

4

Social science methods and human rights

The primary function of method in the social sciences is to establish a direct connection between the main research question, the empirical theories used to provide possible answers to the question, the propositions they make about the social phenomena under investigation (or the observable implications of a particular theory), and the collection of evidence that may or may not support these propositions (King, Keohane and Verba 1994; Landman 2000a, 2003). In this way, methods establish the ground rules of any enquiry, specify the types of knowledge that are possible given the theoretical expectations and assumptions of the researcher, and set the parameters for how evidence is collected and analysed. There is thus no one preferred method. Rather, method is a function of the research question that is posed, the theories used to help answer the question, and the epistemological orientation of the researcher. This volume has made it clear that at a bare minimum, social scientific analysis of human rights should be able to make meaningful analytical statements about observed (and in some instances, unobserved) social phenomena that either specifically involve human rights or are related in some important way. It is committed to enhancing the types of inferences that can be drawn about the variable protection and lived experience of human rights around the world.

With these initial statements in mind, this chapter provides an overview of the methodological options available to scholars and practitioners wishing to explain and understand human rights problems from a social scientific perspective. This overview is organized across three inter-related dimensions that frame the overall analysis of human rights problems. The first dimension represents an ‘epistemological continuum’ ranging across methodological approaches that vary according to the types of knowledge claims they make (universal vs. particular), the type of reasoning that connects their theory and evidence (inductive vs. deductive), the balance between evidence and inference, and the scope of coverage for their evidence (from sub-national units and single countries to global comparisons). The second dimension deepens the discussion concerning the degree to which cross-national comparisons and empirical generalizations frame the analysis of human rights problems; examining the general methodological trade-off between the ‘ladder of abstraction’ (Sartori 1970; Mair 1996) in the concepts that are used and the scope of units that form the basis of the analysis (Landman 2000a, 2002, 2003). The third dimension concerns the relative balance or mix between quantitative and qualitative evidence used in making inferences about human rights problems. These three methodological dimensions and their implications for human rights research are considered in turn.

An epistemological continuum

In an influential essay on the divisions in the discipline of political science, Gabriel Almond (with Genco 1977:489) argued that Karl Popper's (1972) metaphors of 'clouds' and 'clocks' to describe the continuum between indeterminacy and determinacy in physical systems equally apply to questions in political science research. On one end of the continuum, the cloud metaphor captures the irregular, disorderly, and unpredictable nature of things (e.g. swarms of gnats and flies), while on the other end of the continuum, the clock metaphor captures the regular, orderly, and predictable nature of things (e.g. pendulums, precision clocks, and motor cars). While Almond applies these metaphors and their challenges to research in political science, it is entirely possible to extend their application to the social sciences to argue that social science methods and research continue to vary a great deal across this continuum from those approaches that concentrate on the indeterminate and unpredictable cloudlike aspects of the social world to those that concentrate on its determinate and predictable clocklike aspects. The continuum has also been described as ranging from 'soft' interpretative and descriptive approaches to 'hard' rigorous and analytical approaches, but as we will see, the notion of rigour can apply to a wider range of methods and research than is typically understood (see King, Keohane and Verba 1994; Brady and Collier 2004). Table 4.1 lays out this continuum and specifies seven main categories of social science methods ranging from 'soft' hermeneutic approaches to 'hard' nomothetic approaches, which are further broken down by the type of reasoning, the balance between evidence and inference, the nature of the knowledge claims that each purports to make, and the scope of their empirical coverage. The table also includes examples from extant human rights research that fall within each of these different categories of analysis.

The hermeneutic and 'thick-description' (Geertz 1973) end of the continuum (Column I) contains those approaches that rely on descriptive and interpretative analysis of the social world, using a variety of qualitative methods such as participant observation, in-depth interviews and ethnographic methods, oral histories and narratives, archival documentation, and formal and informal discourses of individuals, as well as images, symbols, constructs, and architectures (see Devine 1995; Travers 2001; Howarth 2005:335–343). The goal of research that adopts these methods is to *understand* the nature and meaning of the social world that is constructed by the subject population under investigation. Less emphasis is placed on *explanation* of that social world, and there is rarely an attempt to make generalizations that extend too far beyond the evidence that has been examined. In this way, these approaches make knowledge claims that are limited to the particular social phenomena under investigation rather than make knowledge claims that have universal applicability. The units of analysis in such approaches tend to be individuals and groups that share common features and identities, sub-national units and geographical spaces, and single countries. Methodological criticisms of these approaches argue that pure descriptive studies have little social scientific value, since they are 'atheoretical' and 'interpretative' (Lijphart 1971:691), 'configurative-ideographic' (Eckstein 1975), and may simply provide 'evidence without inference' (Almond 1996:52). But scholars who adopt these approaches (and many who do not) argue that

such in-depth studies, while not seeking universal applicability, have tremendous inherent value, are full of inferences that add to our pool of knowledge about the social world, and have significant practical and policy implications (Hirschman 1970; Geertz 1973; Almond 1996; Landman 2000a, 2003; Flyvberg 2001).

Table 4.1 The epistemological/methodological continuum of the social sciences

<i>Range</i>	<i>I</i>	<i>II</i>	<i>III</i>	<i>IV</i>	<i>V</i>	<i>VI</i>	<i>VII</i>
Type of approach	Hermeneutic/thick description	Discourse analysis	Theory-driven empirical	Theory-driven empirical	Theory-driven empirical	Theory-driven empirical	Nomothetic
Reasoning	Inductive	Inductive and analytical	Inductive and analytical	Inductive and analytical	Inductive and analytical	Deductive and analytical	Deductive
Evidence vs. inference	'Evidence without inference'	Meaning and understanding from language and action	Qualitative evidence and inference	Quantitative/qualitative evidence and inference	Quantitative evidence and inference	Inference with confirmatory evidence	'Inference without evidence'
Nature of knowledge claim	Particular Context-specific	Particular Context-specific	Universal with room for exceptions	Universal with room for exceptions	Universal with room for exceptions	Universal	Universal
Scope of coverage	Single countries Sub-national analysis	Single countries Limited comparison	Comparative and single-case analysis	Comparative	Global comparative	Small-N comparative	Theoretical constructs only
Examples in human rights research	Goldstein (2004) <i>The Spectacular City</i>	Norval (1996) <i>Deconstructing Apartheid Discourse</i>	Risse, Ropp and Sikkink (1999) <i>The Power of Human Rights</i>	Foweraker and Landman (1997) <i>Citizenship Rights and Social Movements</i>	Poe and Carey (2004) <i>Understanding Agents of Human Rights Atrocity Violations</i>	Mitchell (2004)	Wantchekon and Healy (1999) 'The "Game" of Torture'
	Huggins <i>et al.</i> (2002) <i>Violence Workers</i>	Roniger and Sznajder (1999) <i>The Legacy of Human Rights Violations in the Southern Cone</i>	Hawkins (2002) <i>International Human Rights and Authoritarian Rule in Chile and Nicaragua</i>	Gómez (2003) <i>Human Rights in Cuba, El Salvador, and Nicaragua</i>	Landman (2005) <i>Protecting Human Rights</i>		

Two examples of empirical research relevant to human rights that fit into this category of research include Huggins *et al.*'s (2002) study of police torturers and murderers in Brazil and Goldstein's (2004) study of vigilante violence in Villa Pagador in Cochabamba, Bolivia. Both studies adopted ethnographic methods, in-depth interviews,

and interpretative analysis to understand why and how human rights violations have occurred in particular economic, social, and political contexts. While their analyses are grounded in the particular contexts of Brazil and Bolivia, the accounts are located in larger questions concerning the motivations for violent action, the conditions and settings in which they take place, and the larger understandings and cultural meanings of the violence that stretch beyond the confines of the subject population under investigation. In the Brazilian study, the analysis examines a sample of twenty-three former and serving police officers whose careers spanned the prolonged period of authoritarian rule (1964–1985) and who became either direct perpetrators or ‘atrocious facilitators’ in carrying out systematic torture and extra-judicial killings. Its methods include network analysis (see Scott 2000) of the sample of police officers, deconstruction of the categories of ‘victims’ and ‘perpetrators’, and storytelling from the violence workers themselves (Huggins, Haritos-Fatouros and Zimbardo 2002:17–28). This combination of methods is used to weave together five significant themes that influenced the memories, perceptions, and consequences of their actions: secrecy, occupational insularity, organizational fragmentation, personal isolation and the changing nature of the Brazilian state (Huggins *et al.* 2002:2).

In similar fashion, Goldstein uses evidence collected through participant-observation to examine the ways in which public ‘spectacles’ of seasonal festivals and vigilante violence represent forms of local expression of citizenship and modes of action that recapture control of individual and collective security in the face of enhanced socio-economic marginality. In making justice ‘by one’s own hands’ the attempted and actual lynchings of thieves by the residents of Villa Pagador ‘form a spectacular cultural performance, a means for people ordinarily excluded from the political, economic, and social mainstreams of Bolivian society to force themselves violently into the public eye’ (Goldstein 2004:3). By comparing these violent activities with other public spectacles such as the annual Fiesta de San Miguel, Goldstein (2004:2–4) argues that not only do they represent ‘vivid, visual displays of collective identity for barrio residents’ but also ‘a claim to citizenship and a demand for citizens’ rights in a context of political, legal, and socio-economic exclusion’. These ‘dramas of citizenship’ (Holston and Appadurai 1999:14) are part of a larger crisis in contemporary Bolivia, where poverty and social exclusion have resulted from the prolonged implementation of neo-liberal economic programmes and the violation of human rights has increased under prosecuting the ‘war on drugs’.

The next category in the continuum (Column II) includes those approaches that adopt discourse-analytic techniques to problem areas in the social sciences. Such analysis is much akin to hermeneutic analysis in that it seeks to elucidate ‘problematized objects of study by seeking their description, understanding, and interpretation’ (Howarth 2005:319). The goal of such analysis is to produce new interpretations about specific objects of investigation through either uncovering phenomena previously obscured and undetected by dominant social scientific theories and approaches or by ‘problematizing existing accounts and articulating alternative interpretations’ (Howarth 2005:320). Such analysis relies on inductive and analytical reasoning that examines the social and political ‘logics’ at work in the construction of meanings, understandings, and articulatory practices. Such meanings and understandings are obtained through an analysis of language and action, which are seen as mutually constitutive phenomena. Like its

hermeneutic counterpart, discourse analysis eschews making universal generalizations and tends to analyse small sub-national units and single countries, although it has begun to explore ways in which to carry out comparative analysis that is 'problem-driven', but does not sacrifice attention to historical context and concrete specificities of the cases under comparison (Howarth 2005:332).

Two examples that fit into this category of method and analysis include Norval's (1996) study of the construction of apartheid in South Africa and Roniger and Sznajder's (1999) comparative study of the legacies of human rights violations in Argentina, Chile, and Uruguay. While Norval's (1996) study is more wedded to post-Marxist and post-structural social theory, both studies analyse the ways in which discursive practices and understandings constructed coercive systems with grave consequences for human rights that have lasted well beyond the actual periods of authoritarianism. In line with the international legal definition of apartheid as a 'crime against humanity' (see Buergethal 1995:67–68; Ghandhi 2002:91–95), Norval's (1996:2) task is to explain how and why the establishment and maintenance of such a system was made possible, who became 'interpellated' by its hegemonic discourse, how that hegemony was called into question in the 1980s, and what the legacies of apartheid are for the new democratic South Africa. Through an analysis of discourses located throughout the period extending from the 1930s and 1940s to the 1980s and 1990s, her account shows that apartheid was not an automatic outgrowth of the 'dislocatory' upheavals in the 1930s and 1940s (e.g. the capitalization of agriculture, rapid urbanization, and World War II), nor was it created in the service of particular economic interests. Rather, it competed with many other discourses of the day, instituted a new form of social division and associated modes of identification, and drew two sets of 'political frontiers' upon which it was situated (Norval 1996:5–11). Such hegemony, however, became unravelled in the 1980s, when apartheid failed to maintain its system of social division, and had to compete with new discourses that challenged it.

