



Project
MUSE[®]

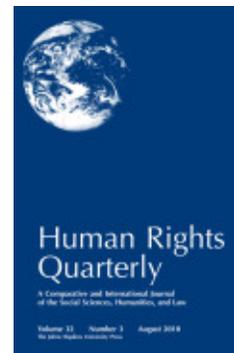
Today's Research. Tomorrow's Inspiration.

The Cingranelli and Richards (CIRI) Human Rights Data Project¹

David L. Cingranelli
David L. Richards

Human Rights Quarterly, Volume 32, Number 2, May 2010, pp.
401-424 (Article)

Published by The Johns Hopkins University Press
DOI: [10.1353/hrq.0.0141](https://doi.org/10.1353/hrq.0.0141)



 For additional information about this article
<http://muse.jhu.edu/journals/hrq/summary/v032/32.2.cingranelli.html>

The Cingranelli and Richards (CIRI) Human Rights Data Project¹

*David L. Cingranelli** & *David L. Richards***

ABSTRACT

The CIRI Human Rights Data Project provides information about government respect for a broad array of human rights in nearly every country in the world. Covering twenty-six years, fifteen separate human rights practices, and 195 countries, it is one of the largest human rights data sets in the world. This essay provides an overview of the CIRI project and our response to some critiques of the CIRI physical integrity rights index. Compared to the Political Terror Scale (PTS), the CIRI physical integrity rights index is focused on government human rights *practices*, can be disaggregated, is more transparent in its construction, and is more replicable because of the transparency of our coding rules. Furthermore, unlike the PTS, the unidimensionality of the CIRI index has been demonstrated empirically. For these reasons, the CIRI index is a more valid index of physical integrity rights.

* *David L. Cingranelli* is Professor of Political Science at Binghamton University, State University of New York. He is the former President of the Human Rights Section of the American Political Science Association. He conducts global, comparative, econometric research examining the causes and consequences of variation in government respect for various types of human rights.

** *David L. Richards* is an Associate Professor of Political Science at the University of Memphis. He is the co-founder and co-director of the Cingranelli-Richards (CIRI) Human Rights Data Project. His research topics have included issues in the measurement of human rights practices, the effect of globalization on respect for human rights, and how the end of the cold war has affected human rights practices. He is currently conducting research on global patterns of violence against women.

1. Besides Cingranelli and Richards, major contributors to the data collection effort include M. Rodwan Abouharb, Chad Clay, Paola Fajardo, Shawna Sweeney, and Craig Webster.

I. INTRODUCTION

The Cingranelli and Richards (CIRI) Human Rights Data Project provides standards-based measures of government human rights practices, using much of the broad range of human rights recognized in the Universal Declaration of Human Rights. The broad scope of human rights represented by our measures is an important feature, as one of the main purposes of the CIRI project was to expand theory building and empirical research about government human rights practices beyond the extant dominant concern with violations of physical integrity rights.

The current focus within political science on physical integrity rights is ironic because one of the earliest attempts to measure the human rights practices of all countries of the world was quite comprehensive. Between 1983 and 1992, Charles Humana published three editions of his *World Human Rights Guide*.² Humana's guides were systematic attempts to track changes in government respect for a wide variety of human rights in a variety of countries over time. The countries and rights included in Humana's guides varied from edition to edition, but it was always broadly comparative. Humana divided his list of human rights into positive rights (freedom to), negative rights (freedom from), legal rights, and personal rights. He scored individually each particular right for each country and assigned a "grade" for each right in each country ranging from full respect to none. Humana suffered many criticisms of his pioneering project; he was criticized for weighting physical integrity rights more heavily than other human rights, for putting too much emphasis on civil and political rights, and for using too many sources of information when making his coding decisions. Still, we admire his effort to measure many human rights, thus providing a fuller picture of the human rights practices of governments around the world. In pursuit of that goal, we follow in his footsteps.

In the first part of this essay, we provide an overview of the CIRI project, including the types of human rights we benchmark, the measurement procedures we use, and our responses to some critiques addressed towards both the CIRI project itself and human rights measurement in general. In the second part, we describe our five-indicator approach to measuring physical integrity rights (the dominant type of human rights investigated by political scientists) and draw comparisons and contrasts with the Political Terror Scale (PTS), a longstanding physical integrity rights indicator. We welcome the opportunity to write this essay and hope that it will serve both as an introduction to our data project for potential new users and as a resource for those who already use the CIRI human rights data.

2. CHARLES HUMANA, *WORLD HUMAN RIGHTS GUIDE* (1st ed. 1983); CHARLES HUMANA, *WORLD HUMAN RIGHTS GUIDE* (2d ed. 1986); CHARLES HUMANA, *WORLD HUMAN RIGHTS GUIDE* (3d ed. 1992).

II. OVERVIEW OF THE CIRI PROJECT

The CIRI Human Rights Project provides policymakers, researchers, teachers, and students with easily accessible, reliable, valid, annually-updated information about government respect for a broad array of human rights in nearly every country in the world. Covering twenty-six years (1981–2007 at the time of writing), fifteen separate human rights practices (plus two indices), and 195 countries, it is one of the largest human rights data sets in the world.³ We engage in this endeavor because we believe that measuring human rights practices helps to understand the determinants and consequences of government respect for human rights and therefore, that measurement is an important part in the global human rights movement that attempts to provide dignity for all persons worldwide. The CIRI data are freely available on our website (www.humanrightsdata.org) where users can either download the entire dataset at once or create a custom dataset for download by choosing only those indicators, years, and countries they wish to use. Starting in 2010, the website will include downloadable color and black and white maps that can be used in presentations, in the classroom, or in publications.

A. Scope

CIRI provides measures of several types of internationally-recognized human rights, including:

1. *Physical integrity rights*: The rights not to be tortured, extrajudicially killed, disappeared, or imprisoned for political beliefs.
2. *Civil rights and liberties*: The rights to free speech, freedom of association and assembly, freedom of domestic movement, freedom of international movement, freedom of religion, and freedom to participate in free and fair elections for the selection of government leaders.
3. *Workers' rights*: The rights to freedom of association, collective bargaining, a minimum age for the employment of children, acceptable conditions of work, and protection from forced labor.
4. *Women's rights*: The rights to legal protection and equal treatment politically, economically, and socially.

The Appendix to this article includes short descriptions of those indicators, other than those that measure physical integrity rights, currently available for download from the CIRI website. In the future, we expect to add measures

3. The basic unit coded is a "country-year." A country-year is a particular country in a particular year. For instance, "Canada 1998" is a particular country-year.

of respect for other human rights, including the right to a fair trial, the right to an independent judiciary, and various economic and social rights.

