2-23-2009

Evaluation of the NAS-ILAB Matrix for Monitoring International Labor Standards

Lawrence S. Root
University of Michigan

Ada A. Verloren
University of Michigan

Evaluation of the NAS-ILAB Matrix for Monitoring International Labor Standards

Abstract

[Excerpt] Labor practices have become a part of many international trade agreements. In response to growing interest in international labor standards, the Department of Labor's Bureau of International Labor Affairs (ILAB) engaged the National Research Council of the National Academy of Sciences (NAS) to recommend a method to monitor and evaluate labor conditions in a country. The focus was on five labor standards: (1) Freedom of Association and Collective Bargaining; (2) Forced or Compulsory Labor; (3) Child Labor; (4) Discrimination; and (5) Acceptable Conditions of Work. NAS developed an approach which included formulating indicators for each of the five labor standards, creating a database of sources of reliable empirical information for assessing these indicators, and suggesting a method for assessing and monitoring compliance that includes a matrix framework for graphically presenting assessments.

The University of Michigan, in consultation with ILAB, evaluated this approach by having 3-person expert panels apply it to assessing and monitoring compliance in three countries. The panel members independently assessed the indicators in terms of level of compliance with international standards ("some problems," "more extensive problems," or "severe problems") and direction of change ("steady state," "improving," or "worsening"). Each panel then convened to discuss their assessments and seek to resolve differences in those assessments.

Observation and analysis of this experience revealed challenges in applying this approach and potential modifications that could improve its operationalization. For most of the indicators, there was not agreement among the independent assessments by the panel members. Information was lacking or outdated for many of the indicators and, when present, was usually not definitive. Panelists raised a number of suggestions for changes in the indicators and matrix approach, more generally. WebMILS, the database created to support such assessments, was seen as a useful resource, although concerns were raised about limiting information sources to those that are available on the Internet. There were also suggestions for updating and modifying the version of WebMILS that was operational during the period of the evaluation. The panelists felt that the process they followed (independent assessments followed by a group meeting) was effective, although there was some feeling that more training might be helpful before undertaking the independent assessments.

The evaluation experience also focused on the challenges of moving from an assessment of individual indicators using the matrix formulation to overall conclusions about a country's compliance with international labor standards. The formulation of discrete indicators for each labor standard provides a basis for a more explicated or transparent discussion of compliance. But the proposed method does not provide guidance for integrating conclusions about individual indicators into a broader assessment of a country's compliance with international standards. This issue raises important questions about the relative importance of different indicators and whether priorities could be established to direct such an effort.

Comments
Suggested Citation

Evaluation of the NAS-ILAB Matrix for
Monitoring International Labor Standards

Project Report

February 23, 2009

Submitted to

Bureau of International Labor Affairs
Department of Labor

DOLB079K26329

Lawrence S. Root, Principal Investigator
    Research Scientist, Institute for Research on
        Labor, Employment and the Economy
    Professor, School of Social Work
    University of Michigan

Ada A. Verloren, Project Director
    Senior Research Associate, Institute for Research on
        Labor, Employment and the Economy
    Lecturer, Center for Afroamerican and African Studies
    University of Michigan

EXECUTIVE SUMMARY

Labor practices have become a part of many international trade agreements. In response to growing interest in international labor standards, the Department of Labor's Bureau of International Labor Affairs (ILAB) engaged the National Research Council of the National Academy of Sciences (NAS) to recommend a method to monitor and evaluate labor conditions in a country. The focus was on five labor standards: (1) Freedom of Association and Collective Bargaining; (2) Forced or Compulsory Labor; (3) Child Labor; (4) Discrimination; and (5) Acceptable Conditions of Work. NAS developed an approach which included formulating indicators for each of the five labor standards, creating a database of sources of reliable empirical information for assessing these indicators, and suggesting a method for assessing and monitoring compliance that includes a matrix framework for graphically presenting assessments.

The University of Michigan, in consultation with ILAB, evaluated this approach by having 3-person expert panels apply it to assessing and monitoring compliance in three countries. The panel members independently assessed the indicators in terms of level of compliance with international standards ("some problems," "more extensive problems," or "severe problems) and direction of change ("steady state," "improving," or "worsening"). Each panel then convened to discuss their assessments and seek to resolve differences in those assessments.

Observation and analysis of this experience revealed challenges in applying this approach and potential modifications that could improve its operationalization. For most of the indicators, there was not agreement among the independent assessments by the panel members. Information was lacking or outdated for many of the indicators and, when present, was usually not definitive. Panelists raised a number of suggestions for changes in the indicators and matrix approach, more generally. WebMILS, the database created to support such assessments, was seen as a useful resource, although concerns were raised about limiting information sources to those that are available on the Internet. There were also suggestions for updating and modifying the version of WebMILS that was operational during the period of the evaluation. The panelists felt that the process they followed (independent assessments followed by a group meeting) was effective, although there was some feeling that more training might be helpful before undertaking the independent assessments.

The evaluation experience also focused on the challenges of moving from an assessment of individual indicators using the matrix formulation to overall conclusions about a country's compliance with international labor standards. The formulation of discrete indicators for each labor standard provides a basis for a more explicated or transparent discussion of compliance. But the proposed method does not provide guidance for integrating conclusions about individual indicators into a broader assessment of a country's compliance with international standards. This issue raises important questions about the relative importance of different indicators and whether priorities could be established to direct such an effort.

Table of Contents

INTRODUCTION

International trade is a major element of most national economies. In 2007, U.S. exports and imports represented 12.2 percent and 17.0 percent, respectively, of the gross domestic product (GDP)—and these figures are projected to double in thirty years.[1] With this expansion, international trade agreements have also grown in importance and scope. Labor conditions and environmental standards are now a part of the discussion of provisions in bilateral and multilateral agreements, regional trade pacts, and special trade remedies.

Incorporating labor standards in trade pacts reflects both moral and practical considerations. Advocates for inclusion of such provisions argue that unfair labor practices and exploitation of vulnerable workers should not be a part of any country's competitive advantage. Those arguing against incorporating labor standards into trade agreements raise concerns about protectionism and other potential adverse impacts on trade. For both advocates and critics, however, there are also questions about how such provisions would be implemented. Are there effective and fair ways to assess and monitor compliance with international labor standards by trading partners?

In 2001 the Department of Labor's Bureau of International Labor Affairs (ILAB) engaged the National Research Council of the National Academy of Sciences (NAS) to recommend an approach to assessing and monitoring labor conditions in a country. The focus was on five labor standards:

1. Freedom of Association and Collective Bargaining
2. Forced or Compulsory Labor
3. Child Labor
4. Discrimination
5. Acceptable Conditions of Work

The first four are "core labor standards," as defined by the International Labor Organization (ILO). These are considered "core" because of their broad acceptance internationally and the general consensus that these standards should be met irrespective of the standard of living or economic resources of a country. The fifth labor standard, "acceptable conditions of work," has emerged in the context of U.S. trade discussions and includes wage levels, hours of work, and health and safety provisions.

The National Research Council established a Committee on Monitoring International Standards to direct the effort to develop an approach to assessing compliance within a country. The result of that effort includes:

a. an explication of the five labor standards, including the formulation of a set of discrete indicators for each labor standard;

b. a database of sources of information for assessing these indicators; and

---

[1] Morrison, W.M. and Cooper, W.H, (2008). *The Future Role of U.S. Trade Policy: An Overview* (RS22914). Washington, DC: Congressional Research Service (CRS).

c. a matrix framework for graphically presenting assessments of indicators and for facilitating discussion of compliance with international standards.

The work of the Committee and its conclusions are documented in several publications, including an overview volume, *Monitoring International Labor Standards*.[2]

The indicators developed for each of the five labor standards are grouped into three categories, reflecting the legal framework in a country, government performance with respect to these indicators (e.g., resources devoted to enforcement of labor provisions), and overall outcomes in the country. For example, there are thirty-eight indicators for Freedom of Association and Collective Bargaining: twenty-one for the legal framework; thirteen for government performance; and four indicators of overall outcomes. The total number of indicators varies across labor standard. Freedom of Association[3] has the most, with thirty-eight indicators, and Forced or Compulsory Labor, with sixteen indicators, has the fewest.

The NAS-ILAB project also created WebMILS, an electronic database addressing international labor standards and the assessment of indicators. WebMILS includes general information about labor standards and compliance, as well as the indicators formulated by the NAS project for assessing the individual labor standards. The database provides an overview of the NAS's recommendations, including many excerpts from *Monitoring International Labor Standards*. A key element of the database is its links to external sites for researching the indicators to assess compliance. WebMILS is implemented on the ILAB section of the Department of Labor's website:

http://www.dol.gov/ilab/webmils/

The third element of the NAS-ILAB approach is a "matrix" formulation, which calls for assessing each of the indicators in terms of the level of compliance with standards (some problems, more extensive problems, and severe problems) and whether problems with compliance are worsening, improving, or staying the same. Based on these two dimensions, each indicator can be placed in a cell in a 3x3 matrix in figure 1, below:

Figure 1: NAS-ILAB Matrix Framework

	Improving	Steady State	Worsening
Some problems			
More extensive problems			
Severe problems			

---

[2] National Research Council (2004). *Monitoring International Labor Standards: Techniques and Sources of Information.* Washington, DC: The National Academies Press.
[3] In this report, "Freedom of Association and Collective Bargaining" is sometimes shortened to "Freedom of Association" for ease of expression.

Although assessment of some indicators is relatively straightforward (e.g., whether a particular ILO convention has been ratified), most assessments rely on at least some element of subjectivity on the part of the individual assessor. The NAS committee did not suggest specific guidelines for operationalizing the levels of compliance because they felt it would not be possible "to set thresholds that are appropriate for every country and every situation." (*Monitoring International Labor Standards*, p. 28) Rather, they provide general guidance for assessing the indicators. "Some problems" are situations in which these problems are "not so frequent, broad, or severe as to seriously undermine compliance." At the other extreme, an indicator has "severe" problems when compliance with international standards is "difficult or impossible." "More extensive problems" refers to situations that fall between those two—"problems that are serious enough to raise questions about compliance in important areas or for particular groups, but not across the board" (p. 28). Criteria for assessing direction of change are not discussed.

Based on the hypothetical example provided in the text, the committee envisions a process in which individuals would independently assess the indicators and then come together to discuss differences in their assessments. In this way, the matrix provides a visual, heuristic framework for displaying the ratings of indicators and a means for comparing the rating of two or more assessors. This NAS-ILAB matrix approach was developed by the Committee as a "middle-road" between a "precise scoring or ranking" of a country and a strictly qualitative effort: "the matrix framework…makes the value judgments and data interpretations of assessors explicit and exposes them to challenges and debate" (p. 32). Although the report describes the matrix layout "as one possible way to analyze the data" (p. 27), it is the only method that the Committee posits for displaying assessments of the indicators and comparing and/or discussing differences between individual raters.

EVALUATION METHOD

The NAS report and matrix framework provides a general direction for an assessment method based on independent assessments followed by a meeting (or other process) to discuss and seek to resolve differences in assessments. The first step in the evaluation was to translate this general goal into a work plan. The overall structure agreed upon with ILAB called for the formation of three panels, made up of three labor rights experts, with each panel assessing indicators for one country. The individual panel members would independently rate the indicators and then convene for a one-day meeting to discuss their ratings and to seek to resolve differences.

In order to get a better idea of what the individual assessments would involve, the research team undertook an initial analysis of indicators for one country both to identify any issues associated with this process and to estimate the time needed to assess indicators. This initial analysis also suggested the importance of developing orientation materials and data input tools to assist individual panelists with the task of assessing a large number of indicators: keeping track of their progress and ensuring that indicators were not overlooked in the assessment process. The evaluation also sought to assess WebMILS as a primary tool for obtaining information about compliance within a country. The steps in the evaluation process are shown in figure 2, below.

Figure 2: Steps in the Evaluation Process
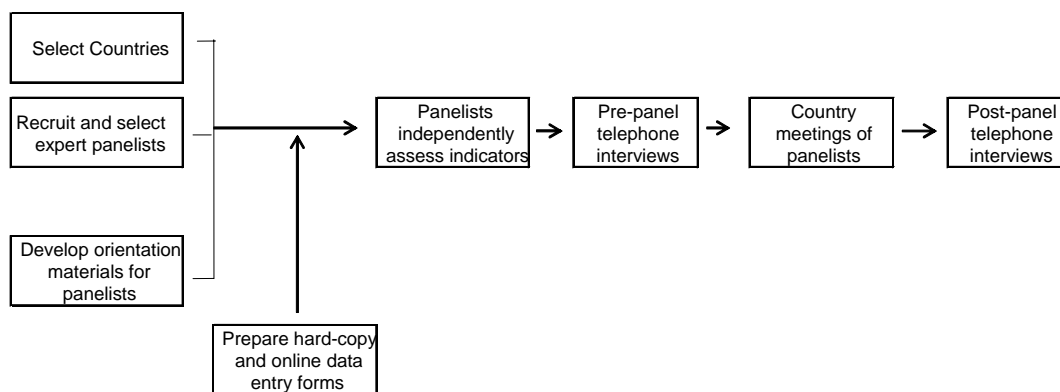
Selection of countries and labor standards assignment

In consultation with ILAB, three countries were selected. The countries used in the pilot testing of the matrix methodology are not identified in this report. This is to ensure that the focus remains on the evaluation of the *method* and not on the panels' findings for the test countries. Although at different stages of development, all three are generally considered "developing" countries. Their populations range in size from 30 million to well over 100 million people. They represent different regions, levels of

export activity, and labor rights traditions. The countries are referred to as Country A, Country B and Country C in this report.

Given the limited time available during a one-day group meeting for discussion, it was determined that instead of each panel assessing all of the labor standards—a total of 150 separate indicators for the five labor standards—each panel would assess two labor standards. All three panels would assess the indicators for Freedom of Association and Collective Bargaining. In addition, each panel would assess a second indicator. The assignments were:

Country A panel:	Freedom of Association and Collective Bargaining Acceptable Conditions of Work
Country B panel:	Freedom of Association and Collective Bargaining Forced or Compulsory Labor
Country C panel:	Freedom of Association and Collective Bargaining Discrimination

This design permitted a comparison of the assessment experience for one labor standard across the three panels/countries, as well as addressing three other labor standards.[4]

Recruitment and selection of expert panelists

For each country, a three-person panel was recruited. Suggestions for potential panelists were gathered from ILAB personnel, discussions with labor rights NGOs, corporate social responsibility personnel, and academic researchers. The backgrounds of the nine panelists who eventually participated on the three panels included:

Country A panel:

- Professor of industrial relations, with teaching and research experience in international and comparative labor relations and human rights
- Professor of law, with focus on international and comparative law, including work on labor rights in central and eastern Europe
- Union-based analyst and program director with focus on international labor conditions and extensive field experience in Asia and Europe

Country B panel:

- Independent consultant on corporate social responsibility, with policy and field experience in labor rights compliance in international supply chains
- Independent consultant on labor conditions and labor rights, with experience working with the NGOs and the U.S. government, particularly on programs in Africa

---

[4] It was decided that the panels in this pilot test would not assess the indicators for child labor. One factor in this logistical decision was the fact that the Department of Labor already has extensive efforts underway on child labor.

- Former ILO official with extensive experience in labor rights, including field, administrative, and policy responsibilities

Country C panel:

- Business school professor teaching U.S. and international labor law, with field experience working on labor rights in south and southeast Asia
- Executive for the U.S. office of a major international union organization, with experience working on labor rights issues with trade union bodies and business groups
- Senior program advisor with several labor rights NGOs, with experience training U.S. government officials on labor rights and assessment techniques

Each panel included at least one individual who had direct labor rights experience in the country being assessed.

Development and distribution of materials for the panelists

Once the panels were formed, each panelist was sent an orientation package. This included a description of the project, their responsibilities as a panelist, and a brief introduction to WebMILS (see Appendix A). They were also provided a copy of the *Monitoring International Labor Standards* text and a bound set of assessment forms for those indicators that they were responsible for assessing. An example of a hard-copy assessment form is displayed in Appendix B.