In similar fashion, Roniger and Sznajder (1999:1–2) argue that post-authoritarian processes of democratization, reconciliation, and construction of the 'truth' are not inevitable outcomes to broad structural changes, but contingent and contested discursive processes where the re-emergent civilian rulers and representatives from different sectors in society are 'caught between contrasting and sometimes polar versions of the past, between normative expectations and the constraints of contingency, and between their will to consolidate democratic rule and the impossibility to do so without grappling with the past legacy of human rights violations'. Their analysis of these processes in Argentina, Chile and Uruguay uses analytical concepts from theories of democratization, globalization and the discourse of human rights, collective memory and identity. It adopts a 'multi-level' approach through a systematic comparison of the three cases by combining a political-institutional approach, a sociological and comparative approach, and a cultural approach that focuses its attention on the discourses of key political and cultural actors (Roniger and Sznajder 1999:2–5). They are less sceptical of comparative analysis than most discourse analysts (see Howarth 2005:332–335) and adopt a version of the 'most similar systems design' (see discussion below, and Skocpol and Somers 1980; Faure 1994; Landman 2000a, 2003), which compares different outcomes across a set of similar countries. In this way, Argentina, Chile and Uruguay share a relatively common background and confronted similar legacies of human rights violations, but differed

greatly in the institutional paths they took as well as the ‘actual interpretations’ of human rights issues among different social sectors (Roniger and Sznajder 1999:2). The key discursive contribution of the study is its analysis of how cultural elites use symbols and models of social order to construct identities and draw boundaries of inclusion and exclusion during the periods of redemocratization across the three cases (Roniger and Sznajder 1999:6).

The next three categories in the continuum (Columns III, IV, and V) share a general orientation to providing theory-driven empirical analysis that is inductive, comparative, and seeks to make broad generalizations that still leave room for some exceptions. While they differ in the degree to which they use qualitative and quantitative analysis and in the number of countries that feature in their comparisons, they are all self-conscious about research methods and the strength of the inferences that are drawn from their analyses. For example, Risse, Ropp and Sikkink (1999) and Hawkins (2002) engage in qualitative comparative and single-country analysis to examine how the transmission of human rights norms from the international level to the domestic level shapes state behaviour towards human rights over time. Gómez (2003:83–96) uses quantitative data to set basic timelines of human rights abuse in the cases of Cuba, Nicaragua, and El Salvador and then uses comparative historical analysis across these ‘most similar’ cases to understand why they have had quite different patterns in human rights abuse. Foweraker and Landman (1997) use a mixture of time-series quantitative analysis and qualitative histories across the four cases of Brazil, Chile, Mexico, and Spain to examine the variable relationship between social mobilization and the protection of individual civil and political rights of citizenship. In contrast, the studies on the human rights violations in Poe and Carey (2004) and the analysis of the growth and effectiveness of the international human rights regime in Landman (2005b) use global quantitative analysis across large samples of countries ($150 \leq N \leq 194$) and over long periods of time ($16 \leq T \leq 30$).

Finally, the last two categories in the continuum (Columns VI and VII) share the same orientation toward providing universal knowledge claims and explanations for human rights violations based on deductive reasoning, while differing in their reliance on evidence. Most rationalist forms of analysis engage in a process of theorizing (see Chapter 3 this volume), where assumptions about human nature are combined with a series of ‘stylized facts’ and then used to derive a set of propositions about social phenomena that can be observed. Approaches in the penultimate category (Column VI) involve deriving a set of propositions deductively from starting assumptions and then testing them with limited empirical evidence, which typically consists of confirmatory case studies. In contrast, approaches in the final category (Column VII) deduce their propositions in exactly the same manner, but do not subject them to empirical testing. Rather, they present their analysis in abstract form only and have thus been criticized for providing ‘inference without evidence’ (Almond 1996:52). A classic example from political science in this category is Anthony Downs’s (1957) *An Economic Theory of Democracy*, which treats democracy as a market, politicians as firms with something to sell (i.e. their policies) and voters as consumers who ‘spend’ their votes on the candidates that best ‘supply’ goods that maximize their preferences.

In the field of human rights, there are two examples from the extant literature that illustrate these methods and their different uses of evidence: Mitchell’s (2004) study of human rights atrocities during civil wars and Wantchekon and Healy’s (1999) analysis of

torture using game theory (see Chapter 3 this volume). Mitchell (2004) develops a quasi-rationalist account using analytical concepts of ‘principals’ and ‘agents’ drawn from the fields of organizational behaviour, management studies, and economics. Principals task agents with certain functions that they carry out with limited degrees of accountability to the principal (e.g. presidents and vice-presidents, generals and foot soldiers, leaders and security personnel). Using the great historical figures of Machiavelli, Dostoevsky’s Grand Inquisitor, and Count Tilly, Mitchell (2004:29–56) argues that principals may be motivated by the desire for political power and/or the pursuit of an ideal, which helps explain their relative propensity to tolerate human rights violations as a means to a greater end, while agents may be motivated further out of pure self-interested greed and sadism.

The analysis of different combinations of differently motivated actors serves to ‘restore the focus to choice and to responsibility in the use of violence’ (Mitchell 2004:6). This simple yet powerful model is then tested across three civil wars—the English Civil War (1642–1651), the Russian Civil War (1917–1921) and the Arab-Israeli conflict (over the period 1948–2002)—in an effort to demonstrate its logic and extend its inferences to other cases. The analysis of torture using game theory outlined in Chapter 3 also restores focus to the role for individual choice, while the stylized specification of ‘the state’ and ‘the torturer’ as actors is in line with Mitchell’s (2004) use of principal and agent. In contrast, however, Wantchekon and Healy (1999) make little reference to the empirical world and do not test their model using evidence. Rather, they derive an equilibrium solution and draw their implications based on the formal analysis of the strategic interaction between the torturer and the victim. Thus, in both examples deductive logic provides a parsimonious framework for analysing human rights problems, which in the former study are examined through comparative cases, while in the latter example they remain inferences defended through formal analysis only.

This section considered a broad range of methodological approaches to studying human rights. It delineated the different methods according to a general epistemological continuum ranging from ‘cloudlike’ to ‘clocklike’ extremes, the type of reasoning (inductive versus deductive), the balance between pure evidence and pure inference, the nature of the knowledge claims, and the scope of geographical coverage in the collection and analysis of evidence on human rights problems. It is clear that extant and new studies located along this continuum ask different research questions or examine similar sets of questions differently, such that the universe of human rights issues and topics for research can be confronted by a plurality of methodological approaches. But to recognize such a plurality is not to adopt an ‘anything goes’ position with respect to social science research methods (see Feyerabend 1993). Rather, it is to adopt the position that a number of significant, valid, and recognized research traditions characterize the social sciences and that each of these traditions asks particular questions of its research material and applies different methods to its objects of inquiry. Using the discussion of this continuum as a backdrop, the next section examines in greater detail the various trade-offs associated with different comparative methods, whose aim is to provide generalizations about human rights problems based on the comparative analysis of countries.

Cross-national generalizations and comparative methods

At a fundamental level, comparative methods provide ways in which to compare similarities and differences across countries to arrive at a series of generalizations about particular human rights problems. There are three general comparative methods available to social scientists of human rights: global comparisons, few-country comparisons, and single-case studies. The trade-offs associated with these methods involve the degree to which each can make broad-ranging empirical generalizations at different levels of theoretical and conceptual abstraction (Mair 1996; Landman 2000a, 2002, 2003, 2005b). Global comparisons tend to make broad-ranging empirical generalizations using concepts and constructs at a fairly high level of abstraction. Few-country comparisons tend to limit their generalizations and lower the level of abstraction in analysing human rights problems across a selection of countries. Single-case analysis tends to limit further its empirical generalizations and concentrates on the contextual particularities of the single case under investigation, but can be constructed in such a way as to contribute to larger theoretical and empirical problems. These three comparative methods and their associated strengths and weaknesses are considered in turn.

Global comparative analysis

Global comparative analysis typically involves the use of large and complex data sets comprising variables that have been operationalized quantitatively (see below) and have been specified in such a way that they can be measured over time and across space (see Chapter 5). Rarely do such studies carry out their analysis using qualitative methods, which focus more in-depth on particular features within countries and is therefore not possible with sample sizes that often exceed 100 countries (see Chapter 6). With such a large number of observations (typical time and space combinations exceed 4,000 such observations), global comparisons make empirical generalizations about relationships between and among variables that have associated degrees of statistical significance. For this reason, such studies are known as ‘variable-oriented’ since their focus is on ‘general dimensions of macro-social variation’ and the relationship between variables at a global level of analysis (Ragin 1994:300). The main strengths of this kind of analysis include statistical control to rule out rival explanations, extensive coverage of cases, the ability to make strong inferences, and the identification of ‘deviant’ cases or ‘outliers’. For example, one typical finding from the extant global comparative literature on human rights suggests that ‘personal integrity rights’ violations are lower in countries that have high levels of economic wealth and democracy (see Chapter 6). The fact that there are numerous wealthy countries and/or democracies that continue to violate human rights does not undermine this general finding. Rather, such countries become significant outliers to the general relationship that are in need of further analysis to explain why their practices appear anomalous. For example, Saudi Arabia is a classic outlier that is wealthy and has significant problems with human rights violations, while Brazil and India are seen as consolidated democracies with exceptionally high levels of torture (Landman 2005b).

Global analysis also has a number of weaknesses, including data availability, validity and reliability of rights and other measures, and its limited application to human rights problems. First, until very recently, there had been a dearth of cross-national data on human rights practices. As the next chapter shows, there are still only five major sources of human rights measures available for global comparative analysis, all of which are limited ordinal ‘standards-based’ scales of human rights practices (see also Landman 2004). The ‘political terror scale’ (Mitchell, Stohl, Carleton and Lopez 1986; Poe and Tate 1994; Gibney and Dalton 1996; Gibney and Stohl 1998), the Freedom House civil and political liberties scales (Gastil 1978, 1980, 1988, 1990; <<http://www.freedomhouse.org/>>), and the torture scale (Hathaway 2002) measure a narrowly defined set of civil and political rights, while the Cingranelli and Richards human rights data set (<<http://www.humanrightsdata.com/>>) includes measures of civil, political and some economic rights. Second, there are serious questions remaining about the validity and reliability of these rights measures, which code qualitative information typically found in Amnesty International and/or US State Department human rights country reports into quantitative scales. Third, global comparative analysis cannot address a whole range of important research questions in the human rights field, since many such topics are not susceptible to quantitative methods. Even if they are, global quantitative analysis provides generalizations that need greater specification and in-depth research that can only be carried out on smaller samples of countries (Landman 2005a).

Few-country comparisons

It is precisely because of the limitations and weaknesses of global comparative analysis that many human rights scholars carry out their analyses on a smaller selection of countries. Variously called the ‘comparative method’, the ‘comparable cases strategy’ (Lijphart 1975), or ‘focused comparison’ (Hague, Harrop and Breslin 1992), comparing few countries achieves control through the careful selection of cases that are analysed using a middle level of conceptual abstraction. Studies using this method are more *intensive* and *less extensive* since they encompass more of the nuances specific to each case. The outcomes that feature in this type of comparison are often seen to be ‘configurative’, i.e. the product of multiple causal factors acting together. In contrast to global comparative analysis, this type of comparison is referred to as ‘case-oriented’ (Ragin 1994), since the case is often the unit of analysis and the focus tends to be on the similarities and differences among cases rather than the analytical relationships between variables. Such comparisons tend to make generalizations that are less broad using concepts and constructs that have been analysed in greater depth across the countries that have been selected for analysis.

In order to make these generalizations, the comparison of the similarities and differences across a small number of countries is meant to uncover the empirical relationship between the presence of key explanatory factors (X_1, X_2, \dots, X_n) and the presence of an observed outcome (Y). The isolation of these explanatory factors and the determination of their relationship to the observed outcome can be achieved through adopting two distinct types of research design: ‘most similar systems design’ and ‘most different systems design’ (Przeworski and Teune 1970; Skocpol and Somers 1980; Faure 1994; Landman 2000a, 2002, 2003). Drawing on J.S. Mill’s (1843) ‘method of

difference', most similar systems design (MSSD) compares different outcomes across similar countries. Comparing countries that share a host of common features allows for the isolation of those factors that may account for an outcome. Typically, regional and area studies analysis compares countries that share similar history, language, religion, politics, and culture and then isolates the remaining factors that vary across the cases to see if that variation is related to the variation in the outcome that is to be explained. In this way, the common features are 'controlled', while the analysis focuses on the relationship between the explanatory factors and the outcome. As noted above, Roniger and Sznajder (1999) compare the different institutional paths and interpretations of human rights issues across the most similar cases of Argentina, Chile, and Uruguay, while Gómez (2003) adopted the same research design to analyse the differences in human rights abuse across the most similar cases of Cuba, El Salvador, and Nicaragua.