Because they cover many different human rights, CIRI data allow for the exploration of a variety of important questions, such as what types of human rights are most and least respected by governments and why? How have patterns of respect for different types of human rights changed over time? To what extent is the distribution of foreign aid affected by the human rights practices of potential recipients? Is there a relationship between violations of human rights and rebellion? How have the spread of democracy and rapid economic globalization since the end of the Cold War affected human rights practices? How have specific policies such as trade liberalization, bilateral foreign aid, and structural adjustment conditions affected government human rights practices, including labor rights and women's economic rights? Is respect for some types of human rights necessary for equitable and rapid economic growth? And do human rights crises such as the current one in Sudan have measurable effects on the human rights practices of neighboring governments?

While we created the CIRI project in 1994 to provide data for the research projects of the project directors and others who conduct quantitative studies of government human rights practices, its data now are also widely used by governments, intergovernmental organizations, non-governmental organizations, think-tanks, and private businesses. Despite the use of its data in domestic and international policymaking, CIRI remains an independent, non-governmental organization. Our data and analyses are independent of governmental influence or the influence of any other external entity. In recent years, the CIRI project has received several grants from the National Science Foundation's Political Science Division. This financial support has allowed us to expand the data set, to update it annually, and to make it freely available to users. CIRI does not accept funds in the form of contracts from government institutions, including US government agencies, and never functions as an extension of any government. In fact, in October 2009, US Senator Tom Coburn (R- Nebraska) publicly criticized the CIRI Project for disseminating information and analyses that were critical of the US government's recent practices related to torture.⁴

In addition, some fear that governments and organizations, like the World Bank and the Millennium Challenge Corporation, who use human rights data in their evidence-based policymaking have allowed measurement methodologies to run ahead of clear conceptual thinking about what human rights mean and how they should be measured. For example, Kate

4. For further information, see Senator Coburn's statement, Coburn Amendment 2631—Prohibits the National Science Foundation from wasting federal research funding on political science projects, *available at* <http://tinyurl.com/coburn-ciri>.

Desormeau notes that the World Bank applies a “governance factor” to arrive at a country’s overall rating for aid eligibility.⁵ Of the six dimensions of governance, three—“voice and accountability,” “rule of law,” and “political stability”—incorporate indicators supplied by the CIRI Human Rights Data Project and the Freedom House project. She observes that these indicators emphasize civil rights (including physical integrity rights) and political liberties and finds the World Bank’s selective attention to some human rights and its omission of others troubling because, for Desormeau, *all* human rights included in the International Bill of Human Rights should be considered.⁶ That is, instead of further exploring and analyzing a country’s human rights record with respect to *all* human rights, aid providers seem to be using only available data to arrive at a country’s respective ranking.

On one level, we agree with this argument. Results and overall conclusions would be more legitimate if aid providers considered government efforts to protect *all* human rights—especially economic and social rights—when making aid decisions. On another level, we disagree. Desormeau implies that unless *all* or nearly all human rights practices are considered when decisions are made about the distribution of economic aid, then *no* human rights practices should be considered. This seems to us a classic case of “making the perfect the enemy of the good.” We applaud the use of human rights criteria in government decision making and hope this tendency takes root and grows.⁷ Over time there will be more and better measures of a greater variety of human rights, and, hopefully, aid providers will incorporate these in ever more sophisticated ways into the policymaking process.

B. Producing the CIRI Data

We use a mixed-methods approach to create our indicators of government respect for various human rights, employing content analysis of qualitative material describing respect for human rights in countries around the world to create quantitative indicators. Our coding criteria were developed both to reflect the meanings of various human rights as defined in international human rights law and to represent the myriad ways in which the expectations of human rights law and actual government behavior intersect. The CIRI human rights scores indicators are considered “standards-based” because

5. Kate Desormeau, *Measuring Rights and Governance: The Quantification of Aid Conditionality 5* (Harvard Kennedy School, Carr Center for Human Rights Working Paper), available at http://www.hks.harvard.edu/cchrp/pdf/Desormeau_Paper.pdf.

6. *See id.* at 7.

7. For an elaboration of this argument, see M. RODWAN ABOUHARB & DAVID CINGRANELLI, *HUMAN RIGHTS AND STRUCTURAL ADJUSTMENT* 227–39 (2008).

we rate actual government practices relative to standards set in international law. We do not rank countries relative to one another.

Measuring government respect for human rights for nearly every country in the world requires systematic qualitative information—meaning standardized information about the same rights for each country, annually. The annual *US State Department Country Reports on Human Rights Practices (Country Reports)* is the only such existing source, and our coders use it to code all variables. When coding physical integrity rights, coders also use Amnesty International's *Annual Report*. When there is a difference between the two sources, our coders treat the Amnesty International assessment as authoritative. Most scholars believe that this step, crosschecking the *Country Reports* assessment against the Amnesty International assessment, is necessary to remove a potential bias in favor of US allies. However, since 1981, there has been substantial and increasing agreement between the two reports.⁸ We crosscheck only physical integrity rights because it is the only type of rights Amnesty International systematically addresses in its reports.

CIRI's indicators are ordinal, meaning that they give information in terms of "more or less" but not precisely how much more or less. For example, our torture measure would say that torture in Country A is "frequent," but it would not say "317 persons were tortured." A methodological reason underpins this decision. We use the ordinal level of measurement because measurement schemes should complement the nature of the source information. Human rights information is far from perfect.⁹ Observers differ in explaining what they witness; witnesses have better or worse memories; witnesses can be intimidated, corrupted, or even killed; victims can be intimidated and, in many cases pertaining to physical integrity rights, killed; the only witnesses to a violation might be the perpetrators and victims (who might be dead); governments have incentives to lie; and violations may be taking place in an ongoing, long-standing conflict where observers cannot properly work. Thus, accurate numbers of violations are rare, and all counts contain an inherent amount of measurement error because of the aforementioned issues and others. We use the latitude provided by ordinal score categories to allow for measurement error resulting from the qualitative information on which our scores are based.

-
8. Steven C. Poe, Sabine C. Carey & Tanya C. Vazquez, *How Are These Pictures Different? A Quantitative Comparison of the US State Department and Amnesty International Human Rights Reports, 1976–1995*, 23 *HUM. RTS. Q.* 650 (2001).
 9. See Michael Stohl, David Carleton, George Lopez & Stephen Samuels, *State Violation of Human Rights: Issues and Problems of Measurement*, 8 *HUM. RTS. Q.* 592 (1986); Herbert F. Spirer, *Violations of Human Rights—How Many? The Statistical Problems of Measuring Such Infractions Are Tough, but Statistical Science is Equal to It*, 49 *AM. J. ECON. & SOC.* 199 (1990); Alison Brysk, *The Politics of Measurement: The Contested Count of the Disappeared in Argentina*, 16 *HUM. RTS. Q.* 676 (1994).

