international human rights law. Within the context of human rights, ‘standards’ are directly linked to international legal norms yet a number of the standard-based scales reviewed above are not expressly linked to these legal standards. Whatever their strengths and weakness as measures of democracy and good governance (or the rule of law), standard-based scales such as that developed by the Freedom House cannot be regarded as a credible indicators of human rights per se.

As we have seen, the measurement of human rights has to some extent proceeded in a piecemeal fashion with indicators for certain rights (freedom from torture, the right to physical integrity etc.) being well developed, whilst the development of indicators for the wide range of other rights contained in international human rights texts mains weak. Table 2 provides descriptive information on and an evaluation of the main initiatives measuring human rights, the type of measure, its geographical coverage, and its temporal range. The table marks with an ‘*’ those human rights measures that have been considered by the team to be ‘winners’ for having stood the test of time and/or provided particularly good geographical and temporal coverage. These initiatives are also marked as such in the database.
<table>
<thead>
<tr>
<th>Initiative</th>
<th>Type of measure</th>
<th>Geographical Coverage</th>
<th>Temporal range</th>
<th>Strengths</th>
<th>Weaknesses</th>
</tr>
</thead>
<tbody>
<tr>
<td>ILO (Böhning 2002)</td>
<td>De jure and de facto labour rights using official reporting mechanisms</td>
<td>Global</td>
<td>2002, but possible for 1980s to the present</td>
<td>Transparent coding rules and coding sources; good potential time-series</td>
<td>Limited information base; no independent sources of information</td>
</tr>
<tr>
<td>Freedom House*</td>
<td>Standards-based scales for political and civil liberties</td>
<td>Global</td>
<td>1972-present</td>
<td>Global coverage; long time-series (since 1972); widely used; updated annually</td>
<td>Lack of transparency in coding; unknown source material mixes economic assessments with political assessments; ideological biases; aggregation problems</td>
</tr>
<tr>
<td>Political Terror Scale*</td>
<td>Standards-based scales for 'personal integrity rights'</td>
<td>Global</td>
<td>1976-1996</td>
<td>Global coverage; long time series (1976-1996); common source material (Amnesty International; US State Department); transparent coding rules; robust inter-coder reliability; explicit set of human rights.</td>
<td>Variance truncation; not updated frequently</td>
</tr>
<tr>
<td>Torture scale*</td>
<td>Standards-based scales for human rights</td>
<td>Global</td>
<td>1985-1999</td>
<td>Global coverage; long time series (1985-1999); common source material (Amnesty International; US State Department); transparent coding rules; robust inter-coder reliability; explicit set of human rights.</td>
<td>Variance truncation; not updated frequently</td>
</tr>
<tr>
<td>Dataset Name</td>
<td>Description</td>
<td>Global Coverage</td>
<td>Time Period</td>
<td>Update Frequency</td>
<td></td>
</tr>
<tr>
<td>---------------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>-----------------</td>
<td>------------------------------</td>
<td>---------------------------------</td>
<td></td>
</tr>
<tr>
<td>(Hathaway 2002)</td>
<td>the degree that torture is practiced</td>
<td>Common source material (Amnesty International; US State Department); transparent coding rules; robust inter-coder reliability; explicit single human right.</td>
<td>Updated frequently.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women’s Rights*</td>
<td>Standards-based scales on the economic and political equality of women</td>
<td>Global coverage; common source material (Amnesty International; US State Department); transparent coding rules; robust inter-coder reliability; explicit set of women’s human rights</td>
<td>1994, but can be updated</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Poe, Wendel-Blunt and Ho 1997)</td>
<td></td>
<td></td>
<td></td>
<td>Variance truncation; single year available only.</td>
<td></td>
</tr>
<tr>
<td>Worker’s Rights*</td>
<td>Standards-based scales for a series of core workers rights</td>
<td>Global coverage; long time series (1981-1999); common source material (Amnesty International; US State Department); transparent coding rules; robust inter-coder reliability; explicit set of women’s human rights</td>
<td>1981-1999</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Cingranelli and Tsai 2002)</td>
<td></td>
<td></td>
<td></td>
<td>Variance truncation; not frequently updated</td>
<td></td>
</tr>
<tr>
<td>Minorities at Risk*</td>
<td>Abstract scales of cultural discrimination against 285 numerically dominant and non-dominant minorities</td>
<td>Global coverage; specific focus on minority rights and discrimination</td>
<td>Annual and quinquennial data for the 1990s</td>
<td>Variance truncation; limited time-series; annual and quinquennial data for the 1990s</td>
<td></td>
</tr>
<tr>
<td>Amnesty International</td>
<td>Events-based or violations-based monitoring of human rights</td>
<td>Global coverage; cross-checked reporting</td>
<td>1977 – present, updated annually</td>
<td>Initial focus was narrow but later years expand mandate to include a fuller range of human rights; monitoring not measuring</td>
<td></td>
</tr>
<tr>
<td>Human Rights Watch</td>
<td>Events-based or violations-based monitoring of human rights</td>
<td>Global coverage; cross-checking</td>
<td>1989 – present, updated</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organization</td>
<td>Monitoring Type</td>
<td>Coverage</td>
<td>Time Frame</td>
<td>Details</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>----------------</td>
<td>----------</td>
<td>------------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>US State Department</td>
<td>Events-based or violations-based monitoring of human rights</td>
<td>Global</td>
<td>1985-present, updated every two years</td>
<td>Global coverage; detailed reporting</td>
<td></td>
</tr>
<tr>
<td>HURIDOCS</td>
<td>Events-based or violations-based monitoring of human rights violations; standardised reporting format</td>
<td>Global application</td>
<td>Set up in 1982 and now a global network of human rights NGOs monitoring human rights violations</td>
<td>Tool for reporting; influenced many NGOs</td>
<td></td>
</tr>
<tr>
<td>FIDH-OMCT</td>
<td>Events-based or violations-based monitoring of repression and harassment of human rights defenders</td>
<td>Global</td>
<td>1997-present, updated annually</td>
<td>Global coverage; cross-checked reporting</td>
<td></td>
</tr>
<tr>
<td>AAAS</td>
<td>Events-based or violations-based indicators of gross violations</td>
<td>Country-specific</td>
<td>Country specific</td>
<td>In-depth event counts; cross-checked source data; disaggregated data (units and time); practical application in international tribunals and Truth and Reconciliation Commissions</td>
<td></td>
</tr>
<tr>
<td>UNDP, UNICEF, OHCHR, UNFPA, UNESCO</td>
<td>Development indicators as human rights indicators</td>
<td>Global</td>
<td>1950s – present</td>
<td>Global coverage; good time-series; readily available</td>
<td></td>
</tr>
</tbody>
</table>