The individualized bound set of forms was intended as a tool for the panelist to facilitate the rating process. An online data entry system was created for the actual recording of their assessments (see Appendix C). This online system was the basis for collecting the assessments of the panelists prior to their meeting. The hard-copy forms, in addition to serving as a "workbook" for the panelists, provided a backup of each panelist's ratings that could be used in case there were technical problems with the online system.

In addition to the assessment of level of compliance and direction of change, the hard-copy and online system collected additional information from the panelists. For each indicator, they were asked whether they used sources from outside the WebMILS database and, if so, what sources they used. An open-ended field was provided for recording any comments about assessing the indicator. The panelists were also asked to indicate how confident they were in making the assessment: "low confidence" (the assessment is highly speculative); "moderate confidence" (some objective evidence); and "confident" (clear indications, strong objective evidence). The level of confidence question was added for two reasons. First, it was intended to provide more information about their individual assessments. The second reason was to encourage the panelists to make an assessment even if they did not feel that they had strong evidence—a common situation based on the experience of the test analysis undertaken by the research team.

Individual assessments by panelists

The panelists received their written instructions for the analysis in late February, with instructions that the assessments were to be submitted at least five days before the day of their respective panel meetings in early May. During this period, there were a few communications from individual panelists, but the research team minimized any feedback that might influence their individual ratings.

After the panelists submitted their assessments, they were interviewed by telephone before they convened for their panelist meeting. This pre-panel interview asked about reactions to the assessment process and the use of WebMILS. These interviews generally lasted from thirty-minutes to an hour. Appendix D lists the questions used in this pre-panel interview.

Preparation for the panel meetings

As discussed in the analysis section of this report, the panelists didn't agree in their assessments of most of the indicators. Recognizing that a one-day panel meeting would be insufficient for discussing all of these indicators, the research team reviewed the ratings and categorized each indicator in terms of whether there was low, medium, or high agreement among the panelists. The panel discussion of the individual indicators was then structured to give priority to those indicators for which there was the most disagreement in the independent ratings. As noted in the analysis section below, the extent of pre-meeting agreement on indicators varied among the panels. For example, the discussion in the panel for Country A had few indicators on which there was agreement in the independent assessments. Their discussion of indicators was completely taken up with those for which there was "low" initial agreement. In contrast, the discussion for the Country C panel, with fewer indicators with low agreement, also was able to discuss indicators with "medium" initial agreement among the panelists.

Panelists' comments in their ratings and in the pre-meeting telephone interviews also were used to structure the panel discussion. For example, as discussed later in this report, some of the indicators do not fit within the matrix framework. In such cases, low agreement typically reflected differences in how individual panelists dealt with this anomaly in using the matrix framework rather than differences in their assessment of a country's compliance. These indicators were discussed as a group at the beginning of the session rather than taken up individually.

The research team generated printouts for the panel members of their own ratings and comments for their use in the meeting. Panel members were not given the ratings of the others on their panel.

Panel meetings

All three panel meetings took place at the University of Michigan and followed a similar format.   The individual panelists arrived in the afternoon before the main meeting day. At a dinner meeting that evening, the panelists were provided with an agenda for the

following day (see Appendix E). As noted above, each panelist was also provided a printout of his/her own assessments and comments compiled from the online database. This initial dinner meeting provided an orientation to the work of the following day. It also served as a "get acquainted" session for the panelists. The identities of the panel members had been purposefully withheld prior to the meeting to avoid any potential impact on the individual assessments. For the most part, the panelists did not know each other, although in a few cases they had crossed paths in the past or knew of each other.

During the main meeting, on the following day, the panel discussed their individual assessments of the indicators as well as the use of the matrix framework and WebMILS. As described above, the discussion of individual indicators at the meeting was structured to focus on those for which there was less agreement among the panelists.

Following the discussion of each of the indicators for which there was lower initial agreement, the panelists were asked to re-rate the indicator in light of the discussion. These re-assessments were collected at the end of the session. In the discussion, panelists often revealed how their assessment changed based on the issues raised by others, but their written re-assessments were not shared with the panel.

Most of the meeting time was devoted to discussion of the individual indicators, along with more general points related to the matrix framework, WebMILS, and the availability and quality of information for making assessments. Toward the end of the session, the panelists were asked to try to reach a consensus on an overall level of compliance for the three broad groups of indicators: legal framework, government performance, and outcomes. Although this step was not a part of the process outlined in the NAS report, it was undertaken to explore issues involved in moving from a discussion of individual indicators to a broader assessment of compliance in a country.

Following the meeting, another telephone interview was conducted with each of the panelists to de-brief about the session. The post-panel interview questions are presented in Appendix F.

The analysis presented in this report is based on the individual assessments by the panel members, the interviews and other communications with the individual panelists, and the experience of the panel meetings. The research team took notes during the interviews and panel meetings. These were also audio-recorded. These recordings were subsequently reviewed to ensure that the major points were captured for the analysis. Atlas.ti, a qualitative software program, was used for coding the content of the interviews and discussions and to support analysis for this report.

ANALYSIS/FINDINGS

The findings from this evaluation are organized into sections addressing: the process used by individual panelists to assess the indicators; agreement/disagreement on those assessments; the nature of the indicators themselves; WebMILS as a resource in making assessments; and the panel process itself. These analyses draw upon data from all three panels. It also may be of interest to look at the individual panels and how they functioned. Appendix G includes a narrative overview of each of the individual panel discussions. This provides a summary of the session as well as identifying some of the key points identified at the time of the panel session.

ASSESSMENT PROCESS BY THE INDIVIDUAL PANELISTS

The panelists brought different backgrounds to the assessment task. Some had extensive experience with ILO conventions and researching those sources. For others, information about many of the individual indicators was unfamiliar. Some of these differences among panelists impacted both the processes that were followed in making the assessments and their willingness to make assessments when faced with ambiguous or missing data.

Using WebMILS for analysis: batch versus iteration

Panelists were encouraged to use WebMILS as the primary source of information, but were given little specific direction on how to approach the assessment task. Based on their reports, there were two general patterns of how they used WebMILS in assessing the indicators. Some went directly to the source materials, searched for information related to the country in question, and then printed that material (or stored it electronically). Once this "batch" search phase was completed, they used that material as the basis for their assessment rather than returning to WebMILS. The other pattern was more iterative: panelists went through the indicators one at a time, going back and forth to the WebMILS source material for each indicator.

Both of these approaches had strengths and weaknesses. The iterative approach led to some inefficiencies because a panelist would have to return to the same source for subsequent indicators. As one panelist remarked:

> *What I did find was that there seemed to be a fair amount of overlap and redundancy, and especially in terms of doing the assessments of Freedom of Association. I found that to be annoying. It just seemed like I just answered this and now I have to do it again... go back to where I was before and copy the same stuff and bring it up here.[5]*

---

[5] The panelists' comments presented in this report are drawn from the notes the researchers made in the course of reviewing the recorded interviews and panel meetings. They were not made by a professional

Most panelists appeared to use some combination of these approaches—printing out some key sources, but also doing some searches on an indicator-by-indicator basis. One panelist, in looking back on the panel discussion, described it this way:

> *We all used different strategies based on our individual work styles. My first attempt was to go one by one and poke around. At some point I said I should just be reading the source material first and get familiar with those because they all relate to each other and it will be more efficient. It took me hours to come to that understanding. That was a learning curve for me. I did not know whether each indicator would command discrete answers from discrete sources. I did not understand the overlapping and complementary nature of the sources.*

Another panelist expressed a similar view of the process:

> *With all of us, our approaches evolved. [Panelist a] did a little bit of what I did, which is to go real deep into the indicators, then looking at the data, then thinking "hey, this is not the best way to do this." then stepping back out of it, checking the body of information and then going through it. I went through the MILS data sources; I collected them and did not really delve into them. After I collected them, I went through the indicators, and then I went deeper to see if I could specifically answer the indicators, and if not, I looked for most of the alternative sources.*

This description suggests that one weakness of only using the "batch" approach may be a greater tendency to not return to WebMILS or seek additional sources when the originally collected material did not have relevant information for an indicator.

Non-assessment of indicators

Most panelists did not assess some of the indicators. Usually non-assessment occurred when a panelist felt that there was insufficient information to make an assessment. Non-assessments also occurred when a panelist felt that the response options were not appropriate. For example, as described in the section below on individual indicators, whether a country has ratified an ILO convention is a "yes-no" question. If the country had ratified the convention, some panelists chose to use "some problems," as the most positive response, while others left it blank, feeling that "some problems" wasn't accurate. A similar problem occurred when a panelist felt that there were no compliance problems, which is not one of the possible responses in the current matrix formulation.

---

transcription service and, while true to the content of the statements, they may not have the verbatim quality associated with professional transcriptions.

PATTERNS OF AGREEMENT/DISAGREEMENT

Appendix H displays the panelists' assessments of the indicators in matrix form for the three countries. This parallels the format provided in the example provided on pp. 30-31 of *Monitoring International Labor Standards*. That hypothetical example compares the assessments of two raters, using bold-face type to highlight the six indicators (of thirty) for which there are differences in the ratings. But the matrix display of ratings in Appendix H is difficult to scan to identify instances of agreement or disagreement on individual indicators. This difficulty arises from the fact that in the pilot test, the ratings of three panelists are displayed rather than two. It also reflects the fact that in the pilot test there were far more instances in which the three raters disagreed than in the hypothetical example. For the purposes of this analysis, patterns of agreement or disagreement in the assessment of the indicators are examined in terms of the whole panel and by comparing pairs of panelists. We first look at agreement among each panel as a whole.

Agreement on ratings by all members of a panel

From the perspective of each panel as a whole, we can examine how often the panelists all agreed in their initial assessments. In practice, non-assessments reduce the number of indicators for which there could be agreement among all three panelists. This is particularly striking for the panels for Country A and Country B. The extent of non-responses is displayed in table 1.

For Freedom of Association and Collective Bargaining, Panel A and Panel B had a high incidence of indicators with non-assessments: 50.0 percent (19 indicators) and 68.4 percent (26 indicators), respectively, of the indicators had at least one non-assessment. This high proportion of indicators with a least one non-assessment is largely the result of one panelist on each of those panels who had a particularly large number of non-assessments. That same pattern of non-assessment is even more pronounced with regard to the second labor standard these panels addressed: Acceptable Conditions of Work (Panel A) and Forced or Compulsory Labor (Panel B).

Looking at agreement by all three members of each panel, we see in table 1 that relatively few indicators were assessed in the same way by all panelists. For example, Panel C, with fewer non-assessments than the other panels, had only thirteen Freedom of Association indicators for which all members rated level of compliance the same way. That represents 34.2 percent of the thirty-eight indicators or 39.4 percent of those indicators that did not have any non-assessments. Agreement on indicators of Discrimination, the second labor standard addressed by Panel C, was also relatively high: 44.4 percent of indicators that had no non-assessments.

The figures in table 1 show more agreement on direction of change than on level of compliance. This may reflect differences in these two types of assessments.

Table 1
Initial Agreement on Assessment of Indicators by Individual Panelists

a. Indicators with at least one non-assessment

Freedom of Association and Collective Bargaining n=38 indicators						Acceptable Conditions of Work n=37 indicators		Forced or Compulsory Labor n=16 indicators		Discrimination n=31 indicators	
Panel A		Panel B		Panel C		Panel A		Panel B		Panel C	
n	Pct.	n	Pct.	n	Pct.	n	Pct.	n	Pct.	n	Pct.
19	50.0%	26	68.4%	5	13.2%	30	81.1%	13	81.3%	13	41.9%

b. Indicators with initial agreement among all panelists

Freedom of Association and Collective Bargaining n=38 indicators						Acceptable Conditions of Work n=37 indicators		Forced or Compulsory Labor n=16 indicators		Discrimination n=31 indicators	
Level of Compliance:											
Panel A		Panel B		Panel C		Panel A		Panel B		Panel C	
n	Pct.	n	Pct.	n	Pct.	n	Pct.	n	Pct.	n	Pct.
2	5.3%	6	15.8%	13	34.2%	1	2.7%	0	0.0%	8	25.8%
Excluding indicators with at least one non-assessment:											
	10.5%		50.0%		39.4%		14.3%		0.0%		44.4%
Direction of Change:											
Panel A		Panel B		Panel C		Panel A		Panel B		Panel C	
n	Pct.	n	Pct.	n	Pct.	n	Pct.	n	Pct.	n	Pct.
9	23.7%	12	31.6%	28	73.7%	5	13.5%	3	18.8%	11	35.5%
Excluding indicators with at least one non-assessment:											
	47.4%		100.0%		84.8%		71.4%		100.0%		61.1%

Compared to level of compliance, there was much less overall variation (i.e., higher agreement) in the assessments of direction of change. This can be seen in table 2, in which the distribution of the ratings given by all panelists is displayed. For "level of compliance," "some problems" was the modal rating, but an assessment of "more extensive problems" was seen almost as frequently. The panelists characterized the level of compliance as "severe" less often, but, still, it accounted for 20.4 percent of the assessments. In contrast, there was much less variation in the ratings of direction of change. "Steady state" was selected 81.6 percent of the time. Compliance was characterized as "improving" in 12.7 percent of the assessments and "worsening" only twenty-eight times, 5.7 percent of the assessments. The dominance of "steady state" may reflect a reality that at most points in time the level of compliance is not changing significantly. It may also suggest that "steady state" tends to function as a default category in the absence of any clear indication of direction of change.

Pair-wise agreement among panelists

Returning to level of compliance, we can consider agreement between pairs of panelists. In contrast to agreement among all members of the 3-person panels, comparing pairs of assessors provides another view of how much agreement and disagreement existed within the panels. It also has the advantage of including more points of comparison because an instance of non-assessment by one panelist does not limit comparing the ratings of the other two. Table 3 presents the pair-wise agreement on level of compliance for the three panels: panelists a and b (a:b); panelists a and c (a:c), and panelists b and c (b:c). For each panel, the number of indicators rated by each pair is shown, along with the number of times the pair were in agreement, and "agreed indicators" as a percentage of the ratings that the pair of panelists had in common (that is, when they both rated an indicator). As noted above, non-assessments diminish the number of indicators for which there is the possibility of agreement. As was the case for panel-wide discussion, non-assessments are a major factor for Panel B and, to a lesser extent, for Panel A.

In general, pairs of panelists tended to agree between about 30 percent to nearly 70 percent of the time. This is considerably lower than the 80 percent agreement in the hypothetical matrix example presented in *Monitoring International Labor Standards*. For Freedom of Association and Collective Bargaining, the labor standard addressed by all three panels, Panel A had lower levels of agreement (31.6 percent to 36.4 percent); the range of pair-wise agreement was higher for Panel B (52.6 percent to 66.7 percent) and Panel C (52.9 percent to 58.8 percent).

Table 2
Distribution of Ratings for Level of Compliance and Direction of Change

Level of Compliance

	Panel A		Panel B		Panel C		Three Panels	
Some problems	76	47.8%	61	49.2%	74	39.4%	211	44.8%
More extensive problems	40	25.2%	45	36.3%	79	42.0%	164	34.8%
Severe	43	27.0%	18	14.5%	35	18.6%	96	20.4%

Direction of Change

	Panel A		Panel B		Panel C		Three Panels	
Steady state	134	80.7%	102	73.9%	164	88.2%	400	81.6%
Improving	13	7.8%	35	25.4%	14	7.5%	62	12.7%
Worsening	19	11.4%	1	0.7%	8	4.3%	28	5.7%

Table 3
Pair-Wise Agreement among Panelists—Level of Compliance

Freedom of Association
n=38 indicators

Pairings of panelists	Panel A				Panel B				Panel C		
	a:b	a:c	b:c		a:b	a:c	b:c		a:b	a:c	b:c
Number of indicators assessed in common	19	26	22		38	12	12		34	38	34
Number of agreements	6	9	8		20	7	8		20	21	18
Percent agreement	31.6%	34.6%	36.4%		52.6%	58.3%	66.7%		58.8%	55.3%	52.9%

	Acceptable Conditions of Work n=37 indicators Panel A			Forced or Compulsory Labor n=16 indicators Panel B			Discrimination n=31 indicators Panel C		
Pairings of panelists	a:b	a:c	b:c	a:b	a:c	b:c	a:b	a:c	b:c
Number of indicators assessed in common	7	24	11	16	4	4	20	21	26
Number of agreements	2	14	4	6	0	3	9	13	13
Percent agreement	28.6%	58.3%	36.4%	37.5%	0.0%	75.0%	45.0%	61.9%	50.9%

Agreement between a pair of raters can also be characterized using Cohen's *kappa*, a statistical measure developed to quantify the level of agreement between raters ("inter-rater reliability") taking into account chance agreements. Values for *kappa* generally range from "0," indicating no more agreement than would have been expected by chance, to "1," indicating complete agreement. Scores of 0.400 or higher tend to be interpreted as indicating moderate agreement between pairs of raters. Negative *kappa* values can occur and generally are interpreted to indicate less agreement than would be expected by chance.