There are some fine event-count human rights data projects, but because of the labor intensiveness of that type of endeavor, the data they produce are typically for a single country over a finite period of time, usually in the past. While these data sets are crucial for work done by entities such as the International Criminal Tribunal for the former Yugoslavia, the International Criminal Tribunal for Rwanda, and the International Criminal Court, scholars and policymakers also demand annually-updated data for each country. Like many others, we too wish that reliable count-data were available annually for all countries. However, at this time, given the quality of existing systematic source information on human rights, only ordinal measurement schemes are appropriate to create data to meet these demands. Finally, some concepts such as “freedom of movement,” “women’s economic rights,” or “electoral self-determination” do not lend themselves to counts. For these, ordinal measures are naturally appropriate.

In this volume, Wood and Gibney critique what they say is CIRI’s use of count-based categories to derive our ordinal physical integrity rights scores.¹⁰ Their critique is based dominantly on information in the Cingranelli and Richards 1999 article, “Measuring the Level, Pattern, and Sequence of Government Respect for Physical Integrity Rights.”¹¹ There, we explained that for the four CIRI physical integrity rights indicators, zero violations would dictate a score of “2” (full respect), 1-to-49 violations would dictate a score of “1” (moderate respect), and 50-or-more violations would dictate a score of “0” (no respect). This scheme was implemented to try to increase intercoder reliability in an era where the source information (the same as used by the PTS) was less systematic than it is at present. They argue that these numeric thresholds are arbitrary. That is true—as it is true for any numeric threshold, such as the 1,000 battle deaths threshold of war used by the Correlates of War Project, which provides widely-used data that have transformed peace studies research.¹² We chose numeric thresholds that we felt, after a substantial reading of years worth of reports, represented the three types of violators that could be reliably ascertained from the source material available to us.

In practice, these numeric thresholds are rarely used to produce scores because usable numeric estimates are seldom provided in our qualitative source material. Thus, most of our physical integrity scores are based on qualitative descriptions, just like the indicators that are not based on countable phenomena. For example, CIRI coders are instructed that “[i]nstances

-
10. Reed M. Wood & Mark Gibney, *The Political Terror Scale (PTS): A Re-introduction and a Comparison to CIRI*, 32 HUM. RTS. Q. 367, 378–80 (2010).
 11. David L. Cingranelli & David L. Richards, *Measuring the Level, Pattern, and Sequence of Government Respect for Physical Integrity Rights*, 43 INT’L STUD. Q. 407 (1999).
 12. Correlates of War, available at <http://www.correlatesofwar.org>.

where violations are described by adjectives such as 'gross,' 'widespread,' 'systematic,' 'epidemic,' 'extensive,' 'wholesale,' 'routine,' 'regularly,' or likewise, are to be coded as a ZERO (have occurred frequently)."¹³ We take great care in putting together these lists of key terms, going so far as to regularly use concordance software to analyze the language of the *Country Reports* so that simple changes in language do not affect scores. When doing content analysis on an annual multi-authored source over time, one needs to ensure that scores are based on practices, not on changes in the way similar things are described from year to year.

This brings us to the issue of data truncation, whereby large amounts of variation tend to exist in a measure's "top" category. All ordinal scales suffer from this issue, to some extent. Wood and Gibney criticize CIRI for having a scale where a country that tortured fifty-one persons would get the same score as a country that tortured 3,001 persons. That is true of CIRI and it is true of the PTS as well. It is very difficult to discern, for example, the difference between a score of 4 and a score of 5 on the PTS, as coding rules such as "[t]he practices of Level 3 are expanded to larger numbers" (for a score of 4) and "[t]he terrors of Level 4 have been extended to the whole population" (for a score of 5) seem more than slightly subjective. What range of violations counts as "larger numbers"? Do they mean "the whole population" in a literal sense? Finally, even were the PTS coding guidelines more detailed, we believe, after much analysis of State Department reports with concordance software, that the available sources of systematic information are neither rich enough, nor reliable enough in type of content offered (e.g., qualitative descriptions of abusive episodes, numerical estimates of victims, etc.), nor consistent enough in the way that similar episodes are described either within the same country or across countries, to support reliable ordinal measures with more than three or four categories.

Wood and Gibney also criticize the CIRI measurement procedure, because even states with the most exemplary human rights records (e.g., New Zealand, Denmark, and Canada) have been accused by CIRI of physical integrity violations and have not received perfect or even near-perfect CIRI index scores. Thus, they brand our scheme as inflexible. However, we object to this characterization because, as noted, CIRI rates countries according to an absolute standard of behavior and does not rank them relative to one another. The fact that these three countries are typically considered among the "best" does not sanitize their existing transgressions for purposes of assigning a score on a human rights measure. For example, based on Amnesty International's criticism of the United States for police and prison guard

13. DAVID L. CINGRANELLI & DAVID L. RICHARDS, THE CINGRANELLI-RICHARDS (CIRI) HUMAN RIGHTS DATA PROJECT CODING MANUAL, version 7.30.08 (2008), at 8, available at http://ciri.binghamton.edu/documentation/ciri_coding_guide.pdf.

brutality—including prison rapes (torture); the imprisonment of suspected terrorists without access to lawyers or the right to a trial (political imprisonment); and the continued use of tasers by police, leading to numerous deaths in instances where lethal force is not necessary (extrajudicial killing)—the United States received low CIRI scores on these rights.

For purposes of quality control, each CIRI data point (e.g., torture in Colombia in 2007) is coded by at least two trained coders. Coders are trained by senior project staff on the measurement scheme for the particular variable on which they will be working and do “practice coding” from sample packets until they code well enough to start coding the countries to which they have been assigned. When both coders are finished with their assigned countries, they meet with senior personnel to resolve any disagreements. Most times, disagreements are based on a coder having missed a piece of information. Sometimes, indeed, a disagreement arises about how to apply a particular coding rule or how to interpret a particular piece of information. In these cases, both of the coders and the senior personnel reread and discuss the original material together to arrive at an agreeable score. If no agreement exists, the senior personnel member decides the score. As part of our quality-control process, even when both coders agree on a score, it is checked against past years, and if there is a change, the original source material is revisited to make sure the change is due to government practice, not mutual coder error. Coders are strictly prohibited from looking at past scores until this stage in the process.