Problems of validity; development indicators are not human rights indicators; tend to focus on economic and social rights; now includes some economic and social rights; monitoring not measuring; compiled by diplomatic staff; traditionally narrow focus on civil and political rights; complex thesaurus of terms and categories that are not immediately accessible; monitoring not measuring; only focuses on human rights defenders; short time-series; monitoring not measuring.
*Those marked are considered main initiatives that have stood the test of time.

rights; tension between progressive realisation and violations approaches to measuring
2.3 GOOD GOVERNANCE

*Analysis of Main Initiatives for Measuring Good Governance*

80. The measurement of good governance has a shorter history than the measurement of democracy and human rights, but since the 1980s academics, IGOs, and NGOs (as well as private sector companies) have been developing different indicators of good governance. The project team found five types of measures of good governance. These include: (1) civil and political liberties or political freedoms as proxy measures for the rule of law, (2) the frequency of political violence as an inverse measure of good governance, (3) expert assessments and opinion of good governance for investment (see discussion of democratic ‘image indices’ above), (4) objective indicators such as ‘contract intensive money’ (CIM) as a measure of individual confidence in the domestic financial institutions (Claugue et al. 1995, 1999; Knack 2002) or the economic rate of return (ERR) of governmental projects (Isham et al. 1997), and (5) mixed measures that combine aggregate data, scales, and expert opinion (e.g. Kaufmann et al. 1999, 2002, 2003).

81. The early work on good governance sought to investigate its effect on economic performance, and the natural instinct was to use so-called proxy measures of the political and legal institutions in terms of liberties and freedom from state constraint. Thus, Freedom House scales have been used to approximate the rule of law and governance, as well as the ‘institutional framework’ (Scully 1988). There has thus been an overlap in usage of the Freedom House scales as an indicator of democracy, human rights and good governance. The main response to the use of Freedom House has been by the Heritage Foundation and the Wall Street Journal, who produce annual scores of economic freedom for 161 countries from 1995-2003, and the Fraser Institute’s Index of Economic Freedom.

82. One response to the dissatisfaction with using Freedom House has been to use political violence indicators (revolutions, military coups, and political assassinations) as proxy measures for political instability, or ‘bad’ governance. Such an indicator is seen to represent the variable respect for property rights (Barro 1991), which economists feel is a central feature of good governance (see the conceptual discussion in Annex III). High levels of political violence mean a low respect for property rights. Behaviouralists have long established global databases containing political stability indicators, including riots, political assassinations, military coups, demonstrations, etc. Russet (1964), Taylor and Hudson (1972), Taylor and Jodice (1983), and Banks (1994; 1997) are the key examples of these kinds of data (see paragraph 21 above).

83. Another alternative to the use of Freedom House scales and political violence indicators are methodologies that poll ‘expert’ opinion on the business and economic environment of countries, which were developed by private companies in the 1970s. Organisations such as Business International, the Political Risk Services, Business Environmental Risk Intelligence (BERI), and Transparency International produce scales that rank countries across a range of indicators, including political credibility, the credibility of rules, the development of the ‘social infrastructure’, the legal system and the judiciary, political stability, the risk of expropriation, repudiation of contracts,
and the prevalence of corruption, influence, and so-called ‘state capture’. In each instance, the organisations are using well-placed business elites to arrive at a general assessment of the government climate in which they do business. They are thus subjective measures of good governance (like the ‘image indices’ of democracy or the de facto rights measures discussed above) and the general public must pay for access to the country rankings. All or selected parts of these measures have been used by the academic community (e.g. Knack and Keefer 1995; Mauro 1995). Outside these private organisations, The United Nations University World Governance Survey Project (WGS), a collaborative effort between academics and the UNDP, has used expert opinion surveys on good governance for 16 developing and transitional countries (Court, Hyden, and Mease 2002).

84. There are also objective measures of good governance. One example is called ‘contract-intensive money’, or CIM (Clague et al. 1995, 1997, 1999), which is in formal terms, the ratio of non-currency money to the total money supply (CIM = \([M2-C]/M2\)). The figures used to calculate the CIM are available from the IMF. The logic of using CIM as a measure of good governance is that ‘individuals will hold a larger proportion of their financial assets in the form of currency in environments where third-party enforcement of contracts is unreliable’ (Knack 2002:12). In other words, CIM is an objective indicator with wide geographical and temporal coverage that measures general faith in economic and political institutions, as well as the enforceability of contracts.