In contrast to MSSD, most different systems design (MDSD) compares similar outcomes across different countries. Drawing on Mill's (1843) logic of agreement, MDSD compares countries that share very few features and then focuses on those factors common across the countries that may account for an outcome. In this way, selecting countries with the same outcome and matching that outcome to the presence of key explanatory factors allows the researcher to establish their empirical relationship. Comparative studies that focus on large historical events such as revolutions, military coups, transitions to democracy, or 'economic miracles' in newly industrialized countries (Geddes 1990:134–141) adopt this basic research design, where these types of outcomes are matched to the presence of key explanatory factor(s). In *Peasant Wars of the Twentieth century*, Wolf (1969) compares instances of revolutionary movements that had significant peasant participation in Mexico, Russia, China, North Vietnam, Algeria, and Cuba. Though these countries share few features that are the same, Wolf argues that the penetration of capitalist agriculture is the key explanatory factor (X) common to each that accounts for the appearance of the revolutionary movements and their broad base of peasant support (Y). In the human rights field, Hayner (1994, 2002) compares the outcomes and impact of similar instances of truth commissions across countries in Latin America and Africa. Her comparisons reveal that ethnic, religious, and group conflict in Africa explains why reconciliation is less likely than in Latin America, where conflict was born of an ideological struggle between forces on the political left and right. Thus the nature of conflict (X) is a key explanatory factor that accounts for the differences in impact of truth commissions (Y).

Thus, both MSSD and MDSD seek to identify a relationship between explanatory factors and outcomes by comparing different outcomes across similar countries or similar outcomes across different countries. Of the two research designs, MSSD is slightly more robust, since it allows for the presence of different outcomes across the countries under investigation. In social scientific terms, this means that the dependent variable (i.e. the outcome) is allowed to vary. In contrast, MDSD does not allow for the presence of different outcomes, and thus has no variance in the dependent variable. Thus, MDSD at best establishes a concomitance of explanatory factors and outcomes since it does not allow for 'negative' instances of the outcomes being examined (see Mahoney and Goertz 2004). In the Wolf (1969) example above, the analysis would have been strengthened through a test of whether there was (a) the presence of significant peasant support for revolution in the absence of the capitalization of agriculture or (b) the absence of

significant peasant support in the presence of the capitalization of agriculture. As it stands, his analysis does not include such alternative cases.

Despite the main strengths of few-country comparisons in allowing for greater attention to the detail and specificity of country cases, such analysis suffers from two major methodological weaknesses. First, such studies may identify a large number of explanatory variables whose full variation far exceeds the number of countries under investigation. This problem is commonly labelled ‘too many variables, not enough countries’ (Dogan and Pelassy 1990; Collier 1991; Hague, Harrop and Breslin 1992), or ‘too many inferences and not enough observations’ (King, Keohane and Verba 1994). For example, a study that specifies three explanatory variables each with several categories (e.g. low, medium, and high) and outcome variable with as few as two categories (e.g. improvement in rights protection or not), and then analyses these variables across only three countries will never really be able to establish a relationship between the explanatory variables and the outcome. Solutions to this particular problem include raising the number of observations (include time, sub-national units, or more countries); resort to MSSD, which controls for the common features; or reduce the number of explanatory variables through adopting MDSD or through better theoretical specification (Landman 2003:40–41).

Second, the *intentional* selection of cases rather than a *random* selection can seriously undermine the types of inferences that can be drawn. This problem is known as ‘selection bias’, and occurs in comparative politics through the non-random choice of countries for comparison, or the deliberate selection by the comparativist (Collier 1995:462). Though selection of countries lies at the heart of comparison, selection without reflection may lead to serious problems of inference. The most blatant form of selection occurs when a study includes only those cases that support the theory. More subtle forms of selection bias, however, occur when the choice of countries relies on values of the dependent variable (Geddes 1990; King, Keohane and Verba 1994; Mahoney and Goertz 2004) and for qualitative studies, the use of certain historical sources that (un)wittingly support the theoretical perspective of the researcher (Lustick 1996). As outlined above, MDSD suffers from selection bias relating to values of the dependent variable, where only those countries with the outcome of interest (e.g. democratic transition, military coup, revolution) have been selected. Relatedly, it is possible to construct a few-country comparison that contains an indeterminate research design by comparing different outcomes across different countries. For example, in *The Power of Human Rights*, Risse, Ropp and Sikkink (1999) present five ‘paired comparisons’ (Kenya-Uganda, Tunisia-Morocco, Indonesia-Philippines, Chile-Guatemala, and Poland-Czechoslovakia) and one single-country analysis of South Africa to examine the transmission of international human rights norms. But the analysis reveals different outcomes across these different cases, thereby limiting the types of inferences that are drawn (Landman 2003, 2005a).

Single-country studies

The field of human rights research is full of single-country studies. By definition, they focus on countries with particularly problematic human rights records and include official reports from international governmental and non-governmental organizations, domestic commissions and NGOs, journalistic and descriptive accounts, and research monographs.

The *Nunca Más* (CONADEP 1984) report from Argentina and the *Nunca Mais* (Dassin 1986) report from Brazil are classic examples of such descriptive accounts of human rights abuse under conditions of authoritarianism, and as discussed above, truth commissions often publish their findings for the general public, such as the South African Truth and Reconciliation Commission and the *Comisión de Verdad y Reconciliación* in Peru (Truth and Reconciliation Commission, CVR) (see Chapter 7). On balance, however, these descriptive accounts are not grounded in any one discipline, and they rarely make larger inferences from the intensive examination of the individual case. The descriptive accounts do, however, serve as the foundation for research monographs, which are grounded in one or more disciplines and tend to locate the country study in a broader set of theoretical and empirical questions relevant to the study of human rights.

As noted above, there have been significant criticisms made about the usefulness of single-country studies for social scientific analysis. But this volume takes the view that single-country studies can make significant and valuable contributions to the study of human rights, and that when properly carried out, they can provide a number of important functions in developing our understanding of human rights problems. First, they provide important contextual description upon which other studies build their analyses. Second, they develop new classifications of events and outcomes not yet observed in other parts of the world. Third, they can be used to generate hypotheses that can be tested in other countries. Fourth, they can be used to confirm or infirm existing theories by providing ‘crucial’ tests, and by extension provide explanations for anomalous outcomes or ‘deviant’ cases identified through global comparative analysis (Eckstein 1975).

Of these important functions, their role in new classifications, the generation of hypotheses and their use as ‘crucial cases’ are the most germane to the discussion here. There are several examples where the development of new classifications has advanced scholarship in describing, understanding, and explaining patterns of human rights abuse. In describing the Franco regime in Spain, Juan Linz (1964) identified a new form of authoritarianism that was different from personalistic dictatorships and totalitarian states. The Franco regime institutionalized representation of the military, the Catholic Church, and the Falange, as well as the Franco loyalists, monarchists, and technocrats. Unlike totalitarian states, the regime relied on passive mass acceptance rather than popular support (Linz 1964). Guillermo O’Donnell (1973) built on Linz’s (1964) work in Spain and established the concept of the ‘bureaucratic-authoritarian state’ in his examination of Argentine politics and the regime’s prosecution of the ‘dirty war’, a concept of state organization that would be applied not only to other authoritarian regimes in Latin America (see Collier 1979) but also to those in Southeast Asia (see Geddes 1990). In another example, patron-client relations and their permeation of state organization identified in Latin America have been developed into models of neo-patrimonialism and predatory states in Africa (Clapham 1982; Bratton and van der Walle 1997; Haynes 2002). In other examples, Kaldor (1999), Gilbert (2003), and Münkler (2005) have specified new forms of warfare that move beyond more traditional understandings of conflict and that have grave consequences for human rights, and Payne (2000) has developed the concept of ‘uncivil’ movements that can ‘travel’ for subsequent comparative studies.

Single-country studies are also useful for generating hypotheses for theories that have yet to be specified fully. As ‘plausibility probes’ (Eckstein 1975:108), single-country

studies explicitly (or implicitly) suggest that the generated hypothesis ought to be tested in a larger selection of countries (Lijphart 1971:692). For example, in their analysis of the effectiveness of international human rights pressure on the Argentine military regime, Weissbrodt and Bartolomei (1991:1034) conclude by arguing 'the lessons of this case study must be tested in cases involving other countries and time periods to determine whether more general lessons can be drawn from this single case'. In similar fashion, in his study of the relationship between international human rights pressure and the transformation of the Pinochet regime, Hawkins (2002) tests the hypotheses generated in the Chilean case in the additional cases of Cuba and South Africa. His analysis of Chile shows that certain 'rule-oriented' factions within the Chilean military became influenced by outside human rights pressure, which ultimately led to gradual concessions by the regime and the transition to democracy. The further testing of his central hypothesis shows that a similar process took place in South Africa but not in Cuba, since there are not significant fissures in the ruling elite that would be susceptible to the influence of international human rights pressure.

Finally, single-country studies are useful if they act as 'crucial' cases drawn from theoretical expectations and propositions about the world. Such crucial case studies can confirm or infirm existing theories and are therefore conducted within the confines of extant generalizations (Lijphart 1971:692). There are two types of crucial case studies: 'most likely' and 'least likely' (Eckstein 1975:118). Least likely case studies select a country where theory suggests an outcome is not likely to occur. If the outcome is observed, then the theory is infirmed, since it suggested such an outcome should not be obtained in that particular country. For example, in their analysis of democratic transition in South Africa, Howarth and Norval (1998) argue that the South African case is the best example of a least likely case study since the longevity and strength of the apartheid regime suggested that a democratic transition was highly unlikely. The fact that there was such a transition invites deeper analysis of the case itself and greater reflection on theories of democratization. Other 'least likely' candidates in this area of research include North Korea and Burma/Myanmar, which over the next coming years may undergo similar unexpected processes of democratic transition.

In contrast, most likely case studies apply a reverse logic to least likely studies by selecting countries where theory suggests the outcome is definitely meant to occur. If the outcome is not observed, then the theory is infirmed. For example, Brazil and the United States are seen to be most likely case studies that have confounded particular social theories that link socio-economic change to political outcomes. For the Brazilian case, many varieties of social theories suggest that Brazil has had all the objective economic conditions necessary for a social revolution and yet no attempt to organize a mass-based revolutionary force has ever been made. In similar fashion, despite its rapid pace of industrialization, expansion of its labour force, and constitutional protection for the rights to assembly and association, the United States has never had a strong socialist party. The task of the analyst is thus to explain these so-called 'non-events' in these particular cases through identifying those factors that have inhibited the development of a fully fledged revolutionary movement in Brazil or a strong socialist party in the United States. In this way, Brazil and the United States represent 'deviant' cases since they fail to fall in line with theoretical expectations. Thus, most likely case analysis provides the means to explain the presence of such deviant cases. Additional candidates for most likely case

analysis include Cuba and China, which have failed to undergo processes of democratic transition despite the Velvet revolutions' and the collapse of Communism in the former Soviet Union.

Single-country studies thus serve larger comparative purposes if they lead to new classifications of social phenomena, generate new hypotheses about important empirical relationships, and provide critical tests of extant theories. Human rights abuses take place across a huge range of different social, economic, and political contexts, and single-country studies provide the richness of contextual description and the analysis of new institutional, cultural, and behavioural phenomena. Such studies should not be seen as 'merely precursory moments' (Howarth 2005:332) in the larger quest for social scientific explanation, but as also having value in and of themselves. As we have seen, however, in making these new classifications and analyses, single-country studies can generate important hypotheses to be tested in other countries and contexts. Moreover, crucial case studies, whether 'most likely' or 'least likely' do not definitively prove or disprove a theory in line with Popper's (1959) notion of scientific falsification, but they do help *confirm* or *infirm* the applicability of social theories to all cases (Eckstein 1975; Landman 1999). They are thus particularly useful in testing the robustness of theories and research programmes in the social sciences that make universal knowledge claims, such as those outlined in the first section of this chapter.