The CIRI coding schemes are very detailed so as to maximize the reliability of the data. Each indicator’s entry in the coding guide includes a full definition of the concept being measured, the measurement scheme, detailed instructions on how to transform the qualitative source information into numeric scores, directions on where to find information for that particular indicator, and country-examples at each scoring point drawn directly from source materials. The CIRI coding guide is well over one hundred single-spaced pages in length.

A reliable indicator is one whereupon different people apply an identical measurement procedure to identical information and come up with identical scores. The fewer the disagreements between coders, the greater the reliability of a measure. In recent years, CIRI has kept systematic records of coder disagreement, and this allows us to assess the level of reliability of our data. For the coding of human rights violations in the year 2004, for example, Krippendorff’s r -bar intercoder reliability statistic, which ranges from 0 (no reliability) to 1.0 (perfect reliability), was .94, indicating very high intercoder reliability.¹⁴ Aside from identifying bad coders (whose work

14. Krippendorff’s measure is appropriate for ordinal data because it takes into account both the number of disagreements and the extent of each disagreement.

is ultimately redone by someone else), assessing reliability in this manner is very important for knowing when improvements must be made in the coding guidelines. This does not necessarily mean that coding rules must be changed. Rather, in most cases, the existing rules need to be better explained or better examples need to be provided. The detailed coding schemes used by coders are available in the "Documentation" section of the CIRI website.¹⁵

It is important to understand that measurement choices can affect reliability. For example, reliability will be lower for indicators with more score-points than for those with less. Also, the introduction of exogenous elements (those not part of the qualitative information source being content-coded) can affect reliability. For example, we believe that the PTS' directive to coders to take the population of a country into account when assigning scores to countries reduces intercoder reliability.¹⁶ Unfortunately, neither the PTS nor the Freedom House projects publicly report reliability statistics.

III. MEASURING PHYSICAL INTEGRITY

In this section, we compare and contrast the CIRI measures of government respect for physical integrity rights with PTS. As noted, physical integrity rights are the human rights individuals have to be free from arbitrary physical harm and coercion by their government. The principal rights in this category are the rights to not be subjected to torture, political imprisonment, extrajudicial killing, and disappearance.

In the CIRI index, torture is defined as the purposeful infliction of extreme pain, whether mental or physical, by government officials or by private individuals at the instigation of government officials. Torture includes the use of physical and other force by police and prison guards that is cruel, inhuman, or degrading. Political imprisonment refers to the incarceration of people by government officials because of their ideas, including religious beliefs; their non-violent religious practices, including proselytizing; their speech; their non-violent opposition to government policies or leaders; or their membership in a group, including an ethnic or racial group.¹⁷ Extrajudicial killings are killings by government officials without due process of law. They include murders by private groups if the private actors are instigated by government.¹⁸ Disappearances refer to unresolved cases in which political

15. David Cingranelli & David Richards, CIRI Human Rights Documentation, *available at* <http://ciri.binghamton.edu/documentation.asp>.

16. Wood & Gibney, *supra* note 10, at 387, 392.

17. Individuals who are imprisoned because they have committed violent acts, regardless of the reasons why they committed those acts, are not political prisoners.

18. Extrajudicial killings may result from the deliberate, illegal, and excessive use of lethal force by the police, security forces, or other agents of the state whether against criminal suspects, detainees, prisoners, or others. Extrajudicial killing excludes combat deaths.

motivation appears likely and in which the victims have not been found. Disappearances and extrajudicial killings are closely related practices. Many victims of human rights abuse who initially are categorized as having been disappeared are later found to have been killed.

The PTS is a five category measure of the level of “terror” (or physical integrity abuse) in a given country. The scale ranges from level 1 (no terror) to level 5 (widespread terror throughout the entire population). For two decades, the PTS has been a widely used measure of physical integrity rights in social science research. However, when we created the CIRI measures of government respect for physical integrity rights, we decided to create a measurement scheme that focused on government human rights practices rather than on overall human rights conditions. We also preferred to create several disaggregated measures of specific types of physical integrity violations. Finally, we wanted to develop a measurement procedure that was more transparent and, therefore, more replicable.

Below we outline some important differences between the CIRI physical integrity measurement scheme, which includes five indicators (one for each component human right plus the CIRI physical integrity index) and the PTS indicator.

A. Practices versus Conditions

The human rights scores reported by the CIRI project represent the human rights practices of governments. Human rights practices refer to the *actions of government officials* and actions by private groups *if instigated by government* directly affecting the degree to which citizens can exercise various types of human rights. In other words, practices are what governments actually do, not what they claim to do or what consequences flow from what they do. Almost all of the scientific research in political science focusing on human rights as a concept is really concerned about the causes or consequences of variations in what governments decide to do to protect or violate the human rights of their citizens.

The PTS measures human rights *conditions*, which refer to the degree to which citizens can exercise various types of human rights, as affected both by governmental and other actors. Nongovernmental groups (NGOs) such as revolutionaries, gangsters, or terrorists also may violate human rights independently from the government and may worsen the human rights conditions in a country. For example, in their 1992 reports, both Amnesty International and the US State Department estimated that more than half of the civilian deaths in Peru were caused by the Shining Path revolutionary group, not the government. Thus, theories constructed to explain the causes of variations in government human rights *practices* would be different from

theories constructed to explain the causes of variations in human rights *conditions*.

In many cases of domestic turmoil (not necessarily rising to the status of civil war), we would expect the PTS and the CIRI physical integrity rights index to differ because of the difference between measuring government human rights practices and measuring overall human rights conditions. For countries involved in insurgencies, overall human rights conditions would tend to be worse than government human rights practices. For the last few years, for example, the PTS has scored Afghanistan and Iraq at level 5 (maximum terror) while the CIRI scores, which only considered the practices of the Afghani and Iraqi governments, indicated less use of state-based terror.

Wood and Gibney dispute our claim that the PTS actually measures human rights conditions. However, we are skeptical for a few reasons. First, the PTS project has explicitly stated that it measures conditions. Gibney and Dalton have stated explicitly that coders who collect information used to construct the PTS are told to “[t]ry to measure government terror, but ultimately be sensitive to all forms of terror.”¹⁹ Coders are explicitly instructed *not* to ignore terror from nongovernmental actors. The aim of the PTS is “to reflect the human rights violations that exist in a country more generally.”²⁰ Also, the *Human Security Report* uses PTS data and bases its information about that measure on a “background document commissioned by the Human Security Centre” written by Linda Cornett and Mark Gibney, called “Tracking Terror: The Political Terror Scale 1980–2001.”²¹ The *Human Security Report* notes that “[t]he Political Terror Scale counts human rights abuses by any group—government or nongovernment. Most abuse, however, is by governments.”²²

The second reason we are skeptical about the claim that the PTS measures government human rights practices is that the PTS project produces scores for all countries of the world for every year—even for those countries and years in which no national government existed. How can one measure government human rights practices if there are no governments? In contrast, the CIRI project does not produce human rights practice scores for any country for any year during which there was a complete collapse of central political authority or a foreign occupation. The PTS project produced scores for the Democratic Republic of the Congo between 1996 and 2000, for Sierra Leone between 1997 and 2000, and for Somalia between 1996 and 2007, all years

19. Mark Gibney & Matthew Dalton, *The Political Terror Scale*, in POLICY STUDIES AND DEVELOPING NATIONS: HUMAN RIGHTS AND DEVELOPING COUNTRIES 73, 79 (David L. Cingranelli ed., 1996).