85. Another good governance measure seeks to maximise the use of a broad range of available indicators on good governance through a data reduction technique called ‘unobserved components model’ (a variant of factor analysis) to combine up to 300 disparate indicators of good governance into six separate indices. These separate indices include: (1) voice and accountability, (2) political instability and violence, (3) government effectiveness, (4) regulatory burden, (5) rule of law, and (6) graft. Produced by Kaufmann, Kraay, and Zoido-Lobatón (KKZ) at the World Bank, these indices of good governance cover over 160 countries for 1996-2002.

86. Of these measures of good governance, both the scale-based measures of economic freedom (Heritage Foundation/Wall Street Journal) and the expert opinion measures have survived the test of time, while there is an emerging hegemony of the combined scores, as well as the call for the use of objective indicators such as the CIM. Table 2 provides descriptive information on these key initiatives in the measurement of good governance, including the source of the initiative, the type of measurement, its geographical coverage, and its temporal range.

**Evaluation of Main Initiatives to Measure Good Governance**

87. The main initiatives on measuring good governance include scale-based indicators, expert opinion indices, and combined measures. Like the measures of democracy and human rights, each initiative has its own strengths and weaknesses, where the more recent attempts at measuring good governance strive for robustness. These strengths and weaknesses are summarised in Table 3, and those initiative identified in the table are flagged as ‘winners’ in the database as having stood the test of time and/or provided good geographical and temporal coverage.
88. In the absence of good governance measures, scholars initially used Freedom House as a proxy measure of the rule of law, but since then many have criticised its validity on the grounds that whether or not it is a measure of individual freedom, it is certainly not the rule of law *per se* (Knack 2002). Moreover, as the concept of good governance has expanded, the rule of law is now seen as only one component out of many (see Annex III). Yet, Freedom House continues to feature as one of the many indices that KKZ combine in their unobserved components model.

89. The big 'growth area' in good governance indicators has arguably been in production of expert opinion polls. These use very small samples of well-placed individuals that offer their expert opinion on a range of topics related to good governance, including bureaucratic delays and red tape, corruption and graft, state capture, rent seeking and so forth. Despite their wide use, regular production, and good geographical coverage, these measures all suffer from severe selection bias, since they only use very small samples of individuals (usually business elites and entrepreneurs). They also suffer from systematic measurement error, since the experts tend to contaminate their assessment of 'doing business' in the country with the underlying economic situation in the first place. Thus, many of these measures are offering expert opinion on the general state of the economy and not on the nature and quality of good governance.

90. The initiative to use objective indicators such as contract intensive money holds much promise. The data are regularly available from the IMF and have good geographical coverage. The measure is simple to calculate and simple to understand, and it has good distributional attributes for secondary statistical analysis. The key question for using CIM as an indicator for good governance is whether or not it is a valid measure. First, since its use rests on an assumption about consumer behaviour under different banking and regulatory systems, does it really measure good governance? Second, if the starting assumption of using the CIM can be upheld, does it capture enough of the concept of good governance? It is possible for a country to have a good banking system in which consumers have great confidence, yet the government may fail in other areas of governance. Third, is it sensitive to the contextual and cultural specificities of different countries? Its use is predicated on a fairly Western ethnocentric understanding of private savings and private consumer behaviour. It is not at all clear that private savings is necessarily a reflection of faith in the banking system, but may be the product of other underlying cultural factors. Thus, users of CIM need to convince their audience that it is not only an indicator of good governance, but also one that captures most of its substantive content and reflects the different habits of consumers.

91. Finally, the combined measures of Kaufmann, Kraay, and Zoido-Lobaton [KKZ] have achieved a new prominence in the literature on good governance. They claim great robustness for their measures since they are using over 300 different indicators to produce their combined indices across six aspects of good governance. To date, they have produced the measure for 1996-2002 for a global sample of countries. The data are readily available and the method for producing the indices is a straightforward data reduction technique called 'unobserved components model' which is a variant of factor analysis that extracts components common to the 300 different measures in the original data. In their case, six such components have been extracted. They argue that using so many original indicators increases both the validity and reliability of the indices, effectively reducing the chances for systematic measurement error.
Despite these main claims for the superiority of the KKZ indices of good governance, there are remaining methodological problems. First, the 300 indicators do not cover all the countries in the global sample such that different indicators cover different groups of countries. Thus, the combined score on one dimension of good governance for country X will be comprised of a different set of initial indicators than the combined score for country Y. Both countries may share some but not all of the same initial indicators. In effect, the scores are combining 'apples' and 'oranges' and calling them 'pears'. Second, the indices that are produced are point estimates with associated standard errors or confidence intervals. KKZ (2002) themselves argue that the margin of error around the point estimates are so large that it is impossible to rank-order the middle 'rump' of countries in the sample, which precludes more advanced secondary statistical analysis. This suggests that after combining over 300 initial indicators, the KKZ indices can only really differentiate among countries with low, middle, and high good governance scores.
<table>
<thead>
<tr>
<th>Initiative</th>
<th>Type of measure</th>
<th>Geographical Coverage</th>
<th>Temporal range</th>
<th>Strengths</th>
<th>Weaknesses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freedom House</td>
<td>Scale-based indicators for political and civil liberties</td>
<td>Global</td>
<td>1972-present</td>
<td>Global coverage; long time-series (since 1972); widely used; updated annually</td>
<td>Lack of transparency in coding; unknown source material mixes economic assessments with political assessments; ideological biases; aggregation problems</td>
</tr>
<tr>
<td>Fraser Institute</td>
<td>Standards-based scales</td>
<td>123 Countries</td>
<td>1975-2002</td>
<td>Global coverage; good time-series (1975-2002); used widely</td>
<td>Mixes objective and subjective measures</td>
</tr>
<tr>
<td>Political Risk Services</td>
<td>Expert opinion on corruption, rule of law, expropriation risk, repudiation of government contracts, quality of the bureaucracy.</td>
<td>140 countries</td>
<td>1982-2003, updated annually</td>
<td>Global coverage; reasonable time-series (1982-2003); updated annually</td>
<td>Unrepresentative sample of opinion; scores may represent underlying economic conditions; narrow measure of good governance</td>
</tr>
<tr>
<td>Business Environmental Risk Intelligence (BERI)</td>
<td>Expert opinion: political risk index</td>
<td>50 countries</td>
<td>1972 to the present, updated three times a year</td>
<td>50 countries</td>
<td>1966 to the present; lack of transparency in coding</td>
</tr>
<tr>
<td>World Economic Forum</td>
<td>Competitiveness scale</td>
<td>Global</td>
<td>1979-2002 updated annually</td>
<td>Global coverage; good time-series (1979-2002);</td>
<td>Mixes objective and subjective indicators (i.e.</td>
</tr>
<tr>
<td>Index Name</td>
<td>Description</td>
<td>Coverage</td>
<td>Time Period</td>
<td>Advantages</td>
<td>Disadvantages</td>
</tr>
<tr>
<td>-------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-------------------</td>
<td>--------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Combined indices KKZ</td>
<td>Unobserved components data reduction of 300 different indicators on six aspects of good governance</td>
<td>Global</td>
<td>1996-2002</td>
<td>Objective indicators; updated annually</td>
<td>Narrow proxy measure of good governance; culturally biased.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Global coverage; robust number of indicators that are combined</td>
<td>Short time-series (1997-1998); combines different sets of indicators; point estimates so insecure that middle 'rump' of countries cannot be ranked</td>
</tr>
</tbody>
</table>
PART 3 - CONCLUSIONS AND RECOMMENDATIONS