Quantitative and qualitative evidence

The final methodological dimension in need of explicit attention is the difference between quantitative and qualitative evidence and its use in studying human rights problems. Quantitative methods seek to show differences in number between certain objects of analysis and qualitative methods seek to show differences in kind. Quantitative analysis answers the simple question, 'How many of them are there?' (Miller 1995:154), where the 'them' represents any object of comparison that can either be counted or assigned a numerical value. Quantitative data can be official aggregate data published by governments on growth rates, revenues and expenditures, levels of agricultural and industrial production, crime rates and prison populations, or the number of hectares of land devoted to agrarian reform. Quantitative data can also be individual, such as that found in the numerous market research surveys and public opinion polls. In the field of human rights, it is possible to count human rights violations, convert subjective accounts of human rights practices into standardized scales, or to collect survey data on human rights practices from random samples of the population. Such measures of human rights (see Chapter 5 this volume) can then be used for statistical analysis that describes and explains the nature, extent, pattern, and causes of human rights violations.

Quantitative methods are based on the distributions these data exhibit and the relationships that can be established between numeric variables using simple and advanced statistical methods. The common tools for estimating simple bivariate measures of association are correlation and cross-tabulation, where statistics help establish the magnitude, direction, and significance of the association between two variables. For example, at the individual level, there is a strong, positive, and significant bivariate relationship between years of formal education and income. At the global level of

analysis, there is a strong, positive, and significant relationship between per capita income and the level of democracy (e.g. Rueschemeyer, Stephens and Stephens 1992; Przeworski, Alvarez, Cheibub and Limongi 2000; Foweraker and Landman 2004). Both of these relationships are represented by significant and positive correlation coefficients or other measures of statistical association. The common tool for estimating more complex and 'multivariate' relationships is some form of regression analysis, which determines the magnitude, direction, and significance of the independent relationships between the two or more explanatory variables and the outcome that is to be explained (see, e.g. Lewis-Beck 1980; Bohrnstedt and Knoke 1988; Fox 1997). For example, the basic relationship between education and income may want to add age of the individual to the analysis, while the relationship between development and democracy may want to add the world position of the countries in the sample (e.g. Burkhart and Lewis-Beck 1994).

In either example, multivariate analysis controls for these other factors to determine whether the original relationship is upheld. The results of this kind of analysis provide measures of association between all the explanatory variables and the outcome variable, which allows the analyst to determine their relative strength, magnitude, and statistical significance. If the original explanatory variable of interest maintains its significant relationship with the outcome variable in the presence of other explanatory variables, then it is possible to conclude that the original relationship has been upheld. The other explanatory variables are considered 'controls'. Thus for the education-income example or the development-democracy example above, age and world position serve as controls. If the relationship between education and income remains significant even after including age, then the original relationship is said to be upheld after controlling for the effects of age. In similar fashion, if the relationship between development and democracy remains even after including world position, the original relationship is upheld. These basic insights are present in all multivariate analyses of human rights violations, where typical control variables include population, regional position, British colonial influence, among others (see Poe and Tate 1994; Poe and Carey 2004; Landman 2005a, 2005b).

Qualitative methods seek to identify and understand the attributes, characteristics, and traits of the objects of enquiry, as well as the meanings, processes, and context (Devine 1995:139; Devine and Heath 1999). The nature of the method necessarily requires a focus on a small number of units of analysis, whether they be individuals, groups, sub-national regions, countries, or supra-national regions. As discussed across many of the examples above, qualitative methods include macro-historical comparison (Skocpol and Somers 1980; Ragin, Berg-Scholsser and de Meur 1996; Mahoney and Rueschemeyer 2003); in-depth interviews and participant observation (Devine 1995); interpretivism, hermeneutics, and 'thick description' (Geertz 1973; Fay 1975); and varieties of discourse analysis (Howarth 2000a; Travers 2001). In none of these types of method is there an attempt to give numerical expression to the objects of enquiry, and in all of them the goal is to provide well-rounded and complete discursive accounts. These more complete accounts are often referred to as 'ideographic' or 'configurative', since they seek to identify all the elements important in accounting for an outcome.

There has traditionally been a deep division in the social sciences between those who use quantitative methods and those who use qualitative methods; however, it seems that this division is a false one for several reasons. First, the strict separation between

quantitative and qualitative methods is minimized if both methods adhere to the goal of making inferences from available evidence (Foweraker and Landman 1997:48–49; Travers 2001:6–9). The same logic of inference ought to apply equally to quantitative and to those qualitative methods that seek to move beyond pure description (see King, Keohane and Verba 1994; Brady and Collier 2004). Second, the qualitative distinction made between and among categories in any attempt to classify social phenomena necessarily precedes the process of quantification (Sartori 1970, 1994). In this sense, social science needs to know ‘what kind’ of object to count before counting it, and this qualitative step as vitally important in the quantification of human rights (see Chapter 5). Third, and related to the first two points, there have been important and significant methodological developments in combining the strengths of qualitative and quantitative techniques by recognizing that both methods are founded on the same logic of inference and linking qualitative distinctions to quantitative representation. These developments include ‘qualitative comparative analysis’ (Ragin 1987, 1994) and the use of Boolean algebraic techniques to identify necessary and sufficient conditions for outcomes; text and content analysis, which codes words into numbers (Franzosi 2004); and the quantitative deconstruction of victim testimonies to truth commissions (Ball, Spierer and Spierer 2000; see also Wilson 1997, 2001).

Qualitative comparative analysis combines the strengths of both the ‘case-oriented’ approaches and ‘variable-oriented’ approaches, while maintaining the scientific rigour of comparative studies that employ quantitative methods (Ragin 1994:304). The analytical framework considers cases holistically as the configuration of conditions rather than relationships between scores on variables (Ragin, Berg-Schlosser and de Meur 1996). It is the combination of variables rather than their independent relationships that is privileged in this form of analysis. In its simpler form, it constructs a matrix, or ‘truth table’ of ‘causal factors’ that are dichotomized according to their presence or absence within each case. These factors are then listed alongside a dichotomized form of the outcome that is to be explained. By examining the presence or absence of the causal factors and their various combinations, the analysis can find a reduced form of a combination of factors that is associated with the outcome (see Mahoney and Goertz 2004:658–660). This kind of analysis has been used to study the necessary and sufficient conditions for social revolution in Latin America (Wickham-Crowley 1992), the division of working-class movements in the process of nation building in Western Europe (Ragin 1987:126–133); the degree of ethnic mobilization among sub-national groups in Western Europe (Ragin 1987:133–149); the relationship through time between social mobilization and the protection of individual citizenship rights (Foweraker and Landman 1997); and the conditions for democracy in inter-war Europe (de Meur and Berg-Schlosser 1994). More advanced forms of this analysis move beyond dichotomous variables to those with multichotomous categories (Ragin, Berg-Schlosser and de Meur 1996:758), and the technique is very promising for carrying out human rights impact assessments (see Chapter 8 this volume).

The development in the quantification of texts draws on a long tradition in behavioural social science, which sought to break down reality into its smallest countable units (i.e. words) and then aggregate these units into larger systems of meaning and understanding (Eulau 1996). For example, in the 1940s in the United States the Experimental Division for the Study of War Time Communications at the Library of Congress analysed

newspaper and newswire stories quantitatively to uncover patterns in perception of the Allied and Axis powers among significant third countries (Eulau 1996:58–59). In drawing on this tradition, succeeding generations of behaviouralists have used newspaper sources to code events such as riots, strikes, protest, assassinations, repression, and general social mobilization across the world (see, e.g. Taylor and Hudson 1972; Taylor and Jodice 1983; Tarrow 1989; Banks 1994; Kriesi, Koopmans, Dyvendak and Giugni 1995), while other analyses have relied on multiple newswire sources to reduce bias and expand coverage (Lichbach 1984; Gerner, Schrodt, Francisco and Weddle 1994). In practical terms a coding vocabulary is applied to the raw text, which converts the text into quantitative information that recorded instances of events and outcomes of interest to the researcher. In this way, the grammar of a story or narrative is converted into a series of countable units (see Franzosi 2004).

There have been similar developments in the use of such coding techniques in the human rights field, particularly in relation to the work of truth commissions. Here, narrative accounts of human rights violations that have taken place during times of civil conflict, authoritarian rule, or foreign occupation are collected by truth commissions and then coded using human rights violation vocabularies. These coded statements are then organized into large relational databases where individual violations serve as the basic unit of analysis. Oftentimes, the official databases of the truth commissions are complemented further with additional data that have been collected by non-governmental organizations. The resulting data analysis estimates the total number of violations that have taken place (with a margin of error), and then disaggregates the pattern of violations across time, space, and characteristics of the victims and the perpetrators (see Ball, Spierer and Spierer 2000; Ball, Asher, Sulmont and Manrique 2003; Landman 2004). Truth commissions in El Salvador, Guatemala, Haiti, South Africa, Peru, Sierra Leone and East Timor have carried out data projects, which have fed into their final reports and have had significant implications for processes of reparation and reconciliation (see Chapter 7).

Methods matter

In its analysis and discussion of the different methodological traditions and options in the social sciences, this chapter shows how and why method matters for the social scientific analysis of human rights problems. Methods link theory and evidence, provide the basic rules of enquiry, and provide the tools that maximize the kind of inferences that are drawn. The chapter argued that there is not one preferred method, since method is a function of the epistemological orientations of the researcher, the theoretical perspective that is adopted, the nature of the research question, as well as the available time and material resources with which to carry out any research project. It is clear that methods vary across the epistemological continuum from ‘clouds’ to ‘clocks’, the full range of comparative analysis from global comparative to single-country studies, and the degree to which qualitative, quantitative, or hybrid methods are adopted. Nevertheless, method is a central feature of all systematic social research and is not separate from the substance of the research that is being carried out. Rather, it is the substance since it sets the parameters over what can be said about the research problem and makes possible the types of knowledge claims that can be safely advanced given the evidence that has been

considered (see King, Keohane and Verba 1994:9). Good human rights scholarship and good human rights arguments need strong methodological foundations that specify the ways in which human rights problems will be addressed, how human rights evidence will be collected and analysed, and how human rights conclusions will be drawn. Analysing human rights problems with bad methods will lead to erroneous conclusions, bad policy advice, and failure to improve human rights conditions on the ground.

Suggestions for further reading

- Brady, H.E. and Collier, D. (2004) *Rethinking Social Inquiry: Diverse Tools, Shared Standards*, Lanham, MD: Rowman and Littlefield.
- Flyvberg, B. (2001) *Making Social Science Matter: Why Social Inquiry Fails and How It Can Succeed Again*, Cambridge: Cambridge University Press.
- King, G., Keohane, R.O. and Verba, S. (1994) *Designing Social Inquiry: Scientific Inference in Qualitative Research*, Princeton, NJ: Princeton University Press.
- Landman, T. (2002) 'Comparative Politics and Human Rights', *Human Rights Quarterly*, 24(4): 890–923.
- Landman, T. (2003) *Issues and Methods in Comparative Politics: An Introduction*, 2nd edn, London: Routledge.

5

Measuring human rights

Social scientific analysis of human rights problems often depends on the use of valid, meaningful, and reliable measures of human rights across the different categories and different dimensions outlined in Chapter 1.¹ There is now a large literature on human rights measurement that has developed through contributions from the academic disciplines of political science, sociology, economics, and law, as well as from governmental and non-governmental organizations (e.g. Claude and Jabine 1992; Green 2001; Landman and Häusermann 2003; Landman 2004). Long seen as the purview of development and economic analysis, the use of indicators has increasingly been brought into mainstream analysis of human rights in the social sciences and in the work of the United Nations and the World Bank. For example, since the late 1970s and early 1980s, political scientists have been using quantitative measures in global comparative analyses of human rights protection (see Chapter 6 in this volume). The 2000 Human Development Report dedicated an entire chapter on why human rights indicators are important and how they could be incorporated into its own work (UNDP 2000:89–111). The Office of the High Commissioner for Human Rights has been exploring the ways in which human rights indicators could be used by the treaty monitoring bodies to assist in their assessments of state compliance with the different international human rights instruments. In his annual reports to the UN General Assembly, the Special Rapporteur on the Right to Health has advocated the use of indicators to gauge different aspects of the progressive realization of the right to health (A/58/427/2003; A/59/422/2004). Finally, the World Bank has incorporated measures of human rights in its work on good governance and its relationship with development (see Kaufmann *et al.* 1999a, 1999b, 2002, 2003; <http://www.worldbank.org/>).