Kappa values were calculated for each pairing of panelists (see table 4). For this analysis, separate *kappas* were calculated for the assessments of level of compliance and direction of change. As noted in the table, *kappas* are not displayed in three cases where there are a very small number of assessments that can be compared (because of a high number of non-assessments). In two other instances, *kappas* could not be calculated because there was no variance in the ratings—all of the ratings of direction of change were "steady state."

The *kappas* presented in Table 4 provide another indication of the low level of agreement between panelists in their individual, pre-meeting assessments. None of the *kappas* for level of compliance were above 0.400 and only two were above that for direction of change.

Agreement and disagreement on assessments in context

The level of agreement—or lack of agreement—in the assessments by individual panelists for each of the three countries is not surprising, given the nature of the assessment task. Consistency among raters can be thought of as a function of (1) the clarity of the task, (2) the uniformity of available information, and (3) the extent to which criteria for ratings are operationalized. In terms of the clarity of the task, the indicators are intended to represent discrete considerations that can inform a broader assessment of the relevant labor standard. But many, if not most, of the individual indicators are themselves complex and subject to interpretation, as discussed more specifically in the next section of this report.

On the second point—uniformity of the information used for the assessments—the panelists all used WebMILS for relevant information, but they frequently referenced different sources of information—both within WebMILS and through use of non-WebMILS sources.

Table 4
Inter-Rater Reliability—Cohen's *kappas* for the
Pair-Wise Agreement among Panelists*

Freedom of Association and Collective Bargaining

	Panel A				Panel B				Panel C		
Pairings of panelists	a:b	a:c	b:c		a:b	a:c	b:c		a:b	a:c	b:c
Level of Compliance	-0.075	-0.009	0.055		0.257	0.127	0.254		0.361	0.244	0.237
Direction of Change	-0.070	-0.033	0.135		0.386	0.213	0.121		[no var.]	[no var.]	0.096

	Acceptable Conditions of Work				Forced or Compulsory Labor				Discrimination		
	Panel A				Panel B				Panel C		
Pairings of panelists	a:b	a:c	b:c		a:b	a:c	b:c		a:b	a:c	b:c
Level of Compliance	-0.111	0.273	0.167		-0.226	-0.231	[<5 cells]		0.082	0.111	-0.157
Direction of Change	0.633	0.191	0.077		0.244	[<5 cells]	[<5 cells]		0.010	0.121	0.428

* The paired assessments used for calculating these *kappas* excluded the few indicators that suggested a "yes-no" response because of the "answer-set problem" described in the next section of this report.

Perhaps the greatest source of variation arises from the operationalization of levels of compliance. As discussed in the introduction to this report, *Monitoring International Labor Standards* provides a conceptual distinction between the ratings that relies on the discretionary judgments of the individual panelist:

> In general, 'some problems' with compliance are those that are not so frequent, broad, or severe as to seriously undermine compliance with a given standard. 'Severe problems' are such as to make compliance difficult or impossible. We expect, however, that many indicators for many countries will fall into the middle category: problems that are serious enough to raise questions about compliance in important areas or for particular groups, but not across the board. (p. 28)

Given the complexity of the indicators, inconsistency and missing information in reports or other substantiating material, and the extent of discretionary judgments involved, it is not surprising that expert assessors reach different conclusions.

Discussion in the panel meetings provided numerous examples of differences arising from each of these factors. The easiest to resolve were those in which panelists based their assessments on different information. It was not unusual in the panel meetings for panelists to remark they had not seen a particular report that someone else had found helpful and, in light of that additional information, they would modify their assessment. It was also common for panelists to acquiesce to changing their assessment when they felt it was a borderline situation, saying something like: "I rated it as *some problems*, but I could live with *more extensive problems*."

Other sources of differences in assessments were less likely to be resolved in discussion. In some cases, there were distinct differences in the interpretation of an indicator. For example, discussion of the legal framework indicators raised differences in whether panelists should consider how the law operates in practice or only the "letter of the law." This issue is discussed with regard to the individual indicators in the following section. Another recurring question was whether compliance within a country should be assessed compared to the ideal standard or to the situation in other, similarly-situated countries. It was not unusual for one panelist to characterize a compliance problem in country X as severe and for another panelist to respond that it was a lot better than in countries Y and Z.

The individual orientation of panelists is another potential source of disagreement. In the course of the panel discussions and the post-panel interviews, the issue was raised of how the personal views of panelists toward labor rights, the role of government, etc., impact their assessments. As one panelist put it:

> *… it is still a judgment call. There are not strict criteria. What constitutes level of compliance will depend on where someone is coming from in terms of approach, point of view, moral commitment, political commitment.*

Analysis of the ratings did not reveal systematic differences in how individual panelists rated level of compliance. The distribution of these ratings for each panelist for all of the indicators that they assessed is shown in table 5. Differences in the backgrounds of the raters do not appear to be associated with systematic differences in how they rated compliance. For example, one of the panelists working with a union group characterized problems as "severe" more frequently than the others on the panel, but the other union-based expert used "severe" less often than the others on his panel. On a panel-wide basis, those assessing Country A characterized compliance problems as "severe" somewhat more often than Panels B and C. It is not clear whether this represents differences in the countries themselves or in the orientations of the panelists.

Another aspect of the orientation of individuals is how confident they were in their assessments. Table 6 shows the distribution of how the panelists characterized their confidence in their assessment of each indicator. The table also includes an average confidence score ("low confidence" = 1; "moderate confidence" = 2; and "confident" = 3). The average "level of confidence" ranged from a low of 1.75 (between low confidence and moderate confidence) to a high of 2.74 ("confident" most of the time). In the course of panel discussions about different ratings, panelists frequently conditioned their remarks by referring to their level of confidence. For example, one panelist described his willingness to change his rating this way:

> *Absent harder data, I did not think it was fair to say the problems were severe. I also said I had low confidence; my conclusion was highly speculative. I said "some problems," steady state. I did not see the Transparency International report. That might have bumped up my level of compliance.*

Reassessing selected indicators

As described in the "Methods" section, the discussion of individual indicators in the panelist meetings focused on those indicators for which there was less agreement in the individual ratings. After discussing an indicator, the panelists were asked to reassess the level of compliance. At the end of the session, these reassessments were collected. The reassessments, along with the panelists' initial rating and their confidence level with that initial assessment, are displayed in Appendix I.

The indicators in Appendix I do not precisely correspond to those indicators that had low initial agreement. For example, depending upon the content of the discussion, some low-agreement indicators that were discussed were not reassessed by the panelists. Also, some indicators which did not have low initial agreement were discussed by the group and, based on that discussion, panelists reassessed them.

The opportunity for reassessing the ratings led to greater agreement on some of the indicators, but not others. In order to assess the extent to which the

Table 5
Ratings of Level of Compliance by Panelists

Panel A

Ratings	Panelist a *Industrial relations professor*		Panelist b *Law professor*		Panelist c *Union-based analyst*	
	n		n		n	
Some Problems	22	43.1%*	16	48.5%	38	50.7%
More Extensive Problems	14	27.5%	4	12.1%	22	29.3%
Severe Problems	15	29.4%	13	39.4%	15	20.0%

Panel B

Ratings	Panelist a *Ind't consultant (business)*		Panelist b *Ind't consultant (NGO, gov't)*		Panelist c *ILO expert (labor rights)*	
	n		n		n	
Some Problems	22	40.7%	28	51.9%	11	68.8%
More Extensive Problems	23	42.6%	18	33.3%	4	25.0%
Severe Problems	9	16.7%	8	14.8%	1	6.3%

Panel C

Ratings	Panelist a *Business/law professor*		Panelist b *Union-based official*		Panelist c *Ind't consultant (NGO, gov't)*	
	n		n		n	
Some Problems	32	54.2%	17	28.3%	25	36.2%
More Extensive Problems	21	35.6%	28	46.7%	30	43.5%
Severe Problems	6	10.2%	15	25.0%	14	20.3%

*The percentages represent how often a panelist rated level of compliance as "some problems," "more extensive problems," or "severe problems."

Table 6
Panelists' Confidence in their Assessments

Panel A

	Panelist a		Panelist b		Panelist c	
Freedom of Association						
Low Confidence	3	9.7%	0	0.0%	1	2.6%
Moderate Confidence	20	64.5%	19	61.3%	8	21.1%
Confident	8	25.8%	12	38.7%	29	76.3%
Average confidence*		2.16		2.39		2.74
Acceptable Conditions of Work						
Low Confidence	4	14.8%	1	3.4%	4	11.1%
Moderate Confidence	3	11.1%	10	34.5%	9	25.0%
Confident	20	74.1%	18	62.1%	23	63.9%
Average confidence		2.59		2.59		2.53

Panel B

	Panelist a		Panelist b		Panelist c	
Freedom of Association						
Low Confidence	12	32.4%	2	5.6%	3	12.5%
Moderate Confidence	13	35.1%	19	52.8%	6	25.0%
Confident	12	32.4%	15	41.7%	15	62.5%
Average confidence		2.00		2.36		2.50
Forced or Compulsory Labor						
Low Confidence	9	56.3%	4	25.0%	0	0.0%
Moderate Confidence	2	12.5%	8	50.0%	3	42.9%
Confident	5	31.3%	4	25.0%	4	57.1%
Average confidence		1.75		2.00		2.57

Panel C

	Panelist a		Panelist b		Panelist c	
Freedom of Association						
Low Confidence	5	13.2%	12	31.6%	6	15.8%
Moderate Confidence	15	39.5%	13	34.2%	14	36.8%
Confident	18	47.4%	13	34.2%	18	47.4%
Average confidence		2.34		2.03		2.32
Discrimination						
Low Confidence	0	0.0%	1	4.0%	8	25.8%
Moderate Confidence	11	52.4%	16	64.0%	10	32.3%
Confident	10	47.6%	8	32.0%	13	41.9%
Average confidence		2.48		2.28		2.16

* Levels of confidence: 1= low confidence; 2 = moderate confidence; 3 = confident

discussion produced greater concurrence among the panelists, the researchers created a *difference index* as a measure of agreement based on the differences between the three pairs of ratings. Assigning a numeric value corresponding to the rank order of the responses ("some problems" = 1; "more extensive problems" = 2; "severe problems" = 3), the sum of the absolute values of the differences between the ratings of the three panelists could be calculated. For example, if an indicator had the following assessments by the three panelists:

	Rating	Numeric value
Panelist a	"some problems"	1
Panelist b	"severe problems"	3
Panelist c	"more extensive problems	2

The difference index for indicator #1 would be 4:

Panelist a – Panelist b	1 – 3 = -2
Panelist a – Panelist c	1 – 2 = -1
Panelist b – Panelist c	3 – 2 = 1
difference index	|-2| + |-1| + |1| = 4

In this hypothetical example, the various ratings by the three panelists results in a *difference index* of 4.

A cumulative *difference index* was calculated for Panels A and C to see the extent to which discussion among the panelists reduced the differences in their individual assessments of level of compliance. (A cumulative *difference index* is not presented for Panel B because the incidence of initial non-assessments renders the measure less useful.) The reduction in the cumulative *difference index* for Panels A and C are displayed in table 7. For both panels, there was less disagreement after discussion, although the size of the reduction was greater for Panel A than Panel C. Partially this reflects the larger pre-discussion differences between the ratings, with Panel A having a larger initial cumulative *difference index* with a smaller number of indicators. From a more impressionistic point of view, the discussion by Panel C also seemed to have more areas of substantive disagreement on interpretation of indicators (e.g., whether assessing the legal framework should go beyond the "letter of the law" to how the law is enforced).

Table 7
Change in Level of Agreement after Discussion

	Panel A	Panel C
Number of indicators reassessed (excluding those with non-responses)	10	16
Total *difference index*		
before discussion	36	34
after discussion	26	28
Reduction in *difference index*	10	6
Percent reduction	-27.8%	-17.6%

Moving to overall assessments

The NAS-ILAB matrix methodology doesn't provide directions for how to move from the assessment of individual indicators to more general conclusions for a country. As a step in this direction for the evaluation, panelists were asked for an overall rating of the level of compliance for each labor standard for the three broad areas: legal framework; government performance; and outcomes. This task was introduced toward the end of Panel A's discussion in order to simulate what would be the next logical step in making an assessment of the country as a whole. The discussion among the panelists provides an indication of some of the complications of moving from individual indicators to an overall assessment. For example, panelists noted that a simple summation of the ratings could be misleading because some indicators were more important than others in their impact on the status of labor rights in a country. The issue of moving from individual indicators to a broader characterization of compliance in a country is discussed later this report.

Another issue that arose was the potential consequences of the overall ratings. Panelists noted that although they might be willing to make assessments under conditions of uncertainty in the context of an evaluation, they would be hesitant to venture such judgments if those ratings meant serious economic sanctions on a poor country.

The discussion about overall assessment re-visited issues that arose about individual indicators, such as incomplete or missing data and whether assessing compliance should compare conditions to those in other countries or to ideal standards. This latter point was most strongly expressed in Panel A, the first panel to have its meeting. Panel A only proceeded as far as an overall assessment of the legal framework for Freedom of Association. Panelists were split between characterizing it as having "more extensive" or "severe" problems. After much discussion, they reached a "reluctant consensus" on rating it as having "more extensive" problems. However, one panelist continued to argue that, according to the criteria in *Monitoring International Labor Standards,* there were severe problems.

Based on the experience in the first panel, the task of making an overall panel assessment for the country was described to the panelists earlier in the process (although the actual task still was undertaken towards the end of the session). More complete overall assessments were reached for countries B and C, although these also were often "reluctant" agreements in the face of differing interpretations and lack of solid evidence. Table 8 displays the overall assessments for countries B and C.

For both country assessments, panelists struggled with what they considered borderline determinations, such as whether the overall performance with regard to Forced or Compulsory Labor reflected "more extensive" or "severe" problems in Country B or the legal framework of Country C had "some" problems of compliance or these problems were "more extensive." Typically, the discussion would involve one of the three panelists acquiescing to the views of the other two. The direction of change for Freedom of Association in Country C reflects a movement toward greater democracy ten years

earlier. The level of compliance is characterized as improving if compared with conditions over a ten-year period, but there appears to have been little change in the last five years.

In the course of seeking agreement—and in some of the earlier discussions—panelists raised the option of having more gradations in the ratings, such as a 5-point scale rather than the 3-point ratings of some, more extensive, and severe problems. This is reflected in table 8, when the agreement is reported as "bordering on…," which reflected a middle ground as a compromise between levels in order to reach consensus. Although some panelists felt that having more gradations available in the response set would be helpful, there was no general agreement that that a 5-point scale would be an improvement over the 3-point scale used in the current formulation.