3.1 CONCLUSIONS

93. The results of this study illustrate that across all three areas of work a good degree of progress has been made. The generation of primary source material, secondary analysis and methodological discussion of extant work demonstrate greater attention to the production of robust measures. There has been a general call for greater transparency in the production and use of measures. There has also been a move towards a more scientific approach in measurement techniques, which is particularly marked in the progress from basic human rights monitoring to more advanced human rights measurement. Such a scientific approach recognises the importance of credible sources of accurate information, preferably cross-checked. It places emphasis on ensuring that the indicators used are relevant to the concept being measured, which itself entails clarity in conceptualisation. The methodology needs to be reliable and replicable.

94. Despite the progress made, a major difficulty remains in the lack of awareness about the rich sources of credible primary source material, particularly in the measurement of human rights and good governance. As a consequence there is a tendency to reuse data produced for other purposes, and an over-reliance on a few source materials such as that produced by Amnesty International, Freedom House and the US State Department. This problem is exacerbated by a tendency of IGOs to collate material from a number of sources and reproduce it uncritically within their own analysis, effectively giving it another layer of credence.

95. It is clear from the study that extant indicators have been produced for a variety of purposes. Failure to recognise this specificity has led to their inappropriate use in some instances.

96. It is also quite clear that there remain problems of communication across and within the IGO, NGO and academic sectors. Furthermore, the initiatives studied indicate missing links between the four disciplines of political science, law, economics and statistics. For example, political scientists have been engaged in what they consider to be the measurement of human rights, frequently without any real reference to the international law of human rights. Similarly, the human rights community has, from its inception, been dedicated to monitoring human rights violations and governmental performance. But it is only within the last decade that the tools of statistics have been introduced into human rights monitoring. There still remain considerable lacuna in human rights measurement, both in terms of the scope of rights being accurately measured and in terms of the number of actors in this field.

97. Similar problems affect the measurement of good governance. The term was originally utilised by economists, but was soon taken up by development donors and has now gained political significance, particularly in the field of overseas assistance. The measurement of good governance is often required in
order to provide a base line of information against which donor countries may
gauge the success or otherwise of their development assistance. But such
measurement becomes problematic when the perceived content of the concept
is so varied. As a result, reliance is often placed on existing initiatives to
measure particular aspects of democracy and/or human rights, which may be
taken out of context. A particular problem may arise when priorities set by
national governments for data collection do not coincide with those required by
donors.

98. There is an underlying desire to rank countries. Quite apart from the fact that
such ranking has political implications, methodologically there remain
difficulties in aggregating different dimensions into composite indices. In the
absence of measures for certain key categories, such rankings fail to produce a
holistic picture and tend to produce inaccurate (perhaps even meaningless)
results.

99. Of the difficulties encountered in measurement, by and large the
methodological problems are being resolved. What remains a significant
problem is the lack of source data and indicators in a number of critical areas
such as the enjoyment of economic, social and cultural rights (particularly
cultural rights), the rule of law, equality of access to justice, the measure of
judicial independence, participation of individuals in public affairs, and the
engagement of civil society. Various choices exist for filling these gaps:
existing indicators could be adapted; existing data could be further
disaggregated to reveal discriminatory practices affecting particular population
groups; or alternatively, new indicators could be developed.

3.2 RECOMMENDATIONS

3.2.1 General recommendations

100. The primary recommendation of this study is the need for conceptual clarity.
Progress cannot be made unless the concepts that are measured and their
component parts are clearly defined and identified. Equally important is
clarification of the precise purpose of measurement.