20. *Id.*

21. HUMAN SECURITY CENTRE, HUMAN SECURITY REPORT 2005: WAR AND PEACE IN THE 21ST CENTURY 95 n.30 (2005).

22. See *id.* n.31.

when these three countries suffered a complete collapse of central authority. The following are examples of the PTS producing scores for countries under foreign occupation. The PTS produced scores for Afghanistan in 2003 and 2004 when the former Taliban government was no longer in power, no new Afghan government had been installed, and government practices were under the control of the United States. Similarly, the PTS project produced a score for Iraq in 2003 and 2004 when no Iraqi government existed and for Lebanon in 1996–2000 when it, too, was under foreign rule. The CIRI project did not report human rights scores for any of those countries in any of those years. Instead, we assigned those country-years one of the special codes developed by the Polity data project.²³ Several studies report results using both the CIRI and PTS indices. The results using the PTS scores always include more observations, but those extra observations result from the PTS project's inclusion of countries with no governments. Because of the different foci of the two projects, using the PTS and CIRI data as mutual robustness checks in empirical research projects is unnecessary and invalid.

B. Level of Aggregation

The PTS is most comparable to CIRI's physical integrity rights index, which is an additive index created from four individual indicators (torture, extrajudicial killing, political imprisonment, disappearance) and ranging from 0 (no respect for any of these four rights) to 8 (full respect for these four rights). Because the CIRI index is created by adding the individual scores for four different human rights practices, anyone interested in what combination of practices produced a particular score for a particular country-year can find out by examining the scores for each of the individual indicators. This option allows scholars and policy makers to gain useful information about respect for physical integrity rights both individually and in combination with one another. It also allows scholars to conduct separate analyses of each of the component human rights practices.²⁴

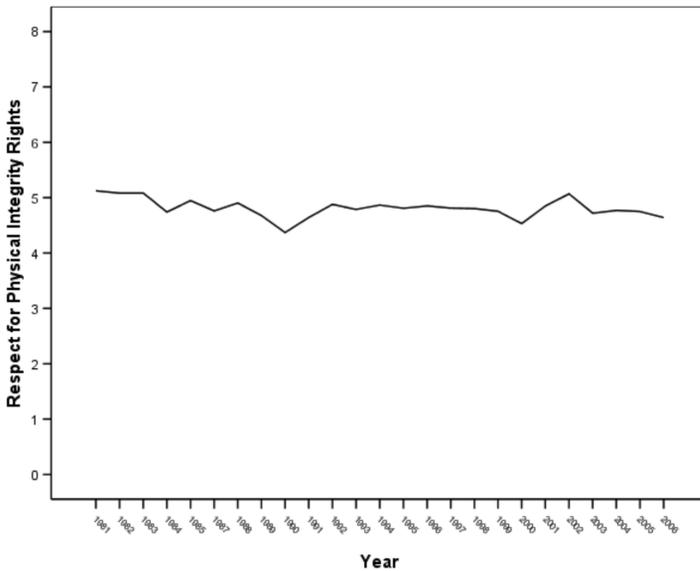
The PTS, however, cannot be disaggregated. This means that an individual cannot deconstruct it to see only the information for a single physical integrity right, for example, torture or disappearance. Indeed, from the PTS' coding scheme, disappearances are not even mentioned until level 4,

23. MONTY G. MARSHALL & KEITH JAGGERS, POLITY IV PROJECT, POLITICAL REGIME CHARACTERISTICS AND TRANSITIONS, 1800–2007 DATASET USERS' MANUAL (2009).

24. See, e.g., James Raymond Vreeland, *Political Institutions and Human Rights: Why Dictatorships enter into the United Nations Convention Against Torture*, 62 INT'L ORG. 65 (2008); David Cingranelli & Mikhail Filippov, *Electoral Rules and Incentives to Protect Human Rights*, 72 J. POLITICS 243 (2010).

implying that the mention of any disappearances leads the PTS staff to code the country at level 4 or 5.

While there are several arguments on behalf of creating disaggregated (or disaggregable) data (like CIRI) rather than completely aggregated indices that cannot be deconstructed (like PTS), we feel that the most compelling reason is that an aggregated index can hide variation in government respect for human rights. For example, the trend line in Figure 1 plots the average CIRI physical integrity index score for 162 countries from 1981 to 2006. Plotting the average PTS scores over time produces a similarly flat line and yields the same story—little change in government physical integrity rights practices over time.



Note: A score of 8 indicates full respect for all four rights from which the index is composed. A score of 0 indicates no respect for any of these four rights.

Figure 1. Aggregated Global Average Level of Government Respect for Physical Integrity Rights, 1981–2006

However, this is an illusion—there is actually considerable variation in these practices. One must disaggregate. Figure 2 plots the trends in average respect for the components of the CIRI physical integrity index for the same 162 countries between 1981 and 2006. Figure 2 shows that there has been much variation in government respect for human rights over the twenty-five-

year period and that aggregated indices (both PTS and CIRI) mask interesting and significant changes. The most important findings displayed in Figure 2 are the steady decline in respect for the right not to be tortured and the steady improvement in the amount of government respect for the right not to be imprisoned for political reasons. The decline in government respect for the right not to be tortured continued steadily across the time period and showed the greatest change from one era to the next. The dramatic increase in average respect for the right not to be imprisoned for political reasons is also statistically significant.²⁵ Respect for the other two physical integrity rights—the right not to be disappeared and the right not to be extrajudicially killed—changed only a little throughout the period and moved in opposite directions. Changes in average respect for the right against disappearance and for the right against extrajudicial killing were both small but statistically significant.²⁶ Throughout the entire time period, government respect was highest for the right not to be disappeared. Because the CIRI data allow us to examine changes in scores for torture, disappearances, political imprisonment, and extra-judicial killing separately, we can determine why the average scores of the aggregated indices are relatively flat. For example, while respect for one physical integrity right has been dramatically increasing (political imprisonment), respect for another has been decreasing (torture). These two changes effectively cancel each other out in the PTS and the aggregated CIRI indices and give the illusion that respect for both rights stayed constant throughout the time period.