Reaching (or not reaching) a consensus

In the course of the meetings, panelists were asked to try to reach a consensus in two areas: the assessment of individual indicators and overall conclusions with regard to the legal framework, government performance, and overall outcomes in a country. With regard to individual indicators, differences in initial assessments arose from differences in interpretation of the indicators, the equivocal nature of data available, and the open-ended nature of the criteria for assigning categories. In the course of trying to seek consensus on indicators, panel members seemed most willing to change their assessments when they felt that new information was presented. For example, one panelist may have based his or her assessment on a report that another panelist had not seen.[6] There were also situations where particular expertise, such as experience in the country or specific knowledge of the law, led a panelist to acquiesce to changing an assessment in order to achieve consensus.

On the other hand, it appears that some aspects of the process worked against reaching a consensus on individual indicators. The fact that each panelist independently assessed the indicators prior to the group meeting created some level of "vested interest" in their individual ratings. In the absence of new information, revising one's assessment was a tacit statement that an assessment was not correct. This didn't appear to be a factor initially, but a "micro-history" developed within each panel as discussion proceeded. If a panelist felt that he or she were more often than others in the position of changing an assessment to reach consensus, there seemed to be a growing reluctance to do so. Previous decisions to change a rating appeared to reduce that individual's willingness to change additional ratings. This dynamic was intensified in the Panel C meeting, when the researchers varied the protocol by asking the panel to publicly reach a consensus on individual indicators rather than only having the panelists re-rate privately (i.e., in writing). As described earlier, this attempt at reaching a public consensus on individual

---

[6] In some instances when a panel member brought up information obtained outside of WebMILS (e.g., a report an NGO or a personal communication from a State Department source), other panelists remarked that they found it difficult to integrate this new information "on the fly." Some panelists also noted that the use of such outside sources was outside of the research protocol and they were unclear how it should be used in the context of the evaluation.

Table 8
Overall Assessments for Panels B and C

Panel B: Freedom of Association

	Level of Compliance	Direction of Change
Legal Framework:	Some problems	Improving
Government Performance:	More extensive Problems	Steady state *(but corruption continuing as a serious problem)*
Overall performance:	More extensive Problems	Steady state

Panel B: Forced or Compulsory Labor

	Level of Compliance	Direction of Change
Legal Framework:	Some problems	Steady state
Government Performance:	More extensive Problems	Improving
Overall performance:	More extensive Problems *(bordering on "severe")*	Steady state *(little information; forced labor improving somewhat because of political pressure; child prostitution common)*

* * * * * * * * * *

Panel C: Freedom of Association

	Level of Compliance	Direction of Change
Legal Framework:	Some problems *(bordering on "more extensive")*	Improving *(over last 10 yrs; "steady state" in last 5 years*
Government Performance:	More extensive Problems *(bordering on "severe")*	Improving *(over last 10 yrs; "steady state" in last 5 years*
Overall performance:	*(no rating)*	*(no rating)*

Panel C: Discrimination

	Level of Compliance	Direction of Change
Legal Framework:	Some problems *(bordering on "more extensive")*	Steady state
Government Performance:	More extensive Problems	Steady state
Overall performance:	*(no rating)*	*(no rating)*

indicators was quickly abandoned when the panel discussion of the first few indicators proved very time-consuming and seemed to be beginning to engender some degree of defensiveness and hostility within the group.

The dynamics of seeking a consensus on a more general assessment of the legal framework, government performance, and outcomes were somewhat different than those for the individual indicators. First, the panelists hadn't already made such an assessment, so there was no question of having to revise a previous assessment. Second, the task was more integrative and called for the participants to step back from the specifics to try to categorize the "big picture." It was in this context that the question of priorities among the indicators was raised. Also, when asked to step back from specific indicators and consider an overall rating, some panelists remarked that the consequences of such an assessment would be a factor in their judgments. As noted above, some panelists said they would be reluctant to venture an assessment—particularly under conditions of uncertainty—that might cut a poor country off from the benefits of a desired trade relationship.

Another aspect of the consensus that was reached with regard to an overall assessment of compliance was the fact that in some instances the agreement was on a middle ground between assessments (e.g., in table 8, "more extensive problems bordering on 'severe'"). Creating this kind of "in-between" assessment facilitated coming to a consensus in a way that hadn't been treated as an option when the panel discussed individual indicators.

CHALLENGES IN ASSESSING THE INDICATORS

The panelists identified a number of issues arising from their efforts to assess the indicators. An over-riding issue is the underlying problems of missing and skewed information in the national sources. This issue, which is also highlighted in *Monitoring International Labor Standards*, was raised by all of the panels. Self-censorship and political/economic motives are sources of potential distortion in government information sources.

Another major problem arises from the fact that official statistics often do not address the "informal" economy, which can account for most employment in developing countries. In Country A, for example, a panelist characterized five different employment situations that impact the extent to which workers are "visible" in national statistics:

(1) Regular employees who are registered and covered by social security
(2) Regular employees who are not registered
(3) Informal employees who do not receive regular pay
(4) Temporary employees and employees of third-party service companies
(5) Agricultural workers/small farmers.

Often, only the first category of workers is included for official statistics and reports.

A picture of the situation in a country may also be skewed by the differential attention received by certain industries. Some export industries, such as apparel and toys, have received more attention than others from NGOs. While conditions in those industries may be representative of conditions elsewhere, this may not be the case. This raises the related problem of how to extrapolate from anecdotal reports to a broader assessment of conditions in a country.

In addition to these general problems related to assessment, a number of issues were identified with regard to individual indicators. A listing of specific issues related to the indicators is included in Appendix J. Most of these issues fall into one of several categories: operationalization of the levels of compliance; the time period for assessing "direction of change;" parameters for assessing "legal framework" indicators; answer-set problems; clarity of wording and/or intent; and relationship to recognized international standards

Operationalization of the levels of compliance

Questions were raised by panelists about the basis for assessing level of compliance. As noted above, the distinction between the levels is based on how widespread and pervasive are the limits on compliance. "Some problems" are situations in which violations are "not so frequent, broad, or severe as to seriously undermine compliance." "Severe problems" is to be used when violations make "compliance difficult or impossible." (p. 28) In *Monitoring International Labor Standards*, they also note, however, that assessments have to consider "the resources available to the government for compliance, the poverty of the country, the number of workplaces, and other demands for government resources" (p. 24). Assessments of level of compliance, therefore, are to be tempered by recognition of the constraints facing a country. In practice, panelists often implicitly or explicitly compared the country they were assessing with other similarly situated countries. A panelist might say that there were serious problems with an indicator but they were not as bad as in a neighboring country. On the other hand, a panelist looking at the same country might provide a lower assessment, focusing solely on the extent to which conditions fell short of the stated international standard.

Time period for assessing "direction of change"

All of the panels raised the question of what time frame should be used for assessing "direction of change." For example, two of the panelists for Country A used changes since 2000 as their basis; the third panelist only considered changes in the last two or three years. As noted in the discussion of table 6 (above), an increase in democratic practices ten years ago in Country C led to improvements in compliance, but there have been no discernable improvements since then. The time period for assessing direction of change in this situation would determine how it was characterized. One panelist suggested that it might be appropriate to think in terms of the time it takes for a change in law to be proposed, adopted, and implemented—perhaps seven or eight years.

Parameters for assessing "legal framework" indicators

Questions emerged concerning whether the way that a law is implemented should be a part of the assessment of the "legal framework" indicators. Should assessors consider how the law has been applied or the outcome of court cases?

The wording of some indicators clearly relates solely to what is written in law. For example, Acceptable Conditions of Work indicator A-2 is whether there are provisions in the "national constitution or law establishing one or more minimum wages." On the other hand, the wording of other legal framework indicators suggests more of an evaluation of the effectiveness of a legal provision. Freedom of Association A-16, for instance, is "the extent to which the government can limit picketing or occupation of the workplace." To assess this indicator, one might have to look at practice, such as how laws related to public order are applied in the case of labor actions. Other legal framework indicators more explicitly call for considering how the law operates in practice. For example, indicator A-9 under Discrimination is:

> whether there are laws that make nondiscrimination in employment
> <u>meaningful in practice</u> with equal access to education, training,
> vocational guidance, and placement services; maternity protection;
> and parental leave (emphasis added).

Including "meaningful in practice" appears to direct the attention of the assessor beyond the letter of the law to how it operates in practice.

Answer-set problems

The matrix approach offers three responses for level of compliance: some problems; more extensive problems, or severe problems. For some of the indicators, this answer set is not appropriate. First, there are situations in which there are "no problems," but this is not one of the possible responses. Second, some indicators call for a categorical (binary) response—"yes" or "no"—rather than a ranking. For example, either a country has or has not ratified a particular ILO convention. As mentioned with regard to indicators that were not assessed, panelists responded to this situation in different ways. If the country in question had ratified the convention, one panelist might use "some problems" (as the most positive response) while another might leave it blank.

A third answer-set problem involves "outcome" indicators, such as union density (Freedom of Association C-1) or strike activity (Freedom of Association C-2). While these indicators are certainly related to union activity in a country, it is unclear how they are related to levels of compliance. In fact, how these indicators should be interpreted is itself unclear. Higher rates of union density may well be associated with greater freedom of association, but does a lower rate indicate problems of compliance in a country? Similarly, does greater strike activity suggest compliance problems or a powerful labor movement?

Clarity of wording and/or intent

Panelists also raised questions about interpreting and assessing specific indicators. These are detailed in Appendix J. While many of these comments are specific to a particular indicator, several recurring issues can be discerned. One problem is an indicator that combines two or more issues—a so-called "double-barreled" question. For example, indicator A-2 under Freedom of Association is: "whether there are legal provisions that entitle workers or employers to establish and join organizations of their own choosing without previous authorization." Panelists noted that adding "without previous authorization" complicates the indicator and makes it difficult to assess. Is the indicator asking whether workers are free to unionize or is it asking about the process by which a union is certified as legitimate?

Freedom of Association A-15 provides another example of an indicator that combines several issues:

> the extent to which the government can forbid certain kinds of strikes (such as "protest strikes," "sympathy strikes," "go slow" strikes, or "work-to-rule" slowdown), or strikes under certain conditions (such as economic or political "crisis"), or strikes requiring a majority of workers involved to authorize a strike

This indicator includes three different potential bases for rating the compliance of governments in their restrictions of strikes: the type of strike; the political/economic context; and the process for strike authorization.

Relationship to recognized international standards

Questions were also raised about indicators that related to situations that are not covered by international labor standards. When an indicator does not correspond to an international standard, it is unclear what is meant by "compliance." This issue was touched on earlier with regard to outcome indicators, such as union density, that are relevant to the context of a country but are not clearly translatable to a measure of level of compliance. Indicator B-7 under Forced or Compulsory Labor provides another example. It asks whether the country is receiving technical assistance. Actively seeking technical assistance may be a useful indicator of a country's intent to address problems of forced labor, but it is not an obligation under international codes and, therefore, not an issue of "compliance."

USING WEBMILS AS A SOURCE FOR ASSESSMENTS

A major element of the NAS-ILAB project was "to construct a database for use in monitoring international labor standards." (*Monitoring International Labor Standards*, p. 16) WebMILS was created as a structure and source for qualitative and quantitative data that could be used to assess current conditions in countries. WebMILS contains key information from the *Monitoring International Labor Standards* volume, as well as links to data sources for assessing the indicators.

The structure of WebMILS is displayed in figure 3, below.[7] From the home page (http://www.dol.gov/ilab/webmils/), one can go to information about the international labor standards and their associated indicators as well as issues associated with assessing compliance and the matrix as a tool in the assessment process. This information draws upon relevant sections of *Monitoring International Labor Standards*. For example, selecting "International Labor Standards" brings one to a narrative description from the book of each of the five labor standards. "Assessing Compliance" allows the user to select a labor standard to see information about compliance with that standard, including the indicators developed for legal framework, government performance, and outcomes. This information pertains to international labor standards generally rather than to the situation of an individual country.

To retrieve information for assessing conditions in a country, one selects "Search WebMILS" and picks the country of interest. This brings the user to a page for that country, with a listing of the labor standards.[8] Selecting a labor standard takes the user to a list of the indicators for that standard followed by sources of information from international bodies, national sources, U.S. government sources, nongovernmental bodies, and other research. With the exception of the national sources, these sources are largely the same for all countries.[9] For example, the link to an international site, such as the ILO's Application of International Labour Standards (APPLIS) database, which documents ratifications of conventions, is the same for all countries. Once users go to that external source, they use the search mechanism at that site to look for information relevant to the country in question.

"National sources," on the other hand, are specific to the country being researched. Depending upon what is available in a country, WebMILS may provide links to a country's ministry of labor, statistical office, description of national laws, and other relevant sites. These websites may be in the national language of the country.

The panelists used WebMILS as their principal source of information for the assessments, although, as noted, additional sources were also used. In general, panelists felt that WebMILS was a useful tool for identifying data sources on the labor rights

---

[7] Figure 3 includes the major elements of the version of WebMILS used during the evaluation. Not displayed in the figure are options for viewing all of the sources of information included in WebMILS, frequently-asked questions, and providing feedback. This report focuses on the aspects of WebMILS that were used in testing the ILAB-NAS matrix methodology. There are some additional aspects of the WebMILS site that did not come into play in this exercise and appear to reflect other aspects of the Committee's work. For example, in addition to the indicators of legal framework (A), government performance (B), and overall outcomes (C), there are references to indicators for associated factors (D), human capital (E), and general country information (F). Although these additional areas enter into the discussion in *Monitoring International Labor Standards* and on WebMILS, they are not formally integrated into the matrix assessment process.

[8] In addition to the four "core" labor standards and acceptable conditions of work, there is also the option to select "Human Capital," because of its association with level of development and compliance with labor standards. Unlike the five other areas, there are no indicators specified. WebMILS does, however, provide sources for relevant background information.

[9] Some nongovernmental organizations and research reports are also specific to a country.

situation in a country. They thought it would be particularly helpful for those who are not experienced researching international labor conditions.

Figure 3
Structure of WebMILS
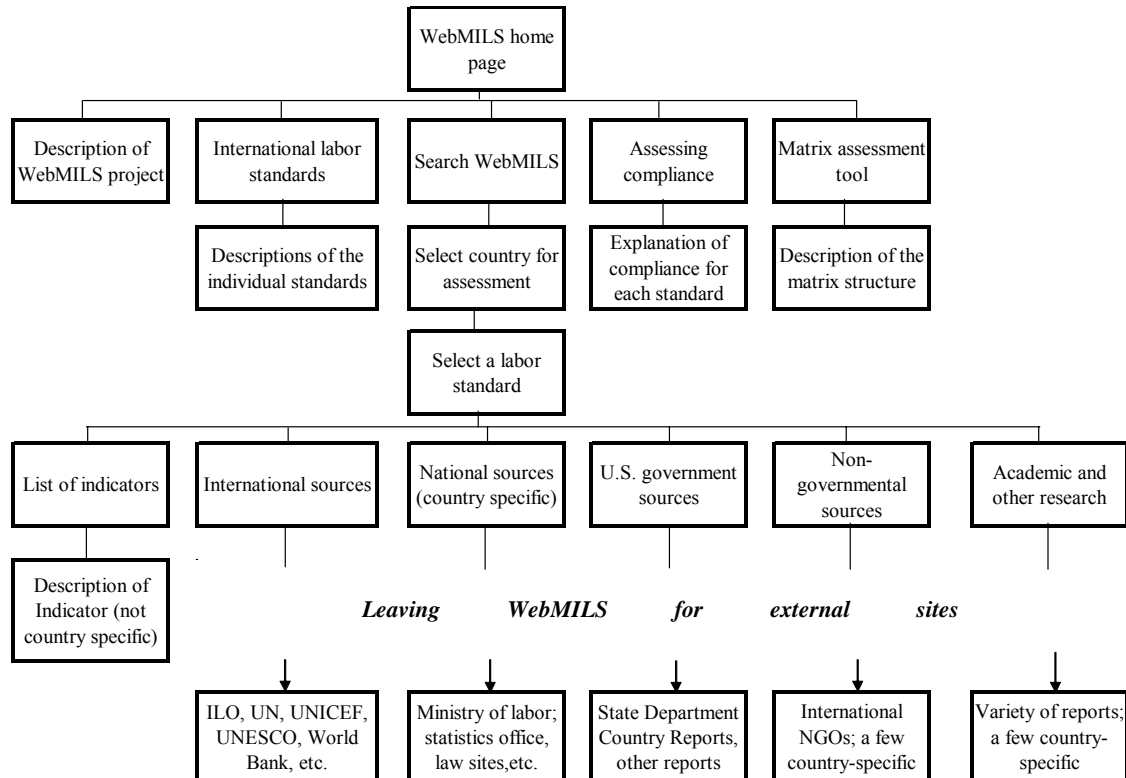
The panelists identified a number of issues related to their use of WebMILS for the assessment exercise. These included questions about the content of the system as well as its structure and functions.