There is thus a groundswell of support and increasing level of interest in the development and use of human rights measures. Such measures serve a variety of important and inter-related functions across the academic and non-academic sectors of the human rights community. First, they allow for *contextual description and documentation*, which provide the raw information upon which measures of human rights are based. Second, they help efforts at *classification*, which differentiates rights violations across their different categories and dimensions. Third, they can be used for *monitoring* the degree to which states respect, protect and fulfil the various rights set out in the different treaties to which they may be a party. Fourth, they can be used for *mapping and pattern recognition*, which provide time-series and spatial information on the broad patterns of violations within and across different countries. Fifth, they are essential for *secondary analysis*, including hypothesis-testing, prediction and impact assessment, the inferences from which can be fed into the policy making process. Finally, they can serve as *important advocacy tools* at the domestic and international levels by showing the improvement or deterioration in rights practices around the world. The accumulation of

information on human rights protection in the world and the results of systematic analysis can serve as the basis for the continued development of human rights policy, advocacy, and education (Rubin and Newberg 1980:268; Claude and Jabine 1992:5–34). Moreover, ‘to forswear the use of available, although imperfect, data does not advance scholarship’ (Strouse and Claude 1976:52), nor does it allow for the kind of continued human rights activism that seeks to eliminate the worst forms of human behaviour.

This increasing demand for human rights measures and the various functions that they perform across a wide spectrum of scholars and practitioners suggests that they will continue to be a central feature of systematic human rights research and applied work in the field. In order to demonstrate how to measure human rights and to examine the many remaining challenges human rights measurement faces, this chapter is divided into three sections. The first section explains how social scientific measurement moves through four different levels ranging from general background concepts to specific scores on specific human rights across specific units of analysis (e.g. a high score on civil rights CR_{\uparrow} in country X in year T). The second section discusses extant measures of human rights, including those that measure rights ‘in principle’ (i.e. *de jure* state commitment), ‘in practice’ (i.e. *de facto* realization), and as a government ‘policy’ (i.e. inputs, outputs, and outcomes). The third section identifies areas in need of better and more appropriate measures to fill important lacunae that remain.

From concepts to indicators

Chapter 1 showed how the scope of human rights includes different categories (civil, political, economic, social, cultural, and solidarity) *and* different dimensions (positive and negative), the combination of which produces six main conceptual ‘boxes’ (see Table 1.1) that need to be operationalized. The different dimensions and categories provide the content for developing ‘events-based’, ‘standards-based’, ‘survey-based’ and other measures of human rights (see the next section). But what are the operational steps that allow us to move from these conceptual distinctions of human rights to the provision of valid, meaningful, and reliable measures? At an abstract methodological level, the process of measurement converts well-defined and well-specified concepts into meaningful quantitative measures or qualitative categories, and has four main steps (Adcock and Collier 2001; also Zeller and Carmines 1980). The first level concerns the *background concept* that is to be measured, which is the broad constellation of meanings and understandings associated with the concept. The scope of human rights outlined in Chapter 1 summarizes what comprises such a broad constellation of meanings and understandings in the field of human rights. The second level develops the *systematized concept*, which specifies further the concept that is to be measured, such as a specific right (e.g. the right not to be tortured) or a group of rights (e.g. civil rights). The third level *operationalizes the systematized concept* into meaningful, valid, and reliable indicators, where decisions are taken as to the type of measure that is to be used and how it captures the positive and negative dimensions of the rights it measures. The final level provides *scores on indicators* for the units of analysis being used (e.g. individuals, groups, countries, regions, etc.). Figure 5.1 depicts these four levels graphically.

Consider a concrete example. The *background concept* to be measured is human rights, the scope of which has been *systematized*. The international community of human rights

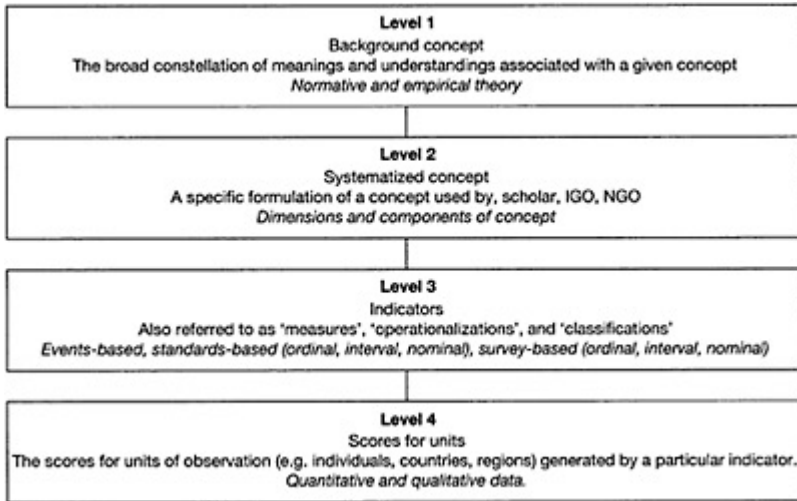


Figure 5.1 Levels of measurement.

Sources: Zeller and Carmines 1980; Adcock and Collier 2001.

scholars and practitioners have spent the years before and since the 1948 Universal Declaration of Human Rights 'constructing' (Donnelly 1999b) and 'justifying' (Sorell and Landman 2005) human rights in conceptual and legal terms. While there have not been agreed philosophical foundations for the existence of human rights (Mendus 1995; Landman 2004, 2005a), the extant international law of human rights provides a general consensus on the core content of those human rights that ought to be protected (Landman and Häusermann 2003), a consensus that found its most forceful and comprehensive expression in the 1993 Vienna Declaration and Programme for Action (Boyle 1995; Donnelly 1999a; Freeman 2002b). Such a core content represents the steps taken from the background concept to the systematized concept within the four levels of measurement depicted in Figure 5.1.

The intersection between the categories and dimensions of human rights is a systematic way of organizing the first step to measurement. Consider the right not to be tortured, which is a systematized concept of human rights that has been identified most notably in the Universal Declaration of Human Rights (UDHR), the International Covenant on Civil and Political Rights (ICCPR), and the Convention Against Torture (CAT). The systematized concept is susceptible to operationalization at Level 3. But given the two dimensions of human rights, the right not to be tortured can be measured at Level 3 both *positively* (i.e. resources a state is investing in procedures, policies, reforms, and training for the prevention of torture) and *negatively* (formal commitment to

international standards on torture and actual incidence of torture). At Level 4, the right not to be tortured is measured for a unit (e.g. Brazil) at a particular time (e.g. 1985), across its positive dimension (e.g. % GDP spent on torture reform, number of police in receipt of torture training, cases of reprimand for torture) and its negative dimension (e.g. incidence of torture revealed through events counting, a scale of torture, or survey estimations on popular experiences of torture). In this way, the right not to be tortured *may have several indicators that measure its core content across its two dimensions*.

Extant measures

In many ways, the proliferation of human rights norms and the promulgation of international human rights treaties since the Universal Declaration have effectively (with minor exceptions) completed the move from background concepts (Level 1) to systematized concepts (Level 2), while the move from providing indicators (Level 3) to score on units (Level 4) comprises the continuing work of scholars and practitioners working on human rights measurement today. There is thus not yet a full set of indicators available that measures all human rights across their different categories and dimensions. Rather, those that are available measure certain categories of human rights across one or another of their dimensions. Extant approaches have measured human rights *in principle* (i.e. as they are laid out in national and international legal documents), *in practice* (i.e. as they are enjoyed by individuals and groups in nation states), and as outcomes of government *policy* that has a direct bearing on human rights protection. Measurement of human rights has taken the form of coding country participation in regional and international human rights regimes, coding national constitutions according to their rights provisions, qualitative reporting of rights violations, survey data on perceptions of rights conditions and experiences, quantitative summaries of rights violations, abstract scales of rights protection based on normative standards, and individual and aggregate measures that map the outcomes of government policies that have consequences for the enjoyment of rights.

Rights in principle

International and domestic law enshrines norms and principles of human rights, which can be coded using protocols that reward a country for having certain rights provisions in place at the domestic level and for having made such rights commitments at the international level through the ratification of international human rights treaties. Such a coding represents a negative dimension of state practice towards human rights since making constitutional provisions at the domestic level and signing and ratifying treaties at the international level carries with it only nominal cost in terms of actual fiscal capacity of the states. It is therefore a formal commitment *in principle* that can be counted. In an important precedent for measuring rights in principle at the domestic level, van Maarseveen and van der Tang (1978) coded constitutions for 157 countries across a multitude of institutions and rights for the period 1788–1975. Their study compares the degree to which national constitutions contain those rights mentioned in the UN Declaration for Human Rights by examining their frequency distributions across different

historical epochs before and after 1948 (see Figures 5.2 and 5.3). Their study is broadly descriptive in nature, but their data allow for global patterns and processes of change in the formal protection of rights at the domestic level to be mapped, while secondary and more advanced statistical analysis could be conducted on the patterns within the data while exploring possible relationships with other indicators. In the spirit of this precedent, Foweraker and Landman (1997:51–52) use an ‘institutional procedural index’ to code rights in principle using the various national constitutions, constitutional amendments, and executive decree laws during the years of political liberalization and

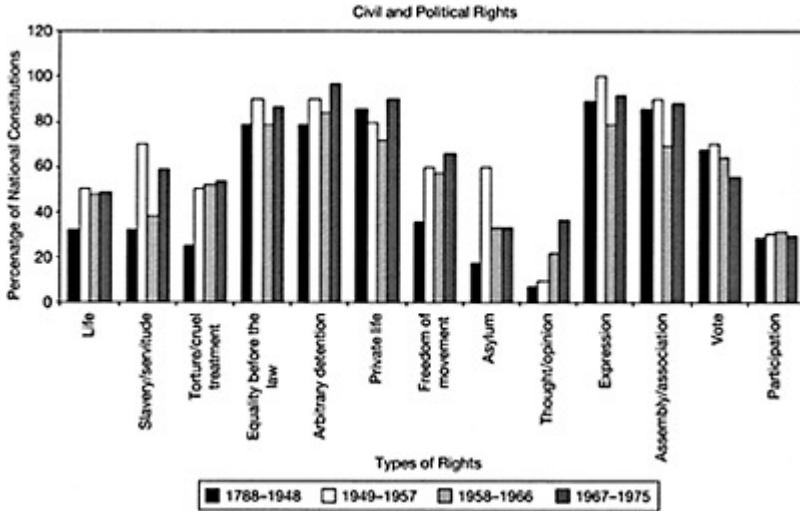


Figure 5.2 National constitutional provisions for rights in principle: civil and political rights.*

Source: van Maarseveen and van der Tang (1978:189–211).

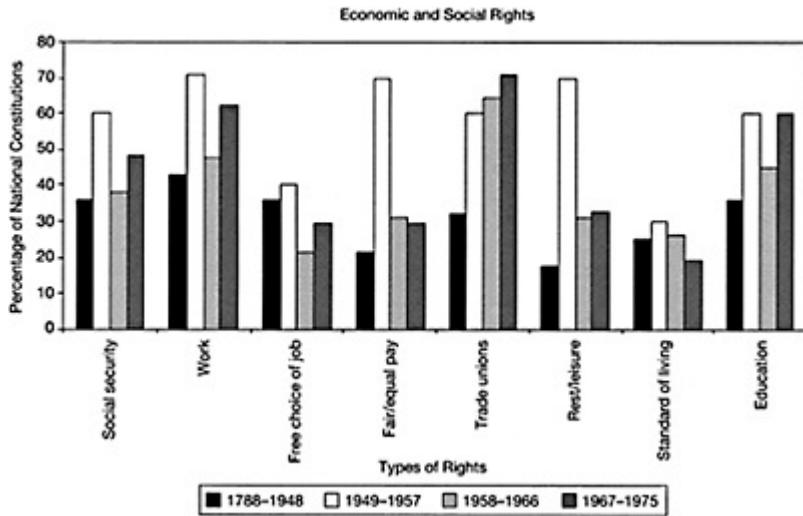


Figure 5.3 National constitutional provisions for rights in principle: economic and social rights.*

Source: van Maarseveen and van der Tang (1978:189–211).