Wood and Gibney's preoccupation with aggregated indices, rather than the disaggregated indicators we prefer, provide fuel for many of their critiques. For example, they provide the following illustration:

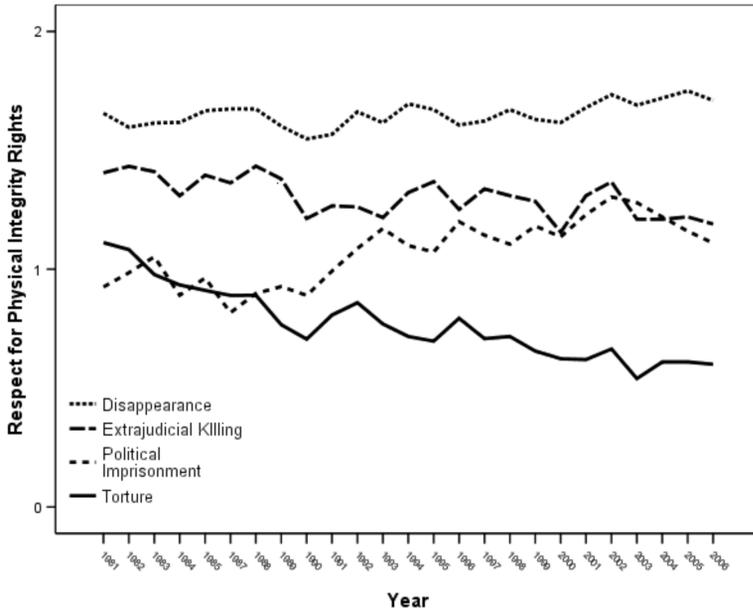
Imagine that in Country A, security officials storm a labor rally and kill 100 labor union members. In Country B, however, 100 labor union members are arrested and imprisoned, tortured, and then killed. According to the approach of the PTS, the level of political violence in these two countries would essentially be the same. However, according to our understanding of the CIRI index, the human rights situation in the second state would be considerably worse than the first state (at least its score would make it appear to be much worse) because each violation would be coded separately.²⁷

We do not know what score either of these hypothetical countries would receive on the PTS because the PTS uses an unknown sliding scale based on population. Perhaps the PTS measurement scheme would assign

25. David L. Cingranelli & David L. Richards, Not Everything's Coming up Roses: Human Rights After the End of the Cold War, Paper presented at the 2006 Annual Meeting of the International Studies Association, San Diego, CA.

26. *Id.*

27. Wood & Gibney, *supra* note 10, at 377.



Note: A score of 2 indicates full government respect for a right, whereas a score of 0 indicates no respect.

Figure 2. Disaggregated Global Average Level of Government Respect for Physical Integrity Rights, 1981–2006

a score of 1 (no terror) for China and of 5 (maximum terror) for Tuvalu for both hypothetical situations. Further, the phrase “essentially be the same” in the quote above implies to us that the PTS would not assign identical scores to the two hypothetical situations. If they would assign different scores, on what basis would they be different? What we do know is that in the CIRI disaggregated data, Country A would receive a 0 for killing. The PTS provides no other information here about the three other practices that, together with killing, comprise the aggregated CIRI index, so we have no idea what the overall index score would be. Country B would receive a 0 for political imprisonment, torture, and extrajudicial killing on the individual CIRI indicators. We do not know what the overall index score would be on our aggregated index, as we do not know what practices were with regard to disappearances that year in Country B. We view this transparency and certainty about how specific combinations of government practices would generate specific scores as a strength of the CIRI measurement procedure.

Some are skeptical of the “no change in human rights practices” stories implied when aggregated indices such as PTS and the CIRI physical integrity

rights index are plotted over time as in Figure 1.²⁸ Qualitative research, they say, suggests that substantial improvements in respect for human rights have occurred since the end of the cold war. They speculate that flat or negative trends have been presented by quantitative scholars mainly because coding practices and standards of evaluation of human rights practices have undergone subtle changes over time, making cross-time comparisons misleading. Another hypothesis is that there is more attention to human rights issues today than there was during the cold war period, so negative trends simply reflect better reporting. While we see evidence of more lengthy human rights reports produced by the US State Department in more recent years, we see little evidence that lengthier reports of human rights practices has led to more negative assessments of those practices. In fact, respect for most human rights included in the CIRI Human Rights Data Project has improved over time. For the thirteen human rights trends we code, when we compared the mean level of respect for each right for the cold war and post cold war periods, we found statistically significant improvements in more than half (seven) of them.²⁹

C. Forming Empirically Verified Human Rights Indices

Indices of different types of physical integrity rights (or other categories of rights) can be formed, but their dimensionality should be empirically justified. The CIRI physical integrity rights and empowerment indices are empirically-verified unidimensional scales.³⁰ To empirically demonstrate the strong unidimensionality of both aggregated CIRI indices, we use a polychotomous extension of a probabilistic cumulative scaling technique known as Mokken Scaling Analysis (MSA). The four components of our aggregated index routinely return an “H Statistic” around .60, indicating, according to Mokken’s rules, strong scalability (unidimensionality).³¹ On the other hand, the PTS assumes both unidimensionality and an *a priori*

28. See, e.g., Emilie M. Hafner-Burton & James Ron, *Seeing Double: Human Rights Impact Through Qualitative and Quantitative Eyes*, 61 *WORLD POL.* 360 (2009).

29. See Cingranelli & Richards, *Not Everything’s Coming up Roses*, *supra* note 25.

30. Cingranelli & Richards, *Measuring the Level, Pattern, and Sequence of Government Respect for Physical Integrity Rights*, *supra* note 11, at 407–08; David L. Richards, Ronald D. Gellensy & David H. Sacko, *Money with a Mean Streak? Foreign Economic Penetration and Government Respect for Human Rights in Developing Countries*, 45 *INT’L STUD. Q.* 219 (2001).

31. For details on Mokken Scaling, see R.J. MOKKEN, *A THEORY AND PROCEDURE OF SCALE ANALYSIS* (1971); Wijbrandt H. van Schuur, *Mokken Scale Analysis: Between the Guttman Scale and Parametric Item Response Theory*, 11 *POL. ANALYSIS* 139 (2003). For an analysis of the CIRI data using this technique, see Cingranelli & Richards, *Measuring the Level, Pattern, and Sequence of Government Respect for Physical Integrity Rights*, *supra* note 11.

ordering among types of physical integrity violations, yet neither has ever been empirically justified.