101. New indicators need to be developed to address the lacuna identified in the
study. In this endeavour: (1) there must be greater clarity of the concepts to be
measured and the purpose of measurement, (2) the key components of the
concepts of democracy, human rights and good governance should be
identified (in co-operation with national statistical offices) so that those data
need to be collected are clear, (3) specific sources of data and appropriate
indicators should be developed to measure each component, and unnecessary
aggregation into composite indices should be avoided, and (4) techniques for
generating data and indicators must have transparent rules for coding and
replication.

102. Human Rights measurement should be firmly rooted in the international law
of human rights and the juridical interpretation of state obligations and the
contents of the rights.
103. Given that development indicators offer only a partial picture of the enjoyment of economic, social and cultural rights, specific attention needs to be given to complementing this information source with rights specific indicators.

3.2.2 Specific recommendations to European Commission (EC)

104. At a bare minimum, the EC needs to: (1) invest in the development of human rights indicators and their utilisation should be a priority, (2) support the development of a core set of indicators to measure the contribution of the adoption of a human rights approach to development towards the achievement of poverty reduction and sustainable human development, (3) add additional layers to the database compiled within the present study to provide links to the initiatives identified in order to compile all the relevant material and indicators, and/or (4) merge the data sets identified in this study into one global data set of measures on democracy, human rights and good governance.
ANNEX I: INVENTORY OF INITIATIVES


Andreassen, B. and Swinehart, T. (Eds.) (various years), Human Rights in Developing Countries: Yearbook, Oslo: Scandinavian University Press


Gurr, T. (1990), Polity II: Political Structures and Regime Change, 1800-1986, Ann Arbor, MI: Inter-University Consortium for Political and Social Research.


Guzman M, (2000), The Investigation and Documentation of Events as a Methodology in Monitoring Human Rights Violations, HURIDOCs.


International IDEA (2003), *International IDEA Voter Turnout Website*, www.idea.int


Londregan, J. and Poole, K. (1990), Poverty, the Coup Trap, and the Seizure of the Executive Power, World Politics, 151-.


Political and Economic Risk Consultancy (2003), Asian Intelligence Reports, www.asiarisk.com


Purdue University (1998), *Purdue Political Terror Scale*, Department of Political Science, Purdue University.


Symonides, J (Ed.) (2000), Human Rights: Concept and Standards, Ashgate and UNESCO Publishing


United Nations University (UNU), *World Governance Assessment (WGA)*, www.unu.edu


Verkuilen, J. (2000), *Comparing Parallel Data Sets with Nonlinear Principal Components. The Case of Democracy Indices*, Unpublished manuscript, University of Illinois at Urbana-Champaign.


**NGO Initiatives**

The American Bar Association Central European and Eurasian Law Initiative (ABA/CEELI)
- Convention on the Elimination of All Forms of Discrimination Against Women (CEDAW) Assessment Tool and Reports
- International Covenant on Civil and Political Rights (ICCPR) Assessment Tool and Reports
- Judicial Reform Index and Reports
- War Crimes Documentation Project
  http://www.abanet.org/ceeli/home.html

American Association for the Advancement of Science, Human Rights Data Analysis Group, Various projects/publications, e.g. Killings and migration in Kosovo, March-June 1999
  http://shr.aaas.org/hrdag/

Amnesty International, Country Human Rights Reports
  http://www.amnesty.org/

  http://article19.org/

Carter Centre, Various projects/publications, e.g. Project on election monitoring in China
  http://www.cartercenter.org/

Centre for Economic and Social Rights, Various projects/publications, e.g. Unsanctioned Suffering – A Human Rights Assessment of UN Sanctions in Iraq (1996)
  http://cesr.org/


CIVICUS – The World Alliance for Citizen Participation, Civil Society Index Project
  http://www.civicus.org/

Coalition to Stop the Use of Child Soldiers, Child Soldiers Global Reports
  http://www.child-soldiers.org/
Committee to Protect Journalists, Attacks on the Press Worldwide Annual Survey
http://www.cpj.org/


End Child Pornography and Trafficking (ECPAT), Various projects/publications, e.g. Annual Reports since 1997 tracing progress on the Implementation of the Stockholm Agenda for Action Against the Commercial Sexual Exploitation of Children
http://www.ecpat.net/eng/index.asp

Forum for Early Warning and Early Response (FEWER), with Carleton University CIFP team, Country Indicators for Foreign Policy
http://www.fewer.org/

Freedom House (www.freedomhouse.org)
- Annual Freedom in the World Survey Nations in Transit
- Press Freedom Index
- Religious Freedom Index
- Human Rights Defenders project

**Global Barometer Surveys**
- New Europe Barometer (since 1991)
- Latinobarometer (since 1995)
- Afrobarometer (since 1999)
- East Asia Barometer (since 2001)

http://www.gilc.org/

Habitat International Coalition Housing and Land Rights Network, Methodology for Monitoring the Human Rights to Adequate Housing (The Tool Kit)
http://www.hic-mena.org/home.htm

Human Rights Watch, Country Human Rights Reports
http://www.hrw.org/

Human Rights Information and Documentation Systems (HURIDOCS), Various projects/publications, eg. Joint economic, social and cultural rights project with AAAS
http://www.huridocs.org/

International Campaign to Ban Landmines, Landmine Monitor Annual Report
http://www.icbl.org/lm/

International Confederation of Free Trade Unions, Annual Survey of Violations of Trade Union Rights
http://www.icftu-apro.org/

International Foundation for Electoral Systems (IFES), Opinion Research Project
http://www.ifes.org/

International Press Institute, Annual World Press Freedom Review
http://www.freemedia.at/index1.html

International Project Unit of the Association of Netherlands Municipalities, NOVIB
and The Habitat Platform, Toolkit and Database on Citizen Participation
http://www.toolkitparticipation.nl/