Content of WebMILS

Panelists identified several content issues that arose in the course of using WebMILS for assessing compliance. As noted in the discussion of individual indicators in the previous section, panelists often found that there was simply insufficient information available to make an assessment. This is not a criticism of WebMILS, but, rather, a commentary on what information is available generally and, more specifically, what is available on the Worldwide Web (Web). Although the range of information sources accessed through the Internet has increased exponentially, there are still major gaps, particularly for developing countries. Even for the ILO, which has an extensive Web presence, key reports may not be accessible electronically. For example, panelists noted, as did *Monitoring International Labor Standards,* that ILO Article 22 reports, in which countries report on their performance with regard to conventions they have ratified, are an important source that is not currently available through the Internet.

WebMILS, by its very nature, draws exclusively from the Web and that creates a *de facto* limit on what is included.

Panelists also mentioned the importance of the user's background for the proper interpretation of some WebMILS sources. For instance, one might be tempted to judge a country's labor relations compliance based on the number of complaints lodged against it. This, however, could be misleading. One panelist noted that there are more labor relations complaints to the ILO about Canada than Colombia, where violence against trade unionists is a major problem. Beyond understanding the inherent limitations of data collection or what is available on the Web, there may also be issues of interpretation of reports. For example, some ILO documents use diplomatic language in which a strong criticism may be expressed in a muted and polite manner. An unsophisticated user might overlook what is intended as a serious concern because of the careful way it is articulated.

All of the panels expressed the idea that media sources, which are not included in WebMILS, can be an important source of information, particularly about current labor situations. The WebMILS designers did not include media sources because of concerns about the objectivity and consistency in news reports. They acknowledge, however, that media reports "may serve as a barometer of labor conditions." (*Monitoring International Labor Standards,* p. 37, footnote). Panelists felt that media sources should be considered, recognizing that care is needed in interpreting news reports. One panelist described newspaper stories as useful for providing facts, such as whether or not there was a strike, but less reliable in their characterization of a strike's causes or legal status. Without having someone "on the ground" in a country, newspapers can provide a picture of the current situation.

Structure and Function of WebMILS[10]

Panelists identified functional aspects of WebMILS that they felt could be addressed to improve its use. First, as with other databases, system maintenance and updating are ongoing challenges. Some WebMILS links to external sites aren't operational and there are instances in which the names of organizations have changed but this is not reflected in the database. Also, although there is a link for users to suggest additional resources, it is unclear whether WebMILS has been updated to include information sources that have emerged since the database was created.

In terms of ease of use, some panelists expressed no concerns. Others, however, described it as unwieldy and less efficient then they would have expected. One individual who was very familiar with international sources, was particularly critical of WebMILS, noting that he found it much easier to just go directly to those sources.

The problems that were identified are related both to the architecture of the site and lack of a search mechanism of the sort that Internet users have come to expect. As discussed earlier, a large portion of WebMILS is devoted to explicating the labor

---

[10] Since the conclusion of research on this project, WebMILS has been revamped and updated. The revision of WebMILS took into account panelists' suggestions described in this report, including those concerning structure, function and search capability.

standards, the indicators, and the information sources. Most information sources are the same for each country. For instance, one is directed to the same ILO, UN, or U.S. government sources whether assessing Afghanistan or Zimbabwe. The current WebMILS structure, however, suggests that the process is more "country-specific" than is the case. For example, to get to sources for assessing a country, one first selects a country (through the "Search WebMILS" link) and then a labor standard. The indicators for that labor standard are listed, but clicking on any indicator takes one to a general description of the indicator rather than to content related to the country that had been selected. When the user scrolls down to the "Sources" part of the page, most of the information sources—with the exception of the "national sources"—are the same for all of the countries.

Panelists also noted that the indicators listed on a country's page are not linked to the relevant sources. Instead, the sources are presented in a separate section, following the listing of indicators. As one panelist remarked:

> *At first I thought that if you clicked on an indicator it would take you to the precise sources. But you have to go to the bottom of whatever sources are there to try to figure it out.*

The current WebMILS structure, by having the user first select a country before going to individual labor standards and their indicators, gives the impression that the page will be more specific to the country than is the case. Once at the listing of indicators for a country, as suggested in the comment above, the user expects there to be a link to the relevant source(s). But this is not currently the case.

Search capabilities and WebMILS

The navigation issues related to the architecture of WebMILS touches upon another recurring theme among the panelists: the desirability of more targeted search capabilities. Searching on the Web has become increasingly sophisticated in the years since WebMILS was first designed and this has changed the efficiency of searching as well as the users' expectations for how an Internet resource can function. One panelist remarked:

> *I do a Google search for so much of my writing. In this case it is so unnatural to be limited to these sources of information. …With the limited number of sources [in WebMILS], they may be more objective and more vetted in a way, but things get lost and you never hear about a problem that could put your mind in a different direction.*

The dominance of searching with Google and other search engines shapes the template for how users expect to use websites. In the words of another panelist:

> *If I could take a magic wand and do one or two things, I would create a search engine [for WebMILS] that had the capacity to look at all of the sources and documents at one time rather than having to go into each and try to figure out how to get to where you needed to go.*

For somebody that does not spend a lot of time on the Internet with these searches, there was a level of frustration. This is hard. If I could go to the DoL WebMILS area and search all of the documents with a fairly robust search tool so that I can refine it in a way that identifies key words and really gets me to where I needed to be, … to not only what the government had on its books in terms of laws but also reaction to the adequacy of those laws from various stake holders, all of that in one place, I imagine it would reduce the effort and allow more access into the sources.

Given this view, it is not surprising that some of the panelists, particularly those who were more familiar with key sources, such as the ILO and the State Department, went directly to those sources, rather than restricting their information collection to WebMILS. This was at variance with the instructions to panelists for this evaluation, but it is probably indicative of how those with experience in international labor rights would proceed in assessing compliance.
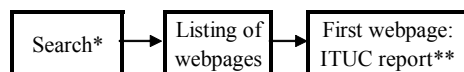
Although general web searches are not limited to information sources that have been vetted for consistency or objectivity, they are a useful tool in assessing compliance. For example, we can look at Discrimination in Country D.[11] The first "legal framework" indicator under Discrimination is whether the country has ratified ILO Convention 100 on equal remuneration. Using WebMILS, a user would first select Country D (from the Search WebMILS link), then select Discrimination. The indicators would be listed. The user would then go down to the Sources section, which follows the list of indicators and select the ILO APPLIS database. Selecting APPLIS would first go to a screen explaining that one is leaving the Department of Labor site for an external one and then to ILO's APPLIS home page, where one can find a list of the conventions ratified by Country D. The steps in this process are displayed in figure 4, below.

Figure 4
Example of Checking an ILO Ratification for Country D

Using WebMILS:

Search WebMILS → Select Country D → Select Discrimination → Indicator: ratif. ILO 100 → Source: ILO APPLIS → Find Country D's ratifications
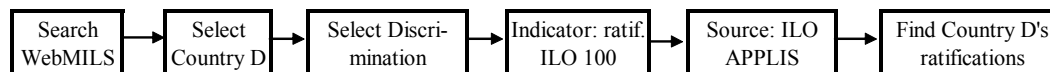
Using Google Search:

Search* → Listing of webpages → First webpage: ITUC report**

   *Search terms: "ILO Convention 100 Country D"
   **ITUC, "*Internationally Recognized Core Labour Standards in [Country D]" (2007* report for the WTO)

---

[11] The examples used in this exercise, are based on a real country. It was not one of the countries examined by the panelists.

An alternative way of finding this ratification information would be to do a Google search instead of using WebMILS. Searching on "ILO Convention 100 Country D," the first four of the hundreds of web pages retrieved are: (1) a 2007 report from the International Trade Union Confederation (ITUC) on the core labor standards in Country D; (2) a report on Country D by the NGO Anti-Slavery International; (3) an ILO news release about "decent work" in Country D; and (4) a Department of Labor-ILAB report about child labor in Country D. The ITUC report included Country D's ratification of ILO Convention 100. It also provided additional commentary/analysis relevant to both the convention and other aspects of discrimination.

Although this Google search on a ratification indicator successfully retrieved useful information, searches on other indicators were less helpful. For example, looking at the first "government performance" indicator—affirmative action—none of the first four web pages retrieved provided information relevant to Country D's affirmative action programs. As noted in Appendix J, however, panelists also found little or no information for this indicator using WebMILS. A Google search on the first indicator under "outcomes," led to UN/UNICEF web pages. These are the same sources found through WebMILS.

Internet searches, such as those with Google, offer an important resource, but they do not target sources that have been evaluated for reliability. The sources available through WebMILS, on the other hand, have been vetted and provide the user a measure of confidence in their accuracy. The experience of the panels suggests that Internet searches can be a useful tool in assessing compliance, although the sources found must be used with care. It also raises the question of whether a search capacity could be added to WebMILS to streamline its use.

THE PANEL PROCESS

The process for assessing compliance envisioned by the NAS-ILAB involves experts individually assessing labor standard indicators and then coming together as a panel to discuss these assessments and resolve differences. In examining the panel process in this evaluation, we can look at the composition of the panels, the pre-meeting orientation of the panelists, and the meeting itself.

Composition of the panels

For the evaluation, the 3-person panels were made up of individuals with a variety of experience in international labor standards—activists, academics, business consultants, and representatives of multi-stake holder groups. Each panel included one person with field experience in the region or country being assessed. Two of the panels also included others who had some experience in the country. Typically, the panelist with experience in the country did not consider himself or herself a "country expert" or specialist in that country; rather, each was an expert in international labor rights, with some direct field experience in that country. This distinction is relevant for the panel process. It is conceivable that a country expert might dominate discussion, particularly in the face of

limited documentary information. The panel discussions did not reveal such a dynamic, although knowledge of the country was an important contribution to the discussions. One of the panelists with country-specific experience remarked during the post-panel interview:

> *Throughout the process I was wondering if I didn't have some knowledge*
> *of [Country A] whether we would have discussed the non-counting of*
> *people, or the non-reporting on significant industrial activity in the*
> *country and the labor situation of those industries.*

In the panel discussion itself, he also mentioned that he felt that "you can't isolate what you know from what you read." Another panelist remarked how his overall view of the country impacted his assessments:

> *This is not a country that is consciously taking large-scale measures to destroy*
> *trade unions. [Country B] is a struggling democracy. It is not Colombia or some*
> *of the other countries. I gave it some latitude for the feeling behind it. If it had*
> *been Myanmar, I would have evaluated everything with a much more suspicious*
> *eye. This assessment of the country is not based on the data sources but on my*
> *experience.*

Knowledge of the country also may have influenced the process that panelists used in their individual assessments. A panelist with experience in the country described how his background led him to use sources beyond what was available on WebMILS:

> *Since I have some experience in [Country C], I went through what I have*
> *already accumulated … I posed a couple of questions to the State Dept.*
> *Regional Labor Counsel and he sent me copies of the new laws. So it was*
> *clear that the whole legal regime had changed. Then I went more*
> *systematically through the indicators and the sources and materials that*
> *were listed on WebMILS and I tried to follow that structure first before*
> *going to what I knew were the other sources.*

Another panelist mused about whether the perspective gained from the sources on WebMILS would produce an understanding that corresponded with the more comprehensive view of someone familiar with an individual country:

> *I was given an area of the world and a country that I did not know well. I*
> *think that was intentional. And it would have been interesting to me to use*
> *the system in a different situation; to see if the tool would allow me the*
> *ability to make judgments and come to conclusions that are consistent*
> *with what I know to be true in the real world…it certainly would be*
> *interesting for me to take this tool and go to a country that I know a lot*
> *better, China or some place in eastern Asia, and undertake a similar*
> *process and look and see if it came out anywhere close to what I know to*
> *be the case.*

The mix of backgrounds and orientations on the panels also enhanced the quality of the discussion. Panelists were aware of these differences and occasionally they directly factored into the discussion. As one panelist remarked:

> *Each of us brought to it our background. So we all had different sorts of takes. Having someone with a legal background was great. Having someone with an academic background around labor and labor movement issues was useful and brought a good perspective to it. And the practical experience I brought was useful...I thought that was a good match.*

That panelist also noted some of the specific contributions of those with different backgrounds:

> *[Panelist b] certainly had really good points to make about how the law works. When she said it, I thought, yes she is right. But as a non-lawyer it wasn't the first thing that came to my mind.*

As another panelist remarked, the discussion, involving individuals with a range of backgrounds "was interesting because more issues came up than I would have expected."

Overall, it was very helpful to have at least one panelist with experience in the country being assessed. That individual was able to identify issues that were not apparent from the Internet sources and to offer contextual perspectives that enriched the panel discussions. The mix of professional background/experience also added to the discussion.

Pre-panel preparation

As described above and in Appendix A, in preparation for the assessment task, panelists were provided with a memo summarizing WebMILS, the matrix approach, and the expectations of them as panelists. They also received a copy of *Monitoring International Labor Standards* and a bound set of response sheets for the indicators they were to assess (Appendix B). Some panelists suggested that more training would have been helpful prior to doing the individual assessments. This training could provide more information about the sources and their use. As one panelists noted in the post-panel interview:

> *As a recommendation for a process that would work better, I would think that a bit more guidance to the assessors would be helpful. It would have been helpful to say you should look at X, Y, Z. Here are some of the basics you should do and if you do anything beyond that you should record it. All three of us would have done at least the basic stuff with a greater degree of consistency. … There is nothing in the evaluation process that we did that suggested we should do any comparative stuff at all. All three of us had experience doing international stuff and had points of comparison. It would be better to systematize that by saying here is a really bad example of setting up a system with regard to the right to organize and bargain collectively. And you might draw on countries*

…where unions are not permitted or…where unions are state dominated.
And here is [a country with strong labor rights that] is completely
different and that is another point of comparison. And you can rate this
country in between.

A panelist on another panel expressed a similar idea:

What I probably would do as a first step is get everyone together to talk
about the process, figure out everyone's perspective—maybe see if there is
a set of better laid out criteria. If X, Y, Z is present we think we are going
to give it this kind of assessment. Then everyone goes off on their own and
then comes back together to compare.

On the other hand, *Monitoring International Labor Standards* doesn't provide such
guidelines. If such pre-assessment training is to be given, these guidelines would have to
be developed.

Another issue is whether a pre-assessment meeting would compromise the
independence of the individual assessments. In the evaluation, panelists were not told
who else was participating on the panel, because of concern that knowing the identities of
the others could possibly influence how they assessed compliance. As one panelist put it:

To do the individual assessments first and then to come together as a panel was
useful. If we had met first we might have had a "group psyche." If panelists met
first they might gauge where things are going. It is useful to come to own
judgments first.

Organization of the panel meeting itself

A number of aspects of the panel meeting—such how the discussion was
structured and efforts to resolve differences in assessments—have already been discussed
in earlier sections of this report and in Appendices E and G. In general, from the
perspective of the panelists the process went well and the discussion brought additional
viewpoints on the country's compliance with labor standards. In the post-panel
interviews, panelists were uniformly positive about their personal reactions to the
process. For example,

-I thought it was really excellent. They were good people. They both had serious
strengths that they brought to this exercise. It was very, very useful.

-It was interesting because more issues came up than I would have expected.
That was a good part of the panel discussion. There were things I thought
about during the discussion that I had not previously thought about. I had not
thought about the common law/civil law context or bias.