*Figures 5.2 and 5.3 are reprinted with permission from *Human Rights Quarterly*, © 2004 Johns Hopkins University Press. All rights reserved.

democratic transition in Brazil, Chile, Mexico, and Spain. In similar fashion, Keith (2004) codes national constitutions and their exceptional clauses to measure the ability of countries to suspend rights protection during states of emergency.

Coding rights in principle is important since it translates qualitative legal information into quantitative information that can be used to track the formal commitment of countries to rights protection against which their actual practices can be compared. Thus, Foweraker and Landman (1997:62–65) use regression techniques to gauge the relative ‘gap’ between rights in principle and rights in practice in Brazil, Chile, Mexico, and Spain (see also Duvall and Shamir 1980:162–163; Arat 1991). Their analysis demonstrates that during the process of political liberalization, authoritarian states can deny rights that they proclaim are protected (a negative gap), protect rights they proclaim are protected (a zero gap), or protect rights that they proclaim are not protected (a positive gap). Keith (2004) uses the state of emergency variable to examine the relationship between rights in principle and rights in practice, while controlling for the independent effects of democracy, wealth, warfare, and other variables.

This development of coding schemes at the domestic level has been replicated at the international level by scoring countries for signing and ratifying major regional and international human rights instruments (see Keith 1999; Landman 2001, 2005b; Hathaway 2002). Rather than code individual rights provisions, the coding schemes

measure the degree to which countries are parties to human rights treaties over time. Keith (1999) and Hathaway (2002) used a simple dichotomous measure of ratification (0, 1), while Landman (2005b) combines a three-point coding scheme for treaty ratification (0, 1, 2) with a four-point coding scheme that rewards and punishes countries for the degree to which they make reservations upon ratification, where a 4 denotes a country that ratifies with no reservations and a 1 denotes a country that has made significant reservations that undermine the object and purpose of the treaty. Combining the ratification variable with the reservation variable produces a weighted ratification variable that ranges from 0 to 8.

The time-series analysis of regime participation shows an expansion in the number of countries that are now formally participating in the international regime (see Figure 1.1 in Chapter 1 this volume); that ‘late ratifiers’ tend to ratify more treaties with fewer reservations than ‘early ratifiers’; and that democracies have a greater propensity to ratify treaties, even though ‘old’ democracies (i.e. those countries that were democratic before the 1970s) ratify fewer treaties with more reservations than new democracies (Landman 2005b). These measures of treaty ratification are also used in global analysis to show that regimes frequently make formal commitments to human rights treaties, but continue to violate human rights. This difference is captured by weak positive or even negative correlation and regression coefficients between ratification and rights variables (Keith 1999; Landman 2001; Hathaway 2002; see also Krasner 1999:122). Moreover, Landman’s (2005b) analysis shows that the significant relationship between treaty ratification and rights protection weakens when taking reservations into account. Carrying out such analyses, however, requires measurement of rights in practice and it is to this that the discussion now turns.

Rights in practice

Rights in practice are those rights actually enjoyed and exercised by groups and individuals regardless of the formal commitment made by a government. While there ought to be a correspondence between formal rights commitments found in national constitutions and international human rights instruments and those enjoyed on the ground, it is often the case that individuals and groups do not enjoy the full protection of their rights (a negative gap in the terminology used above). Ideally, there ought to be in place a legal appeals procedure, mechanisms for seeking domestic and international remedies, and a subsequent ‘correction’ in national practices to uphold the rights to which regimes have made formal commitments. In the absence of such systems or in the face of weak systems, the role of many human rights practitioners is to provide meaningful and accurate information on the degree to which human rights are being violated. Indeed, greater concerns over human rights since World War II have led to an explosion in the number of domestic and international human rights NGOs collecting information on violations (see Chapter 2 this volume). Such NGOs have been given greater status in international governmental organizations, and their activities include setting standards, providing information, lobbying, and giving direct assistance to those suffering abuse of their rights (Forsythe 2000:163–190; Welch 2001a:1–6; Landman and Abraham 2004).

The increase in the salience of human rights as an issue combined with organizations dedicated to documenting human rights violations means that there is greater availability

of comprehensive information on actual practices of states and the conditions under which individuals live. But this information is necessarily lumpy and incomplete, since reporting of human rights violations is fraught with difficulties, including fear within victims, power of the offenders, comprehensive evidence, quality of communications technology, among others. In recognizing this problem, Bollen (1992:198) argues that there are six levels of information on human rights violations: (1) an *ideal* level with *all* characteristics of *all* violations (either reported or unreported), followed by (2) recorded violations, (3) known and accessible violations, (4) locally reported violations (nation state), (5) internationally reported violations, and (6) the most biased coverage of violations, which may include only those reported in US sources.

Work in this area seeks to obtain lower levels of information in much greater detail. For example, the *Torture Reporting Handbook* (Giffard 2002) and *Reporting Killings as Human Rights Violations* (Thompson and Giffard 2002) are manuals that define specific rights, outline the legal protections against their violation, and provide ways in which testimony and evidence from victims can be collected. The Human Rights Information and Documentation System (HURIDOCs), founded in 1982, provides standards for human rights violations reporting, and now represents a vast network of human rights groups (Dueck 1992:127). While such increased information at all levels is helpful for systematic human rights research, there remains a trade-off or tension between micro levels of information gathering and the ability to make systematic comparative inferences about human rights (see Chapter 4 this volume). In order for equivalent measures to ‘travel’ for comparative analysis, there will necessarily be some loss of information, while the comparability of measures allows for stronger generalizations about human rights violations to be drawn.

These issues about levels of information and the commensurability for cross-national analysis delineate the three types of data available for measuring human rights in practice: (1) *events-based*, (2) *standards-based*, and (3) *survey-based*. Events-based data chart the reported acts of violation committed against groups and individuals. Events-based data answer the important questions of what happened, when it happened, and who was involved, and then report descriptive and numerical summaries of the events. Counting such events and violations 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. 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, Spierer and Spierer 2000; and see Chapter 7 this volume). Standards-based data establish how often and to what degree violations occur, and then translate such judgements into quantitative scales that are designed to achieve commensurability. Such measures are thus one level removed from event counting and violation reporting, and merely apply an ordinal scale to qualitative information. Finally, survey-based data use random samples of country populations to ask a series of standard questions on the perception of rights protection. Such measures track individual-level perceptions of rights violations and may even capture direct or indirect individual experiences of rights violations, particularly in countries that have suffered from prolonged periods of civil conflict, authoritarian rule, or foreign occupation (see Chapter 7).

There are by now many examples of each of these kinds of data that measure the negative dimensions of human rights. Events-based data analysis has a long tradition,

where one of the first applications of statistics to the study of violence analysed the distribution of more than 15,000 'quasi-judicial' executions carried out during the height of the Reign of Terror (March 1793 to August 1794) after the French Revolution. Using the archived documents of the tribunals that sentenced people to death, Greer (1966) analyses the patterns of sentencing and executions over time, space, and by social class (nobles, upper middle class, lower middle class, clergy, working class, and peasants). Figure 5.4 shows the time-series pattern in executions by social class, while Figure 5.5 shows the number of executions across space (*départements*). The peasants and working classes suffered the largest number of executions, where the majority of the executions (52%) took place in the West, followed by the Southeast (19%) and Paris (16%). In this analysis, the individual victim serves as the basic unit of analysis, which allows for the kind of secondary analysis shown in the figures and the further testing of hypotheses about the causes of the violations, such as those based on class, political, economic, and/or religious variables (Greer 1966:4).

Ball and Asher (2002) conducted a similar style of statistical analysis on the patterns of killings and refugee migration of Albanians in Kosovo between 24 March and 22 June 1999. The analysis tests whether the violence and migration during this period were due to the activities of the Kosovo Liberation Army (KLA), the NATO air attacks, or a systematic campaign by Yugoslav forces. Using detailed border records of the refugee population that left Kosovo combined with UNHCR refugee data and four sources (American Bar Association, ICTY exhumation data, Human Rights Watch, and the OSCE) of data on the killings that took place during the period, allowed the refugee migration and killings to be plotted over time (see Figure 5.6). The more advanced statistical analysis of these patterns that controlled for region, KLA activities, and NATO bombing, determined that neither the KLA nor NATO could be held responsible for the killings, while KLA activities were associated with increased refugee flows in the Northern and Eastern regions of Kosovo (Ball and Asher 2002:22). In the absence of detailed data on Yugoslav troop movements, the analysis could neither support nor reject the hypothesis that those forces were responsible for the migration and killings. However, it was able to reject the two hypotheses about the KLA and NATO, a finding that when presented to the ICTY undermined significantly any attempt by the defence team to shift blame away from the Yugoslav forces (Ball and Asher 2002:24).

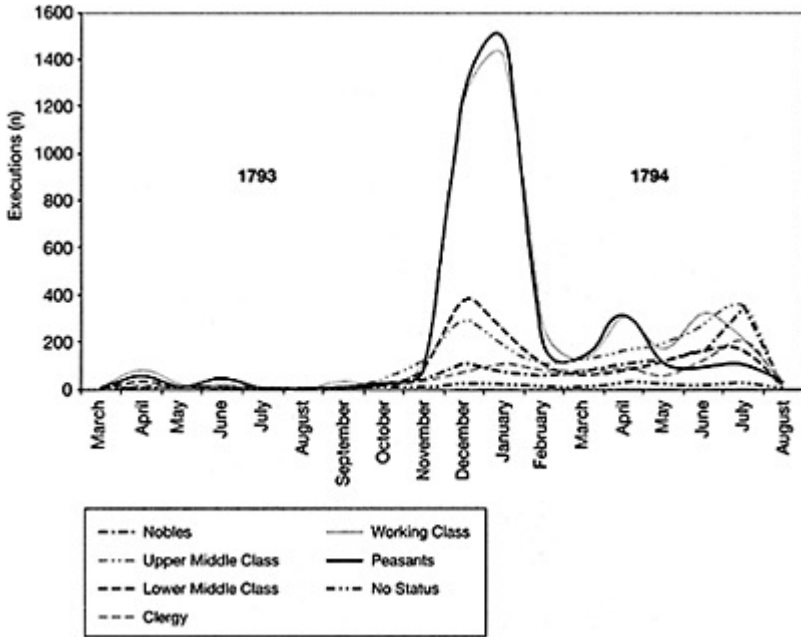


Figure 5.4 Number of executions during the Reign of Terror by social class.

Source: Greer (1966:135–166).

This discussion of events-based data illustrates that highly disaggregated data from particular historical contexts provide an important means for explaining what happened and how, as well as ruling out rival explanations (i.e. how such events could not have happened). Events-based analysis has been used to chart the progression of the Rwanda genocide (<http://www.genodynamics.com/>), to estimate civilian mortality rates before and after the invasion of Iraq in March 2003 (Roberts, Lafta, Garfield, Khudhairi and Burnham 2004), and to estimate the total number of civilian deaths as a result of the war in Iraq (<http://www.iraqbodycount.net/>). It has also become a central feature of many truth commissions (see Chapter 7), such that limited time-series events data have been collected, most notably in Argentina, Chile, Guatemala, El Salvador (still under UN seal), Haiti (not yet published), South Africa, Peru, Sierra Leone, and East Timor.