One related problem with the PTS is that the combinations of human rights violations in the middle categories do not begin to exhaust the logical possibilities. As a result, the human rights practices in many countries may not match the PTS categories in which they are placed.³² To the extent that countries do not fit these patterns, the PTS is subject to measurement error. In contrast, because it is a simple additive index created from disaggregated components, the CIRI physical integrity index makes no *a priori* assumptions of patterns of government practices for purposes of its construction. Moreover, in previous work, we provide evidence using MSA to analyze the disaggregated CIRI indicators that there is indeed a discernible sequence of government respect for particular physical integrity rights across space and time.³³ That is, governments tend to violate some rights first and others later. In that same article, we outline the theoretical pattern of violations asserted by the PTS and use the MSA-based patterns to demonstrate that the pattern asserted by PTS is not empirically supportable.³⁴

As previously stated, no mention is given in the PTS scheme to disappearance until level four. While we are left to believe this is because PTS assumes that this type of rights violation does not occur until then, we are unsure. At any rate, this amounts to an *a priori* theoretical underpinning of the PTS that is demonstrably untrue in practice. For example, Argentina gets a 3 on the PTS for 2006 but experienced some disappearances late in the year. In that same year, Azerbaijan scored a 2 on the PTS but also experienced disappearances.

D. Transparency and Validity

CIRI scores are more transparent than PTS scores, and we believe that this transparency is necessary to assess validity. One of the most difficult and important issues in social science measurement is validity. To be valid, a measure should measure the concept it claims to measure. One of the difficulties of this in social science measurement is that we typically measure latent, or unobservable, concepts. Thus, social scientists are in the business of building proxy measures. The better the proxy represents the latent concept, the less the measurement error. Thus, a highly abstract concept will likely

32. See also James M. McCormick & Neil J. Mitchell, *Human Rights Violations, Umbrella Concepts, and Empirical Analysis*, 49 *WORLD POL.* 510 (1997).

33. See Cingranelli & Richards, *Measuring the Level, Pattern, and Sequence of Government Respect for Physical Integrity Rights*, *supra* note 11, at 413.

34. *Id.* at 414–15.

produce more measurement error. However, social scientists still measure abstract concepts such as “democracy,” “corruption,” “power,” “class,” and “human development.”

Therefore, measures of abstract concepts are controversial, and measurement specialists have a special duty to provide clear definitions of the concepts they are measuring. Even the existence of detailed definitions of various types of human rights violations contained in specialized human rights agreements such as the Convention against Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment and the Convention on the Prevention and Punishment of the Crime of Genocide do not create universal agreement as to whether a particular practice or incident in a society constitutes an act of torture or an instance of genocide. For example, whether the practice of water boarding (simulated drowning) constitutes torture as understood by international and US law has been the source of much debate since 2004.

The world brings many conceptual challenges, and a measurement specialist must explicitly address these and document his or her choices as part of a clear, conceptual definition. For example, do deaths from inappropriate police use of tasers count as extrajudicial killings? Does military hazing count as torture? Do voluntary restraints on clergy as part of tax-free arrangements qualify as restrictions on freedom of religion? What type or length of arrest or incarceration counts as political imprisonment? Is being put in jail for witchcraft political imprisonment or a violation of freedom of religion? Both? Neither? To answer these questions and others like them, the CIRI project provides detailed information about what included concepts like “torture,” “political imprisonment,” and “freedom of movement” mean according to the CIRI Data Project. Thus, users of our data who have a question about a particular score can, using our online coding guide and the publicly available source material, code the cases themselves and see how it works out.

In contrast, in this volume, Wood and Gibney tout the PTS coding practices as more flexible than the CIRI Data Project practices.³⁵ We read “flexible” to mean “subjective.” The fact that there exist several slightly different versions of the PTS scale in print, all containing many ellipses, and that there is no codebook publicly available containing coding rules and conceptual definitions is, to us, an unnecessary and important lack of transparency that negatively affects potential users’ ability to assess the reliability and validity of PTS data. Furthermore, their coding decisions are not obvious. In their article, they tell us that they count “violent expulsions” as

35. Wood & Gibney, *supra* note 10, at 378–79.

violations of physical integrity rights,³⁶ but there is no place in their brief coding categories that they consider coerced expulsions as violations of physical integrity rights. Their coding decision is reasonable, but the lack of transparency in their coding rules is not. Another problem is that, taking the PTS coding scheme at face value, violations seem to count towards a score even if they do not happen in the year being coded. The PTS notes that at level 3, “[t]here is extensive political imprisonment, or a recent history of such imprisonment.”³⁷

E. Population

When coding respect for physical integrity rights, CIRI coders are instructed to ignore the population of the country in question.³⁸ We think it is a mistake to take population into account when assigning human rights scores because, as noted above, doing so introduces too much subjectivity into the coding process. Taking population into account requires that the coders know the population size of each country. Then they must make a subjective decision about whether a certain number of incidents of torture, for example, indicated that the use of torture in that country in that year was “rare or exceptional” (PTS, level 1). They must also decide what number of violations indicated that the practice of torture had “been extended to the entire population” (PTS, level 5). Coders must make the same calculations for the three other physical integrity rights considered by the PTS and then figure out how to weight all four practices to produce a final PTS score.

This subjectivity reduces the reliability of the measure. This critique does not discount population as an important explanatory variable to be used in social science models. We encourage users of the CIRI physical integrity human rights scores to include population size as a control variable in equations designed to explain variation in these scores. The zero order correlation between population size and its CIRI physical integrity rights scores, whether disaggregated or in an index, will always be negative.

If we could, we would record the actual number of violations of each right by each government in each year. Then users of our data could divide the number of actual violations each year by the country’s population and get a

36. *Id.* at 391.

37. This language is not included in the version of the PTS coding scheme excerpted by Wood and Gibney in this volume, but it is included in the version of the PTS coding guidelines found in David Carleton & Michael Stohl, *The Foreign Policy of Human Rights: Rhetoric and Reality from Jimmy Carter to Ronald Reagan*, 7 *HUM. RTS. Q.* 205, 212 (1985) (emphasis added).

38. Note that all other scores for human rights practices in the human rights data set are not based on counts and, therefore, cannot be criticized on this basis.

score representing the number of violations per million people. Unfortunately, as we mentioned earlier, the source materials are not that specific about the number of violations that occur in each country in each year. Instead, we are only able to assign ordinal scores with any reliability or validity. This means that the governments of China and India almost always receive our lowest scores, in part because they have such large populations.