-I thought the whole process was quite well thought-through. This was not
a light, frivolous exercise. Spending 30 hours on a website is quite
intensive. And to spend 8 hours in a room comparing notes with two

> *other people—that was not a frivolous exercise. The conclusions that*
> *you are likely to draw from it will be substantial and substantiated*
> *because the process was rigorous.*

But, if the goal of the process is to reach a consensus on assessments of compliance for the indicators, none of the panels reached that outcome. In part, this may be a function of time constraints of a one-day meeting. The dinner meeting on the previous evening, in the words of one panelist, "made it a little easier for us to dive in the next day." Despite this, even without trying to reach a consensus on the indicators in dispute, their sheer number presents a serious time challenge. One panelist quipped that they could have spent a half day on each indicator.

One the other hand, a significant portion of the discussion time in the evaluation was devoted to discussing the task itself—use of WebMILS, the meaning and wording of individual indicators, and other aspects of the NAS-ILAB method. Presumably this would take up less time once those issues were resolved, although it is likely that any set of experts would have their own questions about the indicators, information sources, and assessments.

In addition to the discussion of the individual indicators, panels might be expected to reach a broader conclusion or set of conclusions about the level of compliance of the country as a whole. The NAS-ILAB matrix methodology doesn't provide directions for how to move from the assessment of individual indicators to more general conclusions for a country. As a step in this direction for the evaluation, panelists were asked for an overall rating of the level of compliance for each labor standard for the three broad areas: legal framework; government performance; and outcomes. As described above, Panel A did not get very far in this task. Panels B and C did reach some tentative conclusions (see table 8). In the course of this effort, panelists noted that the indicators address a wide range of issues with different levels of importance. One panelist, referring to the effort to create an overall rating, noted:

> *We did not have a discussion about prioritizing. We did not attempt to*
> *prioritize. But in our own heads we gave more attention to some. And it*
> *did come out in the summary at the end.*

This question of prioritizing the indicators came up in each of the panels. One of the panelists for Country C described the assessment task, with its total of sixty-nine indicators for the two labor standards (Freedom of Association and Discrimination), as "artificial" when it came to trying to draw an overall conclusion.

> *With this exercise you had to spend a lot of time searching for information*
> *that would not necessarily have been prioritized [if I were doing this*
> *myself]… Reporters in the State Department too will be picking and*
> *choosing from things. But for the purpose of this experiment obviously*
> *you wanted each and every indicator. In a sense that is artificial.*

Although there was general agreement that some indicators were more important than others, the issue of how one would set priorities among the indicators is complicated and raises a number of issues. One panelist talked about this complexity, referring to Colonial-period legislation that gave local African chiefs authority to conscript citizens into labor:

> *Prioritizing indicators is a tough question. My training tells me that violation of any provisions of an instrument is relevant. I would leave it to the evaluation, which of these is a killer problem. For example, if there is something in the legislation that allows chiefs to impose forced labor but in fact no authorization has been given for that for 20 years—then it is a theoretical rather than a national problem. I would want to leave it to the evaluation which is more important.*

> *Prioritizing is a theoretical question that is difficult to answer. For example, if you have an overall system of freedom of association that allows workers and employers to organize, that allows collective bargaining and even promotes it, but it does not allow for workers' organizations to use premises on the work place, it is a violation but not a big one. It is not insignificant but it is of a lesser order than a refusal to register trade unions.*

CONCLUSIONS AND RECOMMENDATIONS

The NAS-ILAB matrix methodology can be understood in terms of three elements. First, a set of indicators were developed for each of the five labor standards. These indicators explicate each of the labor standards and, in doing so, increase transparency in assessing compliance with international standards. This is an important advance and the panelists applauded this effort. As one put it, "Kudos to the effort . . . setting up a grouping of indicators on these labor standards is important, and useful, and largely successful." It was noted that the indicators and WebMILS could be very useful to other audiences—such as businesses considering international investments and the associated risk analysis.

In the course of assessing the indicators, the panelists identified a number of modifications that are documented in this report. Some are matters of wording and clarification. For others, however, questions were raised about whether the indicators reflected international standards or, rather, highlighted other issues that were relevant to labor rights in the country but were not properly characterized in terms of "level of compliance."

The second part of the NAS-ILAB approach was establishing WebMILS as a source of information for assessing the labor standard indicators. The evaluation analyzed the content and functioning of WebMILS. Panelists felt that WebMILS linked the user to relevant sources and, by including only sources that they felt were relatively objective and reliable, it was particularly helpful for a user who is less familiar with these sources. The question of maintenance and updating of the site was raised in response to encountering non-functioning links and changes in the labor standards environment that were not reflected in the database.

A number of issues emerged about the structure and function of WebMILS. In broad terms, it appears that the structure of WebMILS does not clearly differentiate between generic and country-specific information. Also, panelists felt that there should be a more direct link between the specific indicators and the relevant information sources. One possible alteration in the architecture of the site is displayed in figure 5, below.

This potential change would more clearly present the information about the WebMILS project, the labor standards and indicators, and the matrix approach in one area (the left branch, in figure 5). For those using WebMILS to assess compliance, the right branch would take them to the relevant indicators, with links to the specific sources for each indicator. As noted in the figure, most of these links would be the same for all countries, while others would go to the "national sources" and any other country-specific sources that were relevant for that indicator.

Figure 5
Possible Modification in WebMILS Architecture
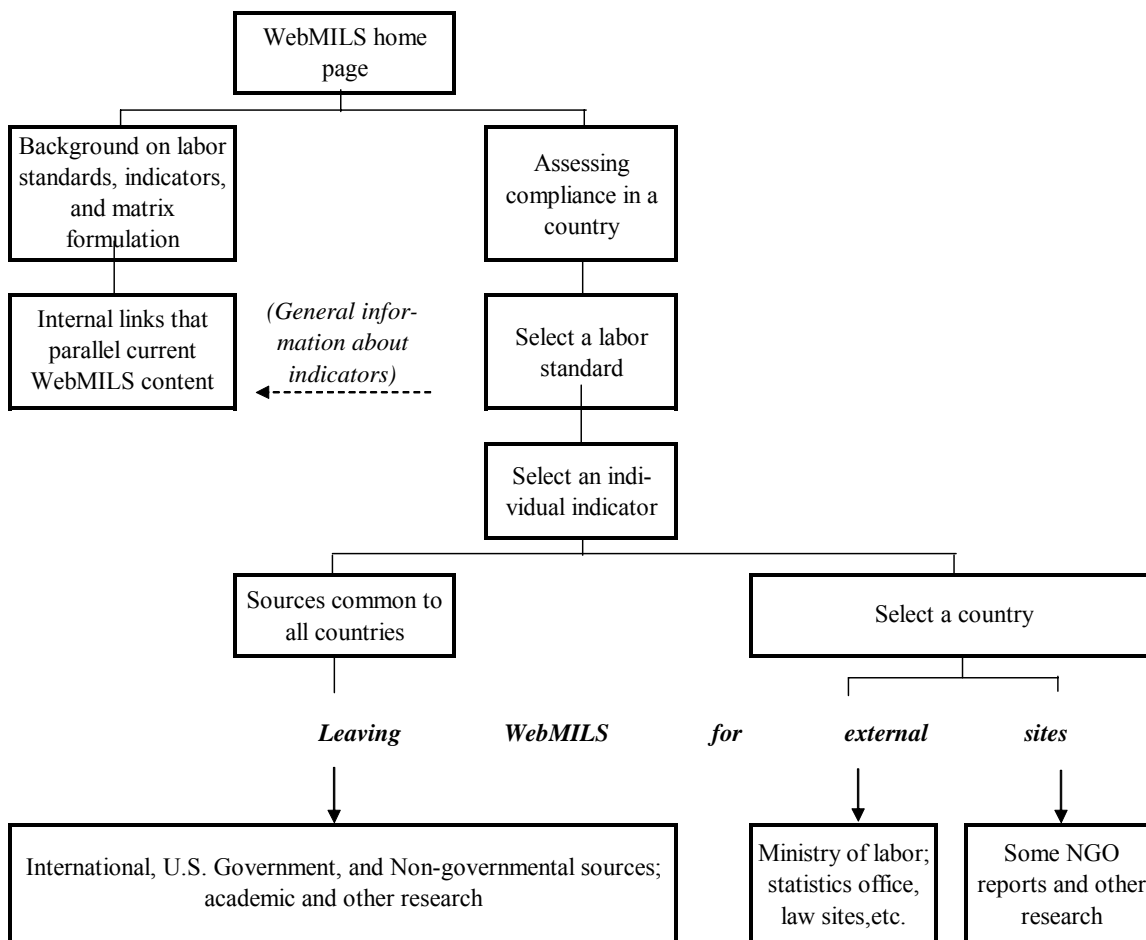
Panelists also raised an underlying concern with using WebMILS as the primary source of information: only using information available through the Internet. There is a great deal of unevenness in what is available in electronic form, particularly in developing countries. For many of the indicators, the panelists found little or no information and there was a recurring theme that what one needed was a local expert— someone "on the ground"—who could provide up-to-date and contextual information. The decision to not include media sources in WebMILS was cited as another limit on securing current information.

The third aspect of the NAS-ILAB method was the matrix layout itself—the characterization of each indicator in terms of three levels of compliance and directions of change. The panelists independently assessed the indicators and then met as a group to discuss their assessments and resolve differences in assessments. This overall process appeared to work well, although there was some suggestion that more training, and perhaps a pre-assessment meeting, might have been useful. Panelists identified a number

of problems in the formulation of the current set of indicators. They also raised the possibility that a scale with more options (e.g., five levels of compliance) might be preferable, although there was no consensus about this.

The matrix formulation is presented in *Monitoring International Labor Standards* as "one possible way to analyze the data" (p. 27). It provides an intuitive way to integrate the individual assessments of indicators for each labor standards. Unlike the hypothetical example in the text, in the evaluation there were relatively few indicators on which most of the panelists agreed in their independent assessments. Although most of these disagreements were relatively minor, the process of reaching a consensus was time-consuming and often unsuccessful.

The specification of indicators and the matrix formulations served its purpose of creating a way to make the process of assessing compliance less opaque. The question remains, however, of how one moves from assessment of individual indicators to a broader assessment of compliance with labor standards in a country. This is not addressed in the specification of the matrix method and led to discussions within the panels of differences in the importance of individual indicators and whether priorities could be established that furthered an integrated, overall assessment. As the authors of *Monitoring International Labor Standard* state, "The daunting policy challenge is to evaluate indicators of compliance and use them as a basis for policy judgments." (p.21)

Based on the experience of the evaluation, there appear to be several directions for next steps in developing the NAS-ILAB approach. These include:

1. Revision of the current set of indicators to reflect the problems identified by the panelists

   This revision could address current problems with wording and the lack of fit with the "answer set." It could also identify those indicators that are more appropriately considered reflections of the conditions in a country rather than "compliance," *per se*.

2. Possible reorganization of the indicators to facilitate discussion

   Some structural issues were identified in the course of the pilot test. For example, as noted in this report, "legal framework" indicators did not always clearly differentiate between the "letter of the law" and how the law is implemented. In such cases, it may be helpful to integrate discussion of the legal indicator(s) with the associated government performance indicators. Although it is important to identify the origin of problems (e.g., the laws versus their enforcement), discussion of compliance in a country might proceed in a more targeted way if it is structured in terms of areas of code compliance rather than always treating legal framework and government performance separately.

3. Modifying the process to include developing a contextual understanding of the country to provide an orientation to the panel assessing compliance

   The NAS-ILAB formulation alludes to contextual issues, such as the resources available to a country, as an important consideration in assessing compliance. Such factors, however, are not formally integrated into the

matrix approach, although there appears to be the expectation that it will impact the assessments of indicators. Some of the indicators—particularly those that have been identified as descriptive of the setting rather than reflecting questions of "compliance"—can be shifted from the matrix formulation to a protocol for developing a contextual description of the country, which can then provide background for panelists assessing compliance in that country.

4. Developing guidance for moving from the assessment of individual indicators to an overall introduction

The effort in the pilot test to make an overall assessment of the legal framework, government performance, and outcomes continually came up against the question of how one pulls together the individual assessments of the indicators. This was often articulated in terms of establishing priorities reflecting the relative importance of indicators. This, however, is a thorny problem. Although it may not be productive to try to establish fixed priorities among the indicators, developing some guidance for how to move from the specifics to an overall assessment would be a significant advance in the use of the indicators.

Appendix A
Orientation Letter to Panelists

INSTITUTE OF LABOR AND INDUSTRIAL RELATIONS
UNIVERSITY OF MICHIGAN
1111 EAST CATHERINE STREET
ANN ARBOR, MICHIGAN 48109-2054
(734) 763-3116   WWW.ILIR.UMICH.EDU

   To: \<panelist name\>
        Panelist for \<Country A\>
 From: Larry Root and Ada Verloren
   Re: Instructions to panelists for the evaluation of the ILAB matrix
        approach to monitoring international labor standards
 Date: February 21, 2008

        Thank you for agreeing to serve as an expert panelist for the assessment of the
Department of Labor's matrix approach to assessing the status of labor standards. As
we've discussed, the National Research Council of the National Academy of Sciences
worked with the Department of Labor's Bureau of International Labor Affairs (ILAB) to
create a method for monitoring the status of a country's adherence to international labor
standards. The University of Michigan is working with DoL-ILAB to test this method
for several countries. We are very grateful for your help in this effort.

        The proposed method for assessing a country's labor standards is described in a
book by the National Research Council, *Monitoring International Labor Standards:
Techniques and Sources of Information* (copy enclosed). Briefly, there are five key labor
standards addressed:

   1. Freedom of association and the right to collective bargaining
   2. Forced or compulsory labor
   3. Child labor
   4. Discrimination
   5. Acceptable conditions of work

For each of the labor standards, indicators have been defined that reflect: (a) the legal
framework related to the labor standard within a country; (b) government performance in
implementing the labor standard (both in terms of effort and effectiveness); and (c)
outcome measures related to the labor standard. These indicators were designed with a
two-fold purpose: to assess the level of compliance in the country (some problems, more
extensive problems, or severe problems) and identify whether the situation is changing
(improving, steady state, or worsening).

        The proposed method calls for assessors to independently evaluate the indicators
associated with a given labor standard as described above.   The findings are organized
within a matrix that has *level of compliance* on one axis and *direction of change* on the
other (see pages 30-31 of *Monitoring International Labor Standards* for an example).

Finally, assessors come together as a panel to discuss their individual assessments, and attempt to reach a consensus evaluation.

The National Research Council and Department of Labor also identified and organized sources of data that could be used to assess the labor standards. These are organized in a DoL-ILAB website called WebMILS, which is located at:

<http://www.dol.gov/ilab/webmils/>

In addition to providing links to key documents for assessing labor standards, WebMILS also contains background information on the labor standards and ILAB matrix approach.

It is expected that the panelists will use WebMILS as the primary tool for locating the information needed to make their assessments.[12] Selecting "Search WebMILS" on the home page brings one to a listing of countries. On this page, the user also has the option of searching by the type of information source (e.g., international organizations such as the ILO or the Asian Development Bank). Selecting a specific country leads to a selection of the broad areas of labor standards. Clicking on one of these areas, takes one to a listing of the individual indicators, followed by sources of information about these indicators. Appendix A-1 provides some screen captures from WebMILS. Although not the primary focus of the evaluation, your assessment of the use of WebMILS for addressing the indicators will form part of the study.

Evaluation plan.

The research design calls for three panelists to assess two labor standards for a country. This assessment is a two-stage process: panelists first provide an independent assessment of compliance and progress for each indicator. The panelists then come together as a group to discuss their individual assessments and seek to reach a consensus evaluation.

> *Stage 1 - Independent Assessment:* Panelists independently consider all indicators for a given labor standard. For each indicator, the panelist provides a rating for the level of country-level compliance and indicates whether there is a trend toward improvement or worsening. Panelists then submit their ratings by means of an online data input system to the project staff at the University of Michigan, who will organize panelist findings into matrices that have *level of compliance* on one axis and *direction of change* on the other. Project staff will use these matrices to compare the ratings across panelists and identify differences in assessments between members of a given panel.