In contrast to events-based data, standards-based scales provide much more aggregated forms of information that have been collected and coded across a large number of countries. The most dominant examples of extant standards-based scales of human rights include the Freedom House scales of civil and political liberties (Gastil 1978, 1980, 1988, 1990; <http://www.freedomhouse.org/>), the ‘political terror scale’ (Mitchell, Stohl, Carleton and Lopez 1986; Poe and Tate 1994; Gibney and Stohl 1996), a scale of torture (Hathaway 2002), and a series of seventeen different rights measures

collected by Cingranelli and Richards (<http://www.humanrightsdata.com/>). Freedom House has a standard checklist it uses

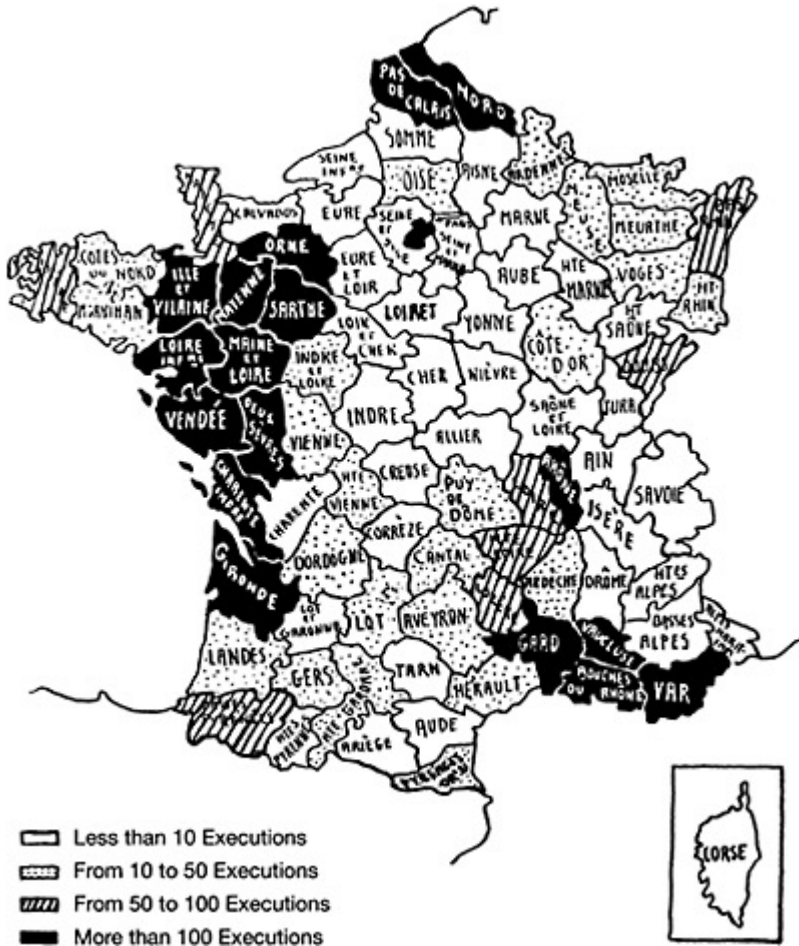


Figure 5.5 Number of executions during the Reign of Terror by *département*.

Reprinted with permission of Peter Smith Publishers Inc., © 1966. All rights reserved.

to code civil and political rights based on press reports and country sources about state practices and then derives two separate scales for each category of rights on a scale that ranges from 1 (full protection) to 7 (full violation). The political terror scale ranges from

1 (full protection) to 5 (full violation) for state practice that include torture, political imprisonment, unlawful killing, and disappearance. Information for these scales comes from the US State Department and Amnesty International country reports. In similar fashion, Hathaway (2002) measures torture on a 1 to 5 scale using information from the US State Department. The Cingranelli and Richards human rights data code similar sets of rights on scales from 0 to 2, and 0 to 3, with some combined indices ranging from 0 to 8, where higher scores denote better rights protections. In addition to a series of civil and political rights, Cingranelli and Richards also provide measures for such rights as women's economic, social, and political rights, worker rights, and religious rights.

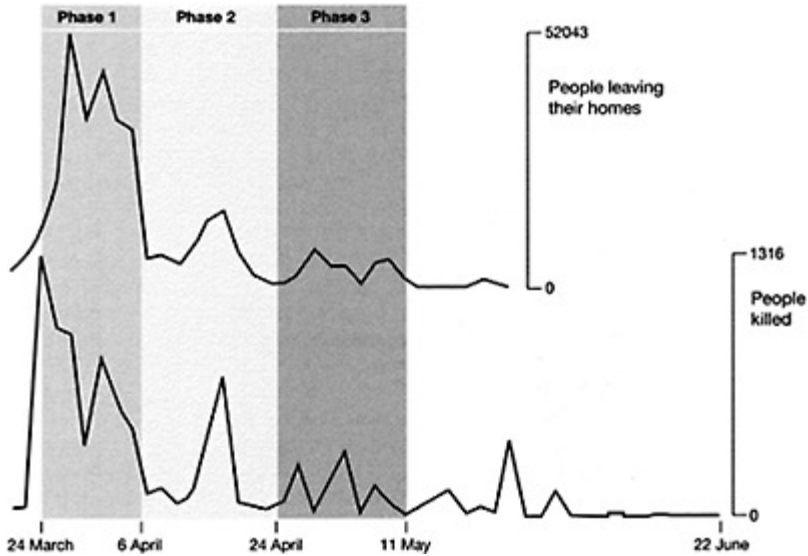


Figure 5.6 Estimated total refugee migration and killing over time in Kosovo.

Source: Ball and Asher (2002:21). Reprinted with permission from *Chance*, © 2002, by the American Statistical Association. All rights reserved.

Figures 5.7 and 5.8 show the Freedom House, political terror and torture scores over time and space for the period 1976 to 2000 (standardized to a common 1–5 scale). Figures 5.9 and 5.10 show the Cingranelli and Richards data for women's economic, political, and social rights on a 0–3 scale over time and space for the period 1981 to 2003.

Survey data have been less used in social scientific research on human rights than either events-based or standards-based measures. They have usually featured more often in research on the support for democracy (e.g. Kaase and Newton 1995), trust and social capital (e.g. Whiteley 1999, 2000), patterns of corruption (<http://www.transparency.com/>), or as components of larger indices of 'post-material'

values (see Inglehart 1997). But increasingly, household surveys have been used to provide measures for popular attitudes about rights and to uncover direct and indirect experiences of human rights violations. Some of the most notable work has been carried out by the NGO Physicians for Human Rights, who conduct surveys of ‘at risk’ populations (e.g. internally displaced people or women in conflict) to determine the nature and degree of human rights violations. Figure 5.11 shows the results of one of their surveys on war-related sexual violence in Sierra Leone based on 991 women (Physicians for Human Rights 2002b:47–48; see also Amowitz *et al.* 2002).

While these examples of human rights measures focus on civil and political rights, Chapter 1 in this volume argued that it is possible to extend the methodological discussion to include these kinds of measurement for economic, social, and cultural rights, as Cingranelli and Richards have begun to do. Indeed, if the denial of economic, social, and cultural rights is the product of particular government practices, then it seems

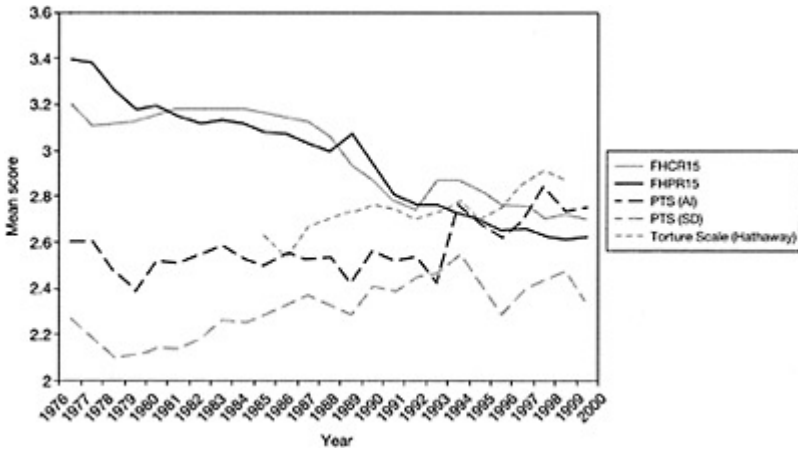
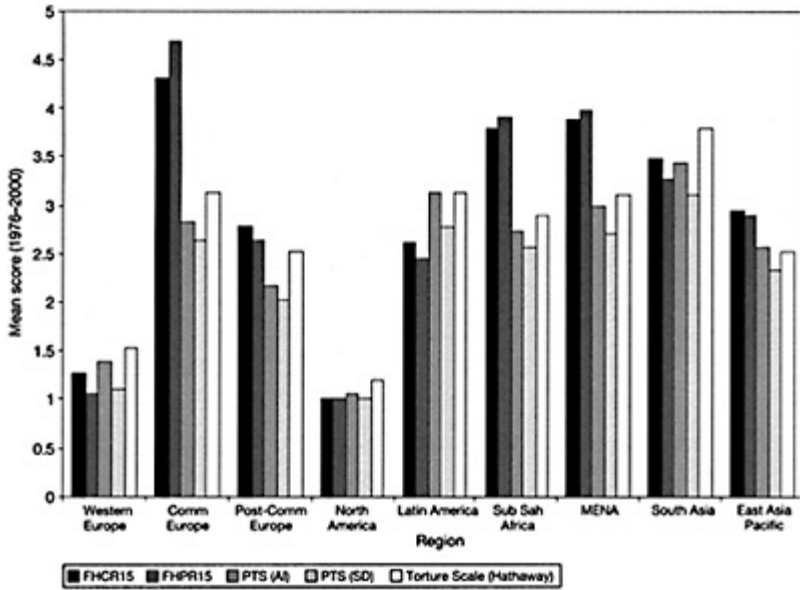


Figure 5.7 Standards-based scales of human rights protection over time, 1976–2000.*

Note: High score=high violations, low score=low violations.

Sources: Landman (2005b, 2005c).



*Figure 5.8 Standards-based scales of human rights protection by region, 1976–2000.**

Note: High score=high violations, low score=low violations.

Sources: Landman (2005b, 2005c).

*Both Figures 5.7 and 5.8 are reprinted with permission from Georgetown University Press, © 2005. All rights reserved.

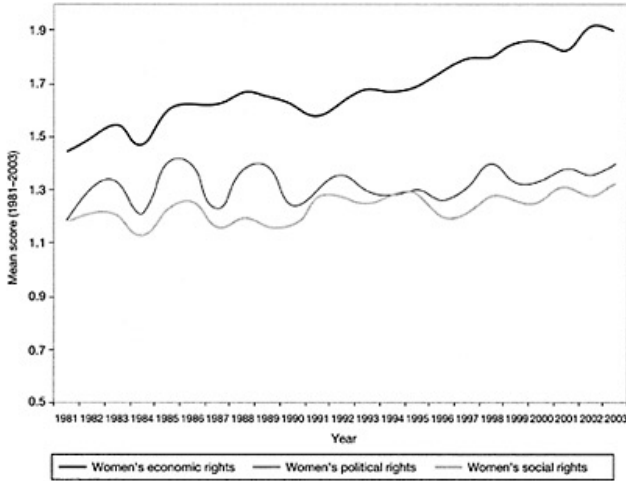


Figure 5.9 Women’s rights over time, 1981–2003.

Note: High=high protection, low score=low protection.

Source: Cingranelli and Richards data set (CIRI) (<http://www.humanrightsdata.com/>).

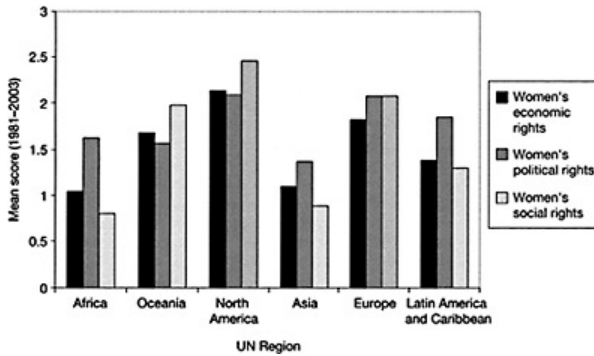


Figure 5.10 Women’s rights by UN region, 1981–2003.

Note: High=high protection, low score=low protection.