Taking population into account, the PTS is more generous to China and India than is CIRI, but on what basis? Exactly what is the sliding scale being employed? Are fifty reported instances of extrajudicial killing more acceptable in China or India because they have billion-plus populations? Does the United States, with a population of approximately 300 million, get a break? International law, according to any reading with which we are familiar, does not give large countries special dispensation on violations because of their population size. No matter a country's population, the standard of judgment is that *no one* is supposed to have his or her physical integrity violated. To the extent that the PTS does not use law as its standard of behavior, it drifts from being a standards-based rating towards being a relative ranking of countries. Further, we must illuminate an inherent ethical dilemma in this method: using a sliding scale based on population means that some deaths count more than do others. This is a violation of the basic principle of dignity that each life is of equal worth.

IV. CONCLUSION

We appreciate the opportunity to introduce the CIRI Human Rights Data Project and explain our mission, including explaining how we measure government respect for physical integrity rights and clarifying why we take the measurement approach we have chosen. We regret that so much of our discussion here has, by necessity, been focused on a defense against the Wood and Gibney critique of our aggregate physical integrity index. We prefer to deemphasize our aggregated physical integrity rights index because we want to encourage users of the data to look more carefully at each of the component human rights. Also, we would have liked to spend more time highlighting the other variables measured in the CIRI Human Rights Data Project. Non-physical integrity rights are currently sorely understudied.

Figure 2 above showed that governments around the world are, on average, reducing their respect for the rights of their citizens not to be tortured but are increasing their respect for the rights of citizens to be protected from political imprisonment. We will never understand why this is if we continue to focus so much attention on explaining variation in aggregate indicators of physical integrity rights like the PTS or the CIRI aggregated index because that variation is hidden by the index scores. In general, we agree with

James McCormick and Neil Mitchell that the focus of the political science discipline on highly aggregated indices, while ignoring the human rights practices that compose them, is hindering the development of theories that might explain why governments decide to violate some physical integrity rights rather than others.³⁹

Aside from paying attention to the issue of level of aggregation, we have argued that our index of physical integrity rights measures government human rights practices while the PTS mixes practices with conditions. Our measure, therefore, is more useful for those who wish to explain why governments choose to respect or violate physical integrity rights. We have presented several reasons why our measures are more transparent and reliable than the PTS. Transparency and reliability are not proof of validity, but they are necessary conditions that must be met before anyone can claim to have a valid measure of any abstract concept.

Furthermore, we hope that someday scholars and policymakers will pay as much attention to government respect to economic, social, and cultural human rights as they have paid to respect for physical integrity rights. Despite the recognition of other types of human rights in international human rights law, until recently, most human rights international nongovernmental organizations (INGOs), including Amnesty International and Human Rights Watch, have focused their reports and activities almost exclusively on identifying and remedying government violations of the physical integrity of the person. Currently, there is a movement towards an integrated human rights approach that reflects a belief in the complementarity, universality, and indivisibility of all rights, and we hope that the CIRI Human Rights Data Project will be an important resource in that movement.

APPENDIX: SHORT DESCRIPTIONS OF CIRI HUMAN RIGHTS MEASURES OTHER THAN PHYSICAL INTEGRITY RIGHTS

Freedom of Speech: This variable indicates the extent to which freedoms of speech and press are affected by government censorship, including ownership of media outlets. Censorship is any form of restriction that is placed on freedom of the press, speech, or expression. Expression may be in the form of art or music.

Freedom of Religion: This variable indicates the extent to which the freedom of citizens to exercise and practice their religious beliefs and convert to other religious beliefs is subject to actual government restrictions.

39. McCormick & Mitchell, *Human Rights Violations, Umbrella Concepts, and Empirical Analysis*, *supra* note 32, at 524–25.

Citizens should be able to freely practice their religion and proselytize (attempt to convert) other citizens to their religion as long as such attempts are done in a non-coercive, peaceful manner.

Freedom of Domestic Movement: This variable indicates citizens' freedom to travel within their own country.

Freedom of Foreign Movement: This variable indicates citizens' freedom to leave and return to their country.

Freedom of Assembly and Association: It is an internationally recognized right of citizens to assemble freely and to associate with other persons in political parties, trade unions, cultural organizations, or other special interest groups. This variable indicates the extent to which the freedoms of assembly and association are subject to actual governmental limitations or restrictions (as opposed to strictly legal protections).

Electoral Self-Determination: This variable indicates to what extent citizens enjoy freedom of political choice and the legal right and ability in practice to change the laws and officials that govern them. This right is sometimes known as the right to self-determination. Its scale ranges from 0 (no respect) to 3 (full respect).

Independence of the Judiciary: This variable indicates the extent to which the judiciary is independent of control from other sources, such as another branch of the government or the military.

Workers' Rights: Workers should have freedom of association at their workplaces and the right to bargain collectively with their employers. This variable indicates the extent to which workers enjoy these and other internationally recognized rights at work, including a prohibition on the use of any form of forced or compulsory labor; a minimum age for the employment of children; and acceptable conditions of work with respect to minimum wages, hours of work, and occupational safety and health.

Women's Political Rights: Women's political rights include a number of internationally recognized rights. These rights include: the right to vote, the right to run for political office, the right to hold elected and appointed government positions, the right to join political parties, and the right to petition government officials.

Women's Economic Rights: Women's economic rights include a number of internationally recognized rights. These rights include: equal pay for equal work, free choice of profession or employment without the need to obtain a husband's or male relative's consent, the right to gainful employment without the need to obtain a husband's or male

relative's consent, equality in hiring and promotion practices, job security (maternity leave, unemployment benefits, no arbitrary firing or layoffs, etc.), non-discrimination by employers, the right to be free from sexual harassment in the workplace, the right to work at night, the right to work in occupations classified as dangerous, and the right to work in the military and the police force.

Women's Social Rights: Women's social rights include a number of internationally recognized rights. These rights include: the right to equal inheritance; the right to enter into marriage on a basis of equality with men; the right to travel abroad; the right to obtain a passport; the right to confer citizenship to children or a husband; the right to initiate a divorce; the right to own, acquire, manage, and retain property brought into marriage; the right to participate in social, cultural, and community activities; the right to an education; the freedom to choose a residence/domicile; freedom from female genital mutilation of children and of adults without their consent; and freedom from forced sterilization.

Physical Integrity Rights Index: This is an additive index constructed from the Torture, Extrajudicial Killing, Political Imprisonment, and Disappearance indicators. It ranges from 0 (no government respect for these four rights) to 8 (full government respect for these four rights).

Empowerment Rights Index: This is an additive index constructed from the Freedom of Movement, Freedom of Speech, Workers' Rights, Political Participation, and Freedom of Religion indicators. It ranges from 0 (no government respect for these five rights) to 10 (full government respect for these five rights).