---

[12] One goal of the evaluation is to assess how individual experts make assessments based on the same set of data/information. If an individual expert on a panel identifies relevant sources for assessment that are not a part of WebMILS, he/she is expected to bring this to the attention of the project staff, who will make it available to the other members of the panel.

Stage 2 - Panel Deliberations: After the individual assessments have been completed, the panelists will convene at University of Michigan for a one-day meeting. At that meeting, differences in individual assessments will be discussed and the panel will attempt to come to a consensus for each indicator on level of compliance and progress for that country. The panelists will also be asked to discuss their experience in making their assessments and ways in which the process could be improved.

Our analysis for the evaluation of the matrix methodology will use quantitative measures of the extent of agreement among panelists in initial independent assessments as well as the extent of agreement following the group discussion. The process/discussion will be observed and recorded in order to identify key points as well as patterns, such as associations between type of indicators and level of agreement. We also plan to interview each panelist by phone before they convene in Ann Arbor and after that meeting.

This evaluation process will be carried out for three countries, with a separate team of panelists for each country. Because this evaluation is intended to evaluate the proposed method of assessment, the identity of the three countries will be treated as confidential for the purposes of any reports, publications and/or presentations.

Expectations for panelists. Each of the panels will be responsible for assessing the indicators associated with two labor standards: (1) freedom of association and collective bargaining and (2) either forced labor, discrimination, or acceptable conditions of work. The assignment of the second core area will be made by the project staff. As a panelist for <Country A>, you will be responsible for assessing the indicators for:

-Freedom of Association and Collective Bargaining and
-Acceptable Conditions of Work

As described above, you will be expected to assess the relevant indicators independently and then to participate in a meeting in Ann Arbor, where you will engage in panel deliberations in an attempt to reach a consensus assessment, and discuss their experience.

Enclosed is a packet of forms with one page for each of the indicators for Freedom of Association and Collective Bargaining and Acceptable Conditions of Work. Panelists will submit their assessments via an online version of the form located on the website of the University's Institute of Labor and Industrial Relations:

www.ilir.umich.edu/ILAB/Matrix/Panelist/Input/

Your Login Name is: <login name>
Your Pass Phrase is: <password>

The hard copy forms that are enclosed are intended to provide you with a working document on which you can make notes. These notes can also serve as reminders for the group discussion and offer a backup in case of technical problems with the online system.

Assessment of the indicators represents a significant work effort on the part of panelists. Some indicators, such as whether or not a country has ratified a particular ILO Convention are straightforward and easy to determine. Others, however, are more difficult to assess. In some cases, a panelist may feel that there is not enough information available to make an assessment. We do, however, expect that you will make an effort to come to an assessment of each of the indicators in the two core areas being addressed by the panel. If there are indicators for which no assessment is made, we request that you use the "comments" section to explain why you have not made an assessment.

Part of the overall evaluation of the matrix is estimating the time necessary to make the assessments. Based on a pilot test, we estimate that panelists should be prepared to devote up to 40 hours to this assessment task. In practice, you may find that it takes less than 40 hours. We request that you keep track of the time it takes you to do the assessments. You are not expected to devote more than 40 hours to the assessment even if there are still indicators for which no conclusion on an assessment has been reached.

As discussed above, each panel will convene at the University of Michigan to deliberate in an attempt to reach a consensus assessment and to discuss their experience with the assessment. We plan to schedule these meetings for the first two weeks in May and will be contacting you imminently to arrange the specific dates the meeting for your 3-person country panel. We expect that the panel will spend one full day during which time they will discuss and seek to resolve differences in their independent assessments. In order for the staff to prepare of the meeting, panelists will have to complete their assessments at least five working days before they are scheduled to meet as a panel in Ann Arbor. In addition to the discussion of the specific indicators and assessments, the panelists are expected to share their perspectives on the process of making the assessments and ways in which this could be improved. Project staff will observe and record this process.

Given the need to coordinate schedules and prepare for each of the meetings in Ann Arbor, it is essential that panelists adhere to their agreed-upon schedule for submission of their assessments. An honorarium of $4,000 will be provided to each panelist in recognition of his/her assistance with this project.

We have an excellent group of expert panelists and we are grateful for your willingness to participate. If you have any questions, please don't hesitate to contact us.

Enclosed:
      Monitoring International Labor Standards (book)
      Panelist Assessment Work Sheets

Appendix A-1
WebMILS

Home screen

Welcome to WebMILS - Microsoft Internet Explorer

Back | Search | Favorites

File Edit View Favorites Tools Help   Address http://www.dol.gov/ilab/webmils/

Links Google CTools Mirlyn Wolverine Access umich UM Directory AA Movies AA Library CNN FlightDelays ILIR Wikipedia SSW-Intranet

U.S. Department of Labor
Bureau of International Labor Affairs

www.dol.gov/ilab   Search / A to Z Index

Find It!: By Topic | By Audience | By Top 20 Requested Items | By Form | By Organization | By Location

January 15, 2008   DOL Home > ILAB > WebMILS

webMILS
MONITORING INTERNATIONAL LABOR STANDARDS

Welcome to WebMILS
En Espanol

WebMILS is an electronic database containing information relating to national compliance
with international labor standards. Users may find the following types of information:

- General information on International Labor Standards;
- Guidance on assessing national-level compliance; and
- Links to a wide range of information sources, which may prove useful in conducting
  specific country assessments.

To access general information on International Labor Standards or to read the guidance
of the National Academies' Committee on Monitoring International Labor Standards
(CMILS) on Assessing Compliance, users should follow those links in the navigation bar on
the right-hand side of the screen. These pages-organized by labor standard-offer
descriptions of indicators that CMILS has recommended for users to consider in order to
exercise "due diligence" in conducting assessments of national-level compliance with
internationally recognized labor standards.

Accessing Information Sources in WebMILS: Conducting Searches

Searches can be conducted from the WebMILS home page by clicking "Search webMILS"
or through the "Sources" link in the blue navigation bar. The search is commenced by
selecting a country and standard or topic from the dropdown menus. In addition to core
labor standards and "acceptable conditions of work," users can choose to search for
human capital or general country information.

WebMILS provides information on the international labor standards
included in the 1998 ILO Declaration of Fundamental Principles and
Rights at Work, with the addition of 'acceptable conditions of work.'
These are:

- Freedom of association and the effective recognition of the right

ILAB Home

WebMILS Home

The WebMILS Project

Search WebMILS

International Labor Standards

Assessing Compliance

Matrix Assessment Tool

Feedback & FAQs

General DoL information

To search for country-specific information

Descriptions of the standards

http://www.dol.gov/ilab/webmils/index.htm

start   DOL-ILAB-MatrixPro... Mulberry (Connected) Message Search   2 Microsoft Office ... Welcome to We

Clicking "Search WebMILS"

Brings up country list for country-specific sources

For example, selecting Albania leads to a screen listing the broad areas

Selecting Freedom of Association and Collective Bargaining, goes to:

Scrolling down

Selecting an individual indicator here takes one to the description of the indicator

Beginning of "Sources" section

[the web page continues…]

Appendix B
Cover Sheet and Example of a Hard-Copy Assessment Form

*[The bound copy for this panelists included 38 assessment forms for Freedom of
Association and Collective Bargaining and 37 for Acceptable Conditions of Work]*

INSTITUTE OF LABOR AND INDUSTRIAL RELATIONS
UNIVERSITY OF MICHIGAN
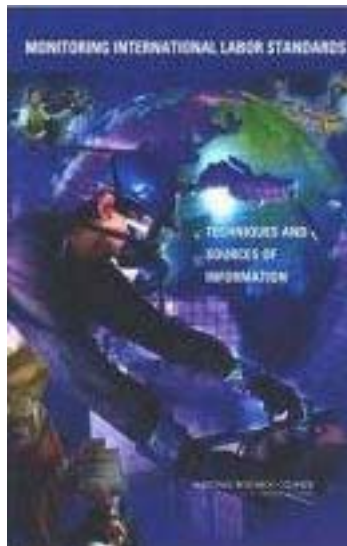
Evaluating the ILAB Matrix Approach to Monitoring International Labor Standards

Panelist Forms for Assessment of Indicators

<Country A>

Labor Standards:

**Freedom of Association
 and Collective Bargaining**

Acceptable Conditions of Work

<u>Panelist Form for Assessment of Indicators</u>
[example of panelist form]

Area: *Freedom of Association*

Indicator: *A-3. the extent to which there are legal restrictions on the ability of
certain categories of workers to organize (such as civil servants, teachers,
or workers without contract of employment), and numbers of workers in
each category*

<u>Assessment</u>:

Level of problems:

□ Some problems   □ More extensive problems   □ Severe problems   □ Not determined

Direction of Change:

□ Improving   □ Steady state   □ Worsening   □ Not determined

<u>Basis of assessment</u>:

WebMILS data sources: _____

_____

_____

Did you use a source <u>not</u> in WebMILS   □ Yes   □ No

  If "yes," please identify source: _____

  _____

Confidence in assessment: □ low confidence—highly speculative
                              □ moderate confidence—some objective evidence
                              □ confident—clear indications, strong objective evidence

Comments: _____

_____

_____

_____

Appendix C
Examples of Online Data Input

Assessor Form for Assessment of Indicators
(demo)

Assessor for Russia
Freedom of Association and Collective Bargaining
Select an Indicator

Thursday 14 August 2008 at 10:29AM

Logout

Back to Select a Labor Standard

When you select one of the **Freedom of Association and Collective Bargaining Indicators** below, the system will display the data entry form for that Indicator.

Completion status

A-1. whether the country has ratified ILO Convention No. 87
☐ Level of compliance
☐ Direction of change

A-2. whether there are legal provisions that entitle workers or employers to establish and join organizations of their own choosing without previous authorization
☐ Level of compliance
☐ Direction of change

A-3. the extent to which there are legal restrictions on the ability of certain categories of workers to organize (such as civil servants, teachers, or workers without contracts of employment), and numbers of workers in each such category
☐ Level of compliance
☐ Direction of change

A-4. whether there are certain sectors where there is no right to organize (such as in export processing zones or in other tradable sectors or in agricultural and informal sectors)
☐ Level of compliance
☐ Direction of change

A-5. whether there are other forms of distinction or discrimination in right to organize, such as race, nationality, sex, opinion, political affiliation, or citizenship (for example, excluding legal immigrants)
☐ Level of compliance
☐ Direction of change

A-6. whether there are legal provisions that permit the government to interfere in freedom of association by workers or employers
☐ Level of compliance
☐ Direction of change

A-7. whether there are legal restrictions on political activities of unions or employers' organizations, either by
☐ Level of compliance
☐ Direction of change

Appendix D

Questions for Pre-Panel Telephone Interview

Panelist: _____ Date: _____

A. The assessment process generally

1. About how much time did you spend making the assessments?

_____hours

Comments: _____

_____

_____

2. In general, did you feel that the goal for doing the assessments was clear?

Very clear	Clear	Neutral	Not clear	Very unclear
☐	☐	☐	☐	☐

Comments: _____

_____

_____

3. In general, did you feel that the criteria for assessing individual indicators were clear?

Very clear	Clear	Neutral	Not clear	Very unclear
☐	☐	☐	☐	☐

Comments: _____

_____

_____

4. How would you characterize differences in assessing the three types of indicators:
   legal framework, government performance, and overall outcomes?

<u>Legal framework</u>

Easier to Assess	Same as other two	More difficult to assess
☐	☐	☐

<u>Government performance</u>

Easier to Assess	Same as other two	More difficult to assess
☐	☐	☐

<u>Overall outcomes</u>

Easier to Assess	Same as other two	More difficult to assess
☐	☐	☐

Comments: _____

_____

_____

5. How would you characterize differences in assessing the level of compliance
   compared to direction of change?

<u>Level of compliance (compared to direction of change)</u>

Easier to Assess	About the Same	More difficult to assess
☐	☐	☐

Comments: _____

_____

_____

6. Were the areas in which you thought there were particular gaps in the information available?

     □ Yes   □ No

Comments: _____

_____

_____

7. Were there information sources that you found particularly helpful?

     □ Yes   □ No

Comments: _____

_____

_____

8. Did you feel that there were biases in sources (both quantitative and qualitative)?

     □ Yes   □ No

     If "yes," were they (check all that apply)
          □ Complaint-driven sources
          □ Reports skewed for political reasons
          □ Based on unrepresentative sample
          □ Other _____

Comments/Examples: _____

_____

_____

9. Did you feel that some of the information sources were flawed to the extent that they were not useful?

  □ Yes   □ No

Examples/Comments: _____

_____

_____

B. Use of WebMILS

10. What is your overall impression of the usefulness of WebMILS for making these assessments?

  Very useful   Useful   Neutral   Not useful
      □   □   □   □

Comments: _____

_____

_____

11. How difficult was it to locate information on WebMILS?

  Very easy   Somewhat easy   Neutral   somewhat difficult   Very difficult
     □   □   □   □   □

Comments: _____

_____

_____

12. Did the data seem to be relatively up-to-date?

      Up-to-date   Neutral   Not up-to-date

        ☐   ☐   ☐

Comments: _____

_____

_____

13 Did you feel you had adequate information about the data sources to judge their reliability?

      ☐ Yes   ☐ No

Comments: _____

_____

_____

14. Were some relevant data only available in non-English language sources?

      ☐ Yes   ☐ No

Comments: _____

_____

_____

15. Are you aware of appropriate data that are not available in WebMILS?

      ☐ Yes   ☐ No

Comments: _____

_____

_____

C. Suggestions for improvements

16. Do you have any suggestions for changes that would improve the assessment process?

    □ Yes   □ No

Comments: _____

_____

_____

17. Do you have any suggestions for changes in WebMILS?

    □ Yes   □ No

Comments: _____

_____

_____

18. Are there other reactions that you have to the assessment experience?

    □ Yes   □ No

Comments: _____

_____

_____

19. Do you have any questions/expectations for the panelist group meeting?

    □ Yes   □ No

Comments: _____

_____

_____

Appendix E
Agenda for Panel Meeting

Monitoring International Labor Standards

Meeting of Panelists for <Country C>

Labor Standards:
    Freedom of Association and Collective Bargaining
    Discrimination and Equality

Panelists: University of Michigan Personnel
  <name of panelist a>   Larry Root
  <name of panelist b>   Ada Verloren
  <name of panelist c>

Dinner meeting (Wednesday, May 21, 2008)

1. Introductions of project and individuals participating

2. Overview of tasks for main session

    a. Discussion of individual indicators, differences in assessments and rationales, and possible re-assessment in light of this discussion

    b. Discussion of the Matrix approach itself as a guide for assessing labor conditions in a country

Main session (Thursday, May 22, 2008)

8:30 – 9:00   Differences in assessments based on answer set or interpretation
Identification/discussion of indicators for which the panelists have suggested that the answer set (level of compliance: some problems, more extensive problems, severe problems) does not provide appropriate responses or the interpretation of the indicator is ambiguous

Discussion of strategies used in such situations

9:00-2:00   Discussion of individual indicators
- Differences in assessments among the panelists
- Rationale and evidence for assessments
- Attempt to come to common assessment when differences exist

12:00-1:00   Lunch

1:00-2:00   (continuation of discussion of individual indicators)

2:00-4:00   Discussion of the matrix approach for assessing labor conditions
- Level of compliance and direction of change
- Range of indicators—indicators to add/subtract
- WebMILS as a source and additional data sources
- Additional/alternative approaches for assessing compliance
- Process: individual assessments followed by group discussion
- Overall assessment for <Country C>

Appendix F

Post-Panel Questions for Telephone Interview

Panelist: _____ Date: _____

1. Did your feel that the meeting addressed what you had expected?

    □ Yes   □ No

Comments: _____

_____

_____

2. Was the organization of time appropriate?

    □ Yes   □ No

Comments: _____

_____

_____

3. Did you feel that the other panelists approached the assessment task in a similar way to
    how you approached it?

    □ Yes   □ No

Comments: _____

_____

_____

4. Were there issues associated with the assessment exercise that you feel were not
   adequately addressed?

   □ Yes   □ No

Comments: _____

_____

_____

5. Is there anything further that occurred to you with regard to the assessment exercise
   (e.g., indicators to add/subtract? ways of prioritizing indicators? additional data
   sources?)