Source: Cingranelli and Richards data set (CIRI) (<http://www.humanrightsdata.com/>).

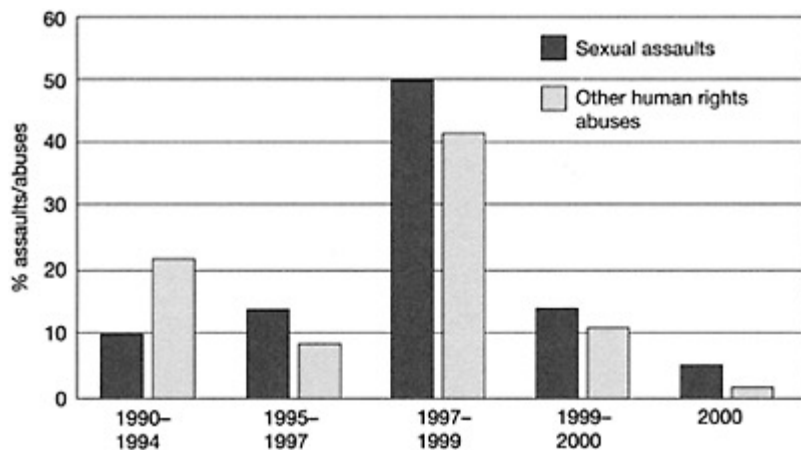


Figure 5.11 War-related sexual violence in Sierra Leone, 1990–2000 ($n=1,157$).

Source: Physicians for Human Rights (2002b:48).
Reprinted with permission.

equally possible to use qualitative information to summarize such practices into ordinal scales like those used for civil and political rights violations. Overt, institutionalized, or implicit discrimination against individuals or groups that prevents their access to education or adequate healthcare constitutes a practice that violates a right. In theory, such a violation can be reported and coded using events-based, standards-based, and/or survey-based data. Cingranelli and Richards are coding practices that violate women's rights using an ordinal scale, while the minorities at risk project codes the degree to which 224 different minority and communal groups experience discrimination also using an ordinal scale (see Gurr 1993, and also Foweraker and Krznicar 2001). These important precedents demonstrate how the negative dimensions of economic, social, and cultural rights can be measured.

Despite their development and increasingly wider use, these three types of data are fraught with methodological problems. Events-based data are prone to either under-reporting of events that did occur or over-reporting of events that did not occur, creating problems of selection bias and misrepresentative data. It is impossible to document every last human rights violation and those organizations collecting such information tend to concentrate on conflict-stricken societies during discrete periods of time and thus cross-country comparisons using such measures is problematic. Sophisticated statistical techniques have been developed to overcome some of the limitations associated with estimation of violations (see Chapter 7), but extant data projects of this nature necessarily concentrate on limited time periods in particular cases, and it may be that some episodes of violence are simply too complex to analyse using these techniques. In contrast, standards-based data establish comparability by raising the level of abstraction, but have a tendency to truncate the variation of human rights protection across different countries.

In other words, their use of a simple limited scale may group together certain countries that actually show a great difference in their protection of human rights. While these scales present a general picture of the human rights situation and are useful for drawing comparative inferences, they necessarily sacrifice the kind of specificity for pursuing direct legal action against perpetrators. Finally, survey data, especially those used across different political contexts are prone to cultural biases, where the meaning of standardized questions on rights protection are differently understood in different countries. In this way, the debate about the universality of human rights affects the method of measuring rights through surveys, since it is not obvious that human rights are understood to mean the same thing across the world. It is important therefore that those measuring human rights in practice *recognize the limits of their data*.

Government policies and outcomes

In addition to rights in principle and rights in practice, it is possible to provide more indirect measures of human rights using aggregate statistics on the outcomes of government policies. Parr (2002) makes the useful distinction between human rights conduct and developmental outcomes that may have a bearing on human rights. She stresses the fact that certain dimensions of conduct and outcomes are simply not prone to quantifiable measurement (see Radstaake and Bronkhurst 2002:31–32). In the language of this volume, her distinction fits well with the difference between rights in *practice* (conduct) and government *policy* (outcomes). But it appears that practices and outcomes are more readily quantifiable than Parr (2002) assumes. Traditionally, development studies and development economics have often relied on quantitative indicators of the outcomes of government policies, including gross domestic product, gross domestic product per capita, income inequality, expenditure on health, education, and welfare, among many others (e.g. see <http://www.worldbank.org/>). Indeed, the UNDP's human development index (HDI) combines per capita income (standard of living) with literacy rates (knowledge), and life expectancy at birth (longevity) (UNDP 1999:127–137). While not providing a direct measure of rights protection *per se*, such measures can elucidate the degree to which governments support activities that have an impact on human rights.

One solution is to combine the HDI with standards-based measures of human rights to get a more holistic picture that captures the interaction between levels of human development and the protection of human rights. Figure 5.12 is a scatter plot between the HDI and a 'factor score' created through principal component extraction from the two versions of the Political Terror Scale, the two Freedom House scales, and the torture measure outlined above. The assumption behind using factor analysis is that each of the five measures is measuring common human rights phenomena. The curvilinear cubic functional form in the figure provides the best overall fit for the relationship between human development and human rights (i.e. has the highest R^2), but using the UNDP's cut-off points for low, medium, and high human development also shows the areas of the world most in need of attention. For example, those countries with low human development and high violations of human rights depicted in the upper left section of the figure form the main 'constituency' for attention by the international community (see Sorell and Landman 2005). Such a combination of measures goes some way in depicting

the inter-relatedness of different categories of rights, and shows how aggregate development indicators can be combined with standards-based measures.

A second solution is to employ development indicators as proxy measures for the *progressive realization* of economic, social and cultural rights. 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 realization of these rights: steps in which states set goals, targets and timeframes for national plans to implement

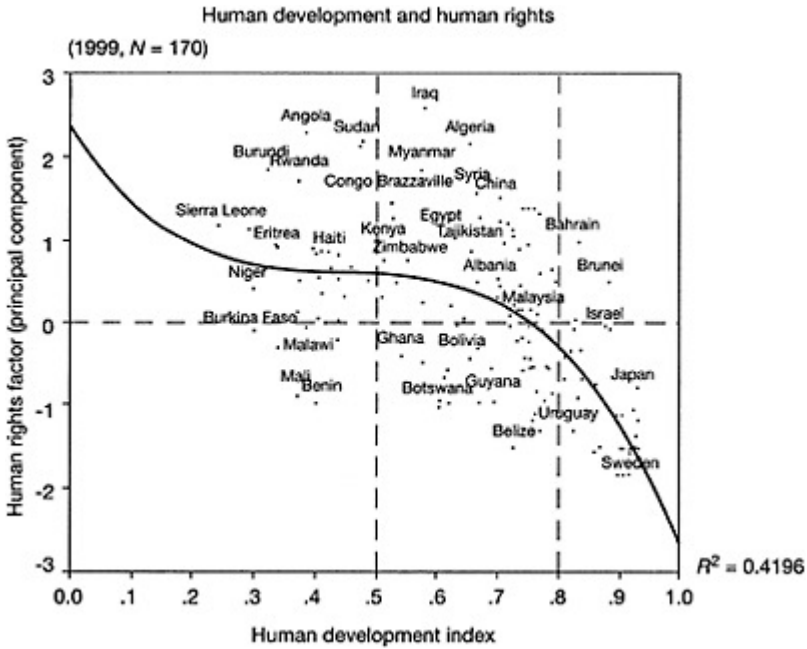


Figure 5.12 Scatter plot for civil and political rights and the human development index, 1999.

Sources: Landman (2005c), UNDP (1999:134–137).

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). Similar such measures can be used as indicators for the right to adequate food and housing.

To date, such development indicators have primarily been applied to economic and social rights, but aggregate statistics can equally be used to measure the positive dimensions of civil and political rights. Following the work of the United States Agency for International Development (USAID 1998a, 1998b), new efforts propose the use of development indicators as potential proxy measures for civil and political rights (e.g. investment in prison and police reform, the processing of cases, the funding of judiciaries, the provision of legal aid and advice to suspects, the amount of time suspects spend on remand, and the proportion of cases taken up by independent reporting and investigating bodies). The extension of such indicators for measuring cultural rights is also possible. The social and spatial mobility of ethnic and cultural minority populations, as well as spending on bilingual education can approximate the degree to which countries are adopting policies to uphold their cultural rights obligations. In short, aggregate measures of provision and outcomes can depict the degree to which governments are committed to putting in place the kinds of resources needed to have a 'rights-protective regime' (Donnelly 1999a).

Lacunae

This chapter has demonstrated how the different categories and dimensions of human rights can be converted, or operationalized into meaningful, valid, and reliable measures. The background concept of human rights has been systematized by the international legal and human rights community such that there is now a known core content of human rights susceptible to social scientific operationalization using a variety of indicators across their different categories and dimensions. These include the positive and negative dimensions of civil, political, economic, social, cultural, and solidarity rights. Efforts to operationalize these different dimensions and categories of human rights have included measures of rights in principle, rights in practice, and proxy measures of government policies and outcomes. To date, most efforts have concentrated on measuring rights in practice and include events-based, standards-based, and survey-based forms of measurement.

It seems clear, however, that we still know more about *what* to measure conceptually and legally than *how* to measure it. Tremendous progress in human rights measurement has been achieved but there are serious and significant lacunae in the field that need to be addressed that include both the *content of rights that remain unmeasured* and an *over-reliance on certain forms of measurement*. With respect to content that remains unmeasured, efforts in measurement have predominantly concentrated on the *negative* dimensions of civil, political and some cultural rights (e.g. minority rights discrimination) and the *positive* dimensions of economic and social rights. There is thus a dearth of measures for the positive dimensions of civil and political rights and the negative dimensions of economic and social rights. The human rights community thus still needs to develop measures for the provision of resources that support the protection of civil and political rights and measures for the violation of economic and social rights (see Chapman 1996). In this regard, the Cingranelli and Richards data set is the first attempt to measure systematically the negative dimensions of some social and economic rights. Moreover, there is less agreement on the content of solidarity rights and at best there have

been some proxy measures offered for them, such as the distribution of global income, trade dependency, and trade openness.

Finally, there has been over-reliance on standards-based ordinal measures of human rights with an emphasis on aggregation into single indices. Such measures maintain a reasonably high level of abstraction suitable for large cross-national comparisons and in part have been driven by that particular research community, but they have problems of validity, reliability, variance truncation, and in many cases follow no obvious aggregation rule. Such measures need to be improved by a greater attention to primary sources in an effort to increase their validity, and greater disaggregation into separate measures of particular human rights. If standards-based ordinal scales are to be used and greater use is made of primary source material then such measures should provide more gradation in their ordinal categories in order to reduce the worst forms of variance truncation. It seems paramount, however, that such an effort needs to be complemented by other forms of data, including events-based and survey-based, and indicators of government policies and outcomes.

Suggestions for further reading

- Adcock, R. and Collier, D. (2001) 'Measurement Validity: A Shared Standard for Qualitative and Quantitative Research', *American Political Science Review*, 95(3):529–546.
- Cingranelli, D. (1988) *Human Rights: Theory and Measurement*, Basingstoke: Macmillan.
- Green, M. (2001) 'What We Talk about When We Talk about Indicators: Current Approaches to Human Rights Measurement', *Human Rights Quarterly*, 23:1062–1097.
- Jabine, T.B. and Claude, R.P. (eds) (1992) *Human Rights and Statistics: Getting the Record Straight*, Philadelphia: University of Pennsylvania Press
- Landman, T. (2004) 'Measuring Human Rights: Principle, Practice, and Policy', *Human Rights Quarterly*, 26(November):906–931.

Online resources

- Cingranelli and Richards Human Rights Data (CIRI)
<http://www.humanrightsdata.com/>
- Freedom House
<http://www.freedomhouse.org/>
- Landman Human Rights Treaty and Reservations Data
<http://www.data-archive.ac.uk/>
- Penn World Tables
pwt.econ.upenn.edu
- Political Terror Scale
www.unca.edu/politicalscience/faculty-staff/gibney.html
- Polity IV Democracy
www.cidcm.umd.edu/inscr/polity
- World Bank World Development Indicators
<http://www.worldbank.org/>