International Rehabilitation Council for Torture Victims (IRCT), Global Torture Victims Information System (GTIS)
http://www.irct.org/

http://iwraw.igc.org/iwraw/

National Democratic Institute for International Affairs (NDI), Various projects/publications, e.g. ‘Media Monitoring to Promote Democratic Elections, An NDI Handbook for Citizen Organizations’
http://www.ndi.org/

The Observer, Human Rights Index (1998-99)
http://www.guardian.co.uk/rightsindex/0,2759,201749,00.html

One World Trust, Global Accountability Index
http://www.oneworldtrust.org/

Philrights – The Philippines Human Rights Information Centre, Various projects/publications, e.g. ‘Monitoring Economic, Social and Cultural Rights – The Philippine Experience (Phase One)’
http://www.skyinet.net/~prights/

Physicians for Human Rights, Various projects/publications, e.g. Juvenile Justice project with the Youth Law Centre
http://www.phrusa.org/

Privacy International, Privacy and Human Rights Annual Survey
http://www.privacyinternational.org/

http://www.derechos.org.ve/ongs_ven/provea/
Reporters sans Frontières, Press Freedom Index
http://www.rsf.org/

http://www.ichrdd.ca/flash.html

Rights and Humanity, Various projects/publications, e.g. ‘Improving Women’s Health: An Introduction to the CEDAW Reporting Process’, prepared for WHO staff working with Member States preparing national reports to the CEDAW Committee.
http://www.rightsandhumanity.org

Save the Children, Annual State of the World’s Mothers Report, including Mothers Index
http://www.savethechildren.org/

Social Watch, Social Development Progress Reports

The Sphere Project, The Sphere Handbook & Key Indicators
http://www.sphereproject.org/

Transparency International
- Corruptions Perception Index
- Bribe Payers Index
- Kenyan Urban Bribery Index
- National Survey on Corruption and Governance (NSCG) / Encuesta nacional de Corrupcion y Buen Gobierno, Mexico
- Transparencia Brasil Vote Buying and Private Sector surveys
http://www.transparency.org/

http://www.wedo.org/

World Audit, World Democracy Audit
http://www.worldaudit.org/

World Movement for Democracy, Various project/publications, especially links to grassroots organisations working on democracy world-wide
http://www.wmd.org/
ANNEX II: REFERENCES CITED IN THE FINAL REPORT


Guzman M, (2000), The Investigation and Documentation of Events as a Methodology in Monitoring Human Rights Violations, HURIDOCS.


ANNEX III: CONCEPTUAL ISSUES

Democracy

1. A procedural definition of democracy, made most notably by Robert Dahl (1971) in *Polyarchy*, has included the two dimensions of *contestation* and *participation*. Contestation captures the uncertain peaceful competition necessary for democratic rule, a principle which presumes the legitimacy of some opposition, the right to challenge incumbents, protection of the twin freedoms of expression and association, the existence of free and fair elections, and a consolidated political party system. Participation captures the idea of popular sovereignty, which presumes the protection of the right to vote as well as the existence of universal suffrage. Figure 1 depicts this definition graphically.

2. Liberal definitions of democracy maintain concerns over contestation and participation, but add more explicit references to the protection of certain human rights. In a new model of liberal democratic performance, for example, Foweraker and Krznaric (2000) argue that liberal democracy contains an *institutional* dimension and a *rights* dimension. Like the procedural definition, this model shows that the institutional dimension captures the idea of popular sovereignty, and includes notions of accountability, constraint of leaders, representation of citizens, and universal participation. The rights dimension is upheld by the rule of law, and includes civil, political, property, and minority rights. The authors of this model are keen to stress that there may be important democratic ‘trade-offs’ between these dimensions and various democratic ‘profiles’ which may differentiate the democratic performance *qua* democracy (see also Foweraker and Landman 2002). This model of liberal democracy is depicted in Figure 2.

3. Social definitions of democracy maintain the institutional and rights dimensions found in liberal models of democracy but expand the types of rights that ought to be protected, including social and economic rights. Such an expanded form of democracy includes the provision of social and economic welfare and the progressive realisation of economic and social rights. This model is depicted in Figure 3. Conceptually, those advocating a pure liberal model of democracy argue that including such social dimensions mixes *intrinsic* and *extrinsic* features of democratic performance (see Foweraker and Krznaric 2000 and below). After all, it is possible for non-democratic regimes to provide social and economic welfare as well as the realisation of their associated rights. This has long been the argument of socialist regimes, particularly those of the former Communist countries of Eastern Europe.

4. The existence of different *definitions* of democracy necessarily means the existence of different *measures* of democracy. Moreover, the existence of such trade-offs in the liberal democratic model have important implications for the measurement of democracy since they preclude the ability to produce an
aggregate index of democracy, which has been the dominant strategy of most measurement efforts to date, and has motivated the most recent effort to produce a single democracy index (Munck and Verkuilen 2000; Munck 2002).

**Good Governance**

5. Like the long conceptual history of democracy extending from Aristotle’s classifications of ancient regimes until today (see Landman 2003), the notion of governance also has a long etymology and evolution, yet ‘good’ governance has only become fashionable within academic and policy circles relatively recently (Weiss 2000). The idea has developed from a neutral and generic reference to the overall set of relations within the public sphere to one that includes an expanding set of normative dimensions (i.e. the ‘good’ aspects of governance).

6. Court (2002: 5), in drawing on a collaborative project that measures good governance, defines governance as the ‘formation and stewardship of the formal and informal rules that regulate the public realm, the arena in which state as well as economic and societal actors interact to make decisions.’ Kaufmann et al (1999a: 1) define governance broadly ‘as the traditions and institutions by which authority in a country is exercised. This includes (1) the process by which governments are selected, monitored and replaced, (2) the capacity of the government to effectively formulate and implement sound policies, and (3) the respect of citizens and the state for the institutions that govern economic and social interactions among them’.

7. In its 1995 Guidelines the OECD uses the term ‘governance’ in accordance with a World Bank definition to denote ‘the use of political authority and exercise of control in society in relation to the management of its resources for social and economic development’ (OECD, 1995). The Commission on Global Governance (1995: 2) defines it as ‘the sum of the many ways individuals and institutions, public and private, manage their common affairs. It is the continuing process through which conflicting or diverse interests may be accommodated and co-operative action may be taken’ (quoted in Weiss 2002: 795-796).

8. Just as procedural, liberal, and social definitions of democracy are systematized concepts of democracy, good governance is a systematized concept of governance. Good governance emerged in the late 1980s and early 1990s primarily in the World Bank, which was concerned about the ways in which governance influenced economic performance (see World Bank 1992). The economic dimension of good governance has variously included public sector management, organizational accountability, the rule of law, transparency of decision-making, and access to information. This idea was taken on board by the OECD and EU and integrated into its requirements for development assistance. It was later expanded by the United Nations Development Programme to incorporate a political dimension that includes government legitimacy, government accountability, government competence, and the protection of human rights through the rule of law.
9. It is arguably beyond question that a government that deliberately violates human rights cannot be described as a ‘good government’. Neither can a government that consistently fails to introduce the steps necessary to ensure the realisation of human rights by all people within its jurisdiction (Häusermann, 1994). The World Bank Institute has identified six-dimensions of the quality of governance: voice and accountability, political stability, government effectiveness, regulatory quality, rule of law and control of corruption (World Bank Institute, 2002).

10. More recently, good governance has spawned the subsidiary concepts of humane governance (Falk 1995) and democratic governance (Fukuda-Parr and Ponzio 2002). But the remit of the map-making project is to focus on good governance, as it is this idea that has gained prominence in policy discourse.

11. It appears that a consensus is emerging on good governance with remaining doubts over whether it requires the existence of democratic institutions and over what kinds of human rights ought to be protected. Particularly in early good governance policies there was a tendency to restrict concern to civil and political rights, to the virtual exclusion of consideration of the relationship between good governance and ensuring respect for the economic, social and cultural rights recognised and protected by international law (Häusermann, 1994). As a result, the initiatives mapped pay scant regard to economic, social and cultural rights. Significantly, and perhaps due to the lack of definitional clarity or the relative infancy of the term, very few of the NGO initiatives identified in the preliminary research purport to be measuring good governance per se.

12. As in the conceptual debates surrounding democracy, measures of good governance will necessarily depend on its definition. Figure 4 depicts the most recent understanding of the concept of good governance.

**Human Rights**

13. In contrast to democracy and good governance, there is much greater academic and global consensus on the content of human rights. Such consensus is found in the international law of human rights found in the Charter of the United Nations, the Universal Declaration of Human Rights and subsequent conventions, six of which are regarded as the ‘core’ human rights treaties (see Figure 5) and each of which has a monitoring body to oversee the implementation of the treaty (Treaty Bodies).

14. The international protection of human rights has been complemented by the development of protection mechanisms at the regional levels of Europe, the Americas, and Africa.

15. The UN and regional human rights mechanisms are making some progress on clarifying the normative content of the human rights recognised in international law (General Comments of the Treaty Bodies). However, there remains a lack of clarity about the precise contents of some rights, for instance the right to social security (International Covenant on Economic, Social and
Cultural Rights, Article 9). The initiatives covered in this map-making study focused both on measuring violations of human rights (Amnesty International country reports; the American Association for the Advancement of Science, Human Rights Data Analyse Group) and on monitoring progress (Social Watch; End Child Pornography and Trafficking, annual reports tracing progress on the implementation of the Stockholm Agenda for Action Against the Commercial, Sexual Exploitation of Children since 1997).

16. Academically, there is consensus on the normative content of human rights, but less so for their philosophical foundations or their internal relativity. Normative and political philosophy have not yet found unshakeable foundations for the existence of human rights, a quest that has included appeals to human nature, the existence of God, and deontological sources (Mendus 1995; Ingram 1994). In response to this failure to find foundations for human rights, pragmatists simply focus on the need for a human rights framework that will protect the world from the worst forms of human behaviour (Mendus 1995; Falk 2000; Ignatieff 2001). While there continues to be much academic debate about the alleged primacy of civil and political rights over economic, social and cultural, at the inter-governmental level this debate was resolved at the World Conference on Human Rights, Vienna, 1993, which recognised all human rights as being ‘universal, indivisible and interdependent and interrelated’ (Vienna Declaration and Programme of Action, 1993, paragraph 5).

17. It is clear from this discussion that much conceptual confusion remains, and the key finding from the first part of the map-making project is that measurement is inextricably linked to conceptual definition such that the variety of measures available reflects a variety of definitions. There has been a tendency to elide or conflate these three ideas as well as to form a ‘chain of equivalence’ (Howarth 2000) among them that permeates international foreign policy documents. Thus, democracy, good governance, and human rights tend to be strung together as if they are equivalent concepts, when in fact each has important exclusive characteristics while at the same time having some shared characteristics.

18. The most popular definitions of democracy and good governance now include reference to the protection of certain categories of human rights, especially civil and political rights. But they also make reference to some economic and cultural rights, such as property rights and the rights of minorities. Definitions of human rights, drawn from the long history of their international legal evolution make reference to the right to participate in public affairs and democratic decision-making, and at times make explicit reference to a right of everyone to take part in the government of his country, directly or through freely chosen representatives (e.g. Article 21(1) of 1948 Universal Declaration of Human Rights). Moreover, many consider democracy to be ‘hollow’ without the protection of civil and political rights (Diamond 1999), while governance is considered to be ‘bad’ without the rule of law and the protection of human rights. Thus, democracy, good governance, and human rights have overlapping concerns but should not be seen as equivalent concepts.
Procedural Democracy

![Diagram of Procedural Democracy]

Liberal Democracy

![Diagram of Liberal Democracy]

Adapted from Schumpeter (1943), Dahl (1970), Hartlyn and Valenzuela (1994)

Adapted from Foweraker and Krznaric (2000)

Figure 1. Graphical representation of procedural democracy

Figure 2. Graphical representation of liberal democracy
Figure 3. Graphical representation of social democracy
Figure 4. Graphical representation of good governance

Adapted from OSCE (1991); World Bank (1992, 1994, 1997); Falk (1995); Weiss (2000); British Council (www.britcoun.org); Good Governance Sourcebook (DG Europe Aid); UNDP (2002); USAID (1998a, 1998b); DFID (2001).
Figure 5. The Core International Human Rights Instruments

Human Rights

United Nations Charter (1945)

Universal Declaration on Human Rights (1948)

International Covenant on Civil and Political Rights (1966)  
Ratified by 149 States*

International Covenant on Economic, Social and Cultural Rights (1966)  
Ratified by 146 States*

Ratified by 165 States*

Ratified by 170 States*

Convention Against Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment (1984)  
Ratified by 132 States*

Ratified by 191 States*

Sources:


http://www.unhchr.ch/html/intlinst.htm
ANNEX IV: DESCRIPTIVE GUIDE TO THE DATABASE OF INITIATIVES

Outline

This Access database provides descriptive and evaluative information of 178 initiatives on developing indicators for democracy (57), good governance (52) and human rights (69) carried out by academics, inter-governmental organisations (IGOs) and non-governmental organisations (NGOs). The Table Initiatives, which is the basic element of the database, stores the actual data of the 178 records. The Form Initiatives provides a more attractive way to enter, display and/or print the data of the table. This form should be taken as an example. It can be easily modified and improved to meet the requirements of the user. In the same fashion, queries and reports can be developed to exploit the basic information.

Description of the Fields

The database contains the following fields:

- **Author**: Author/s of the initiative (name of author or organisation)
- **Initiative**: Title of the initiative
- **Date**: Year of publication of the initiative
- **Origin**: Origin of the initiative (academic, IGO, NGO, private sector, etc)
- **Concept**: Main concept/s measured (democracy, good governance, human rights or other) by the initiative
- **Dimension**: Dimension/s of the concept/s measured
- **Component**: Component/s of the dimension/s measured
- **Measure**: Measure/s, indicator/s or data sets presented or developed
- **Method**: Means of measuring utilised (standards-based scales, events-based data, survey-based data, aggregate performance indicators, etc)
- **Level**: Measurement level (nominal, ordinal, interval, etc) and range
- **Time**: Time coverage of the data
- **N**: No. of countries covered
- **Scope**: Geographical coverage of the measure
- **Frequency**: Frequency of publication (One time, annual, updated periodically, etc)
- **Type**: Type of measure (measure of rights in principle, measure of actual practices, measure of outcomes, etc)
- **Purpose**: Purpose of measure (description, empirical testing, advocacy, etc)
- **Main initiative**: Whether the measure is a main initiative or not (yes/no)
- **Strengths**: Strengths of the main initiative
- **Weaknesses**: Weaknesses of the main initiative
- **Reference**: Bibliographical references
- **Comments**: Further comments and additional information
Recommendations

This flexible database can be improved by working in three main directions.

The number of observations can be increased by identifying, evaluating and coding new initiatives. At this stage, the database can be ‘normalised’ by closing some of the ‘open’ fields (i.e. those with multiple and infinite entry possibilities) and standardising some of the existing information. The construction of the database was grounded on two methodological choices. First, the fields in the assessment protocol and the database remained open categories in order to capture as much information as possible. Second, the strategy of coding the sources was to take them at face value. Both choices involved important methodological advantages, but also some costs in terms of standardisation. The normalisation of the database can improve its performance in terms of queries and reports as well as the reliability of the coding process.

The effort can be concentrated on the sub-sample of main initiatives (e.g. creating new tables describing this sub-sample of initiatives). The unit of analysis of this work was the individual efforts (initiatives) made by academics, IGOs and NGOs on measuring democracy, good governance and human rights. Some of those initiatives concerned the production of composite indexes, aggregate measures or data sets, which contains hundreds of individual indicators and involves the use of a great variety of means of measuring, measurement levels and ranges and aggregation and weighting rules. The database can be accommodated to capture the full richness of that information by creating new descriptive and evaluative fields or dividing some of the existing fields.

The initiatives can be linked to either web source material or to a global data set that collates all the initiatives into one file with all the indicators that have been identified.

Excel file

The information of the Access database is also presented in Excel format. The excel file contains the following information. The Sheets MeasuresGG, MeasuresD and MeasuresHR show information on the initiatives on measuring good governance, democracy and human rights respectively. The Sheet Main Initiatives outlines the information corresponding to the sub-sample of main initiatives, while the Sheet Initiatives Abridge lists the 178 initiatives by a selected number of fields. Finally, the Sheet Literature Review presents a list of 86 conceptual, empirical and methodological discussions on the measurement of democracy, human rights and good governance.