   □ Yes   □ No

Comments: _____

_____

_____

6. Did you find that discussion/interaction with other panelists useful?

   □ Yes   □ No

Comments: _____

_____

_____

7. Do you have any further thoughts about your overall assessment of the matrix approach to monitoring international labor standards?

□ Yes   □ No

Comments: _____

_____

_____

Appendix G
Notes from Panel Discussions

Each panel began with an evening meeting followed by a full day panel meeting. The work plan for each panel meeting included discussion of the individual indicators, with a focus on those indicators for which there was greater disagreement based on the individual assessments undertaken prior to the meeting. Discussion of indicators was followed by discussion of the ILAB-NAS matrix approach generally and WebMILS as a key information source.

Country A (May 11-12, 2008)

        Labor Standards: Freedom of Association and Collective Bargaining
                                              Acceptable Conditions of Work

        Panelists:

- Professor of industrial relations
- Union-based analyst and program director with focus on international labor conditions
- Professor of law, with focus on international and comparative law

This first panel meeting followed the agenda and provided the model which was largely replicated in the two subsequent panel meetings. At the dinner meeting on the evening before the full-day session, we discussed the project, the backgrounds of the individual panelists, and the plans for the main session the following day. It turned out that two of the panelists had previously met at a small labor rights conference several years earlier. One of those two had also worked on projects with labor-based organization where the third panelist works.

The main session the following day began at 8:30 a.m. In preparation for the session, the University of Michigan research staff had reviewed the assessments of each of the panelists, grouping the indicators into three groups by the extent of agreement among the panelists on their assessments: high, medium, or low agreement. "Low" agreement generally included those indicators for which each of the three panelists rated the indicator's level of compliance differently. In some cases, there might have been two different ratings plus differences in rating the direction of change. "High" agreement usually meant that the panelists rated the level of compliance the same.

The pre-panel review of individual ratings also noted panelists' comments that raised questions for discussion about individual indicators, the assessment process itself, and WebMILS as an information source. For example, the research staff noted that panelists identified indicators that did not fit with the answer set—that is, they called for a "yes/no" response rather than "some problems, more extensive problems, or severe problems." Questions about whether or not a particular ILO convention had been ratified were a clear instance of this issue. Such questions had been called to the attention of the researchers by panelists in our pre-meeting phone interviews. The individual comments also identified indicators which raised questions about interpretation. For example, comments suggested that panelists did not know how to interpret an indicator that asked about number of strikes and, if there were many strikes or strikes of long duration, what that might indicate about freedom of association

Given the time constraints, the research staff estimated that trying to discuss each of the 75 indicators (38 for Freedom of Association and Collective Bargaining; 37 for Acceptable Conditions of Work) would result in just a few minutes discussion of each. Instead, the discussion was structured as follows:

-Initial discussion of the indicators for which the answer set (level of compliance: some problems, more extensive problems, severe problems) did not provide appropriate responses or the interpretation of the indicator is ambiguous. Rather than focus on possible differences in assessments for these indicators, the panelists would discuss the problems of using the answer set and how the individual panelists addressed this situation

-Discussion of individual indicators for which there was the most disagreement among the panelists (agreement was "low"). Nine of the 75 indicators were rated as having low agreement, all but one of which were part of Freedom of Association and Collective Bargaining. The discussion of "low agreement" indicators was followed by discussion of indicators for which there was some disagreement ("medium" agreement).

Following the discussion of each of these indicators, the panelists re-rated the indicator by hand. These assessments were collected at the end of the session.

-General discussion of the matrix approach as a guide for assessing labor conditions and WebMILS as a data source for this assessment.

The panel discussion, while following the overall format, was free-flowing, so that general points about the matrix framework and WebMILS often were interspersed with discussions of the indicators.

A number of issues/points came up repeatedly in the discussion, such as:

1. General concern that data sources for many areas were lacking sufficient information to make an informed assessment, for example
    a. There is little information about many aspects of government performance; panelists often mentioned that one might need "people on the ground" locally to get this kind of information
    b. The government data generally does not address the informal sector or unregistered workers—a major portion of the workforce in Country A
    c. WebMILS does not seem to have much useful information on many areas of government performance (e.g., FoA-CB, B10-13)
    d. The panelists felt that more guidance would be useful to help an assessor interpret reports from various sources

2. Need for more guidance on criteria for making assessments of level of compliance; for example, should the assessor be comparing the situation in a country to some ideal standard or to other countries?

3. Mismatch between the answer set and the response that the panelist felt was appropriate for an indicator, for example
    a. Some indicators are clearly a "yes" or "no," rather than "some problems, more extensive problems, severe problems, e.g., all of the ratification questions)
    b. For some indicators, such as questions about legislation, the wording suggests a "yes" or "no" questions (such as "Whether _____ is protected in law"); different wording might make it more of an assessment question (such as "To what extent is _____ protected by law")
    c. With the answer set, there isn't an option to say that there is "no problem" with regard to an indicator
    d. The panelists felt that having a 5-level answer set might have provided a better set of options, providing a "no problem" option and some distinction between "more extensive problems" and "severe problems"

4. In assessing whether compliance on an indicator is changing, need to have some guidance on the time frame for making such an assessment.

    In practice, two of the panelists used the year 2000 as a rough beginning point for assessing change. The third panelist used the last 2-3 years.

5. Media coverage could be a useful additional source of information.

    Media coverage is specifically discussed in the *Monitoring International Labor Standards* book as not intended as a primary source because of difficulties in assessing accuracy. The panelists felt that media sources could be a useful addition to WebMILS (e.g., http://labourstart.org/ provides a searchable listing of international news reports relevant to trade unions)

6. Court cases, such as from the European Court of Justice, the European Court of Human Rights, and the inter-American system, could provide useful information, especially because the legal briefs often provide a thorough view of the context from the point of view of both the plaintiff and defendant.

7. Some questions appear to be addressing two or more different issues. For example, FOA-CB B-1 includes job actions taken as retaliation along with murder—qualitatively different actions.

8. There aren't any indicators that focus on "civil liberties" (e.g., right of assembly, freedom of speech)—important part of labor rights environment

9. Under Acceptable Conditions of Work, unclear how to interpret/rate questions about the minimum wage/average wage levels. Not clear what the criteria should be.

<u>Country B</u> (May 14-15, 2008)

        Labor Standards: Freedom of Association and Collective Bargaining
                                    Forced or Compulsory Labor

        Panelists:

- Independent consultant on corporate social responsibility
- Independent consultant on labor conditions and labor rights
- Former ILO official with extensive experience in employment and human rights

        The County B panel went as planned. A representative from ILAB attended the panel meeting as an observer. There was somewhat less variation in the individual ratings in the Country B panel compared with the Country A panel. For example, there were only six indicators that we characterized as having "low agreement" compared with nine for Country A. The discussion of individual indicators therefore more quickly turned to those for there was "medium" agreement.

        The issue of answer-set problems for "yes-no" indicators was noted. It was suggested that it is not appropriate to talk about "compliance" for indicators for which there is not a formal obligation in international conventions. For example, according to the panelist with extensive ILO experience, indicator B-7 under Forced or Compulsory Labor (FCL) concerns receipt of technical assistance. Because there is no international obligation, "level of compliance" is not formally applicable.

Some of the points raised in discussion:[13]

1. Indicators sometimes included more than one issue. This could raise problems of interpretation/application. For example, in FoA A-2—"whether there are legal provisions that entitle workers or employers to establish and join organizations of their own choosing without previous authorization"— what is meant by the addition of the phrase "without prior authorization?" This seems to make the indicator less clear.

2. With regard to ratification questions, there are countries that have not ratified ILO #87 but they have ratified the two covenants associated with it.

3. From the perspective of businesses, "risk assessment" is a central concern. The matrix approach would be useful for three key goals: assessing the risk of doing business in a country; some measure of whether a country's labor practices are in compliance; and risks associated with potential violations of treaty obligations. From this perspective, the question about specific indicators was whether they helped in assessing the risk of doing business in a country. It was noted that businesses typically know little of the ILO and its conventions.

        During the panel, there was reference to the Equator Principles, a set of social and environmental standards created in 2003 by a number of

---

[13] In these notes on the individual panels, some points noted with regard to an earlier panel are not repeated. General points, however, may be reiterated when additional examples are provided.

international financial institutions (IFIs) such as the World Bank, to guide international investment decisions. The Equator Principles are intended to help financial institutions "to be able to better assess, mitigate, document and monitor the credit risk and reputation risk associated with financing development projects." [from website: http://www.equator-principles.com/faq.shtml] The Equator Principles are not very specific and the labor elements appear to draw heavily upon the International Finance Corporation (IFC), whose goals statement is summarized as "Reducing Poverty, Improving Lives." [http://www.ifc.org/] These efforts can be contrasted with the approach used in the "Doing Business" ratings, in which more extensive labor protections in a country tends to worsen their rating as a business environment. For example, in the "Doing Business" ratings, protections for workers can translate to scores indicating greater "rigidity of employment."

4. It might be helpful to have some general "context" indicators or a narrative way of setting the context for a country. [Note: A similar point was raised in the Country C panel, when it was suggested that the "outcome" indicators might be more useful as information at the beginning of an assessment process—setting the context for looking at legal framework and government performance.]

5. In assessing the legal framework, it is sometimes unclear whether a law was actually in force/implemented or not.

6. The 3-level response set for compliance needs an "is fine" or "no problems" category. It would help in the assessment to have more gradations than the current three.

7. One of the panelists raised the issue of the indicators having a "common law bias" in the sense that they focus on "rights" rather than "remedies" (which is more associated with civil law systems). An example might be FoA A-2 in which the law may provide a right to join a union but that right may be meaningless if the remedy—the penalty when that right is violated—is trivial.

8. In assessing the legal framework, there were questions about whether regulations or "decrees" should be considered.

9. An additional indicator might be whether or not the government collects and publishes relevant data on labor rights outcomes—where there is no information, one can make a judgment about the government's lack of interest or decision not to report honestly.

10. There was support for using newspaper data bases, if they exist—with the idea that media reports can provide "facts" (did a strike occur) but interpretations may reflect specific orientations. From a business perspective, the Business and Human Rights website (UK) was also mentioned as a useful source of news reports (http://www.business-humanrights.org/)

11. Transparency International (http://www.transparency.org/) was also mentioned as a possible source for WebMILS. It focuses on civil society with an emphasis on corruption.

12. Outcome measures, such as union density, are value neutral, purely factual— This may be relevant in describing a country, but "compliance" categories are not applicable.

13. With regard to the role of trafficking in forced or compulsory labor, reference was made to the UN's Palermo Protocol as possibly relevant for assessments.

14. Additional or alternative indicators for forced or compulsory labor:
    -is the government taking steps to prevent forced labor
    -extent of migrations, forced or not
    -extent of illegal migration
    -presence or absence of labor recruiters
    -extent of government regulation of labor recruiters, including regulations/limits on money or wages that can be withheld by recruiters
    -governmental role in sending/receiving remittances
      [a panelist suggested consulting a 2005 research report on exploitation of migrant workers by Verite, "Protecting Overseas Workers" (http://www.verite.org/)]

15. In some instances, the group discussion appeared to lessen individual panelists' confidence in their assessments

16. The discussion revealed some overlap between "child labor" and forced or compulsory labor—e.g., a panelist mentioned the lack of attention to child soldiers in the indicators for forced labor, however this issue is included under child labor

17. With regard to WebMILS, panelists suggested that a search tool would be very helpful, along the lines characteristic of Google searches.

18. In the discussion of overall conclusions about the country, it was suggested that some form of prioritizing or a "weighting" system would be helpful to recognize differences in the practical importance of indicators.

19 There was a generally positive response to the matrix approach as a tool for clarifying and assessing the labor standards.

(May 21-22, 2008)

        Labor Standards: Freedom of Association and Collective Bargaining
                                    Discrimination

        Panelists:

- Business school professor teaching U.S. and international labor law
- Executive for the U.S. office of a major international union organization
- Senior program advisor with several labor rights NGOs

The meeting began with a somewhat abbreviated discussion of the "either/or" questions, followed by a discussion of individual indicators, the matrix approach, and WebMILS. There was somewhat more agreement between individual panelists on their initial assessments of indicators. With fewer instances of "low" agreement, the discussion of individual indicators addressed more that had "medium" agreement.

The discussion of individual indicators was slightly varied in two ways based on the experience of the previous panel. First, at the beginning of the main session, we provided each panelist a printout with the ratings of all three for each indicator (not including their comments). In earlier discussions, panelists orally indicated how they rated the indicator at the beginning of the discussion of that indicator; in this iteration, panelists had the ratings in writing before they began the discussion of individual indicators. This change appeared to quickly provide an initial framework for discussing an indicator. It did not appear to change the nature of the discussion or outcome from that in the prior two panels.

The second variation was tried to test whether the panelists could reach a consensus on their rating of individual indicators. The plan was to have the each panelist re-assess the indicator in writing following discussion. Once this individual re-assessment was completed, the panelists would seek to agree on a group assessment. The intent of this approach was to be able to analyze the individual's re-assessment (as with the other panels), but also to observe the more public effort to achieve consensus. It quickly became clear that this approach was not productive and the results were revealing about the dynamics of trying to achieve a consensus on individual indicators when differences are not large and the evidence is incomplete or ambiguous. In practice, the effort to reach a consensus on individual indicators was very time-consuming, with extensive discussion about small differences. Panelists seemed to feel social pressure to make small modifications based on little firm data. In the interests of time and maintaining a positive environment for the panel discussion, this attempt to reach an oral consensus on individual indicators was dropped and panel followed the pattern used in the previous two meetings, focusing on changes in individual re-assessments as an indication of the impact of the group discussion.

Some points raised in the discussion:

      1. Initial discussion of the indicators of the legal framework suggested some differences in interpretation. Some questions seemed to suggest that the indicator was intended to include the application of the law rather than just the "letter of the law." For example, FoA A-14 ("whether the principle of a strike as a means of action of organizations is generally recognized")

suggested to some that the indicator was including how the law is implemented. If the indicator was worded "whether the principle of a strike as a means of action of organizations is guaranteed," it would be more clearly limited to the law itself. Similarly, Discrimination A-9 seems to emphasize how a law in implemented:

> whether there are laws that make nondiscrimination in employment *meaningful in practice* with equal access to education, training, vocational guidance, and placement services; maternity protection; and parental leave [emphasis added].

FoA A-20 also seems to ask about the effectiveness of the law:

> "assessment of extent to which "right to work" laws or other "free rider" provisions undermine the ability of workers to organize and the extent to which laws require workers to join a given trade union as a condition of employment or that new workers be hired through a given trade union"

How should "practices" influence an assessment of the effectiveness of laws. If strikers are stopped from picketing because they are "disturbing the peace" (not a part of labor law), is this evidence that "the government can limit picketing…" (FoA indicator A-16)?

The panel discussion raised the question about whether regulations and court decisions and precedents should be considered part of the legal framework. It was noted that in a "common-law" system, court cases are a key part of understanding the law. Court interpretations are also relevant for understanding the operation of civil law regimes. The question was raised about whether a "right" that is enunciated is really protected if the "remedy," in terms of responses to violating this right, fails to ensure the respecting of the right. For example, FoA A-8 refers to "whether national laws protect workers from discrimination if they join a union or participate in union activities." Assessing this might require knowing what the punishment is for violating this right.

In discussing another aspect of the complexity of assessing the legal system, FoA A3, it was noted that the law protecting civil servants' rights in Country C was conditioned on the passage of a separate law about civil servants that was never passed.

2. The group talked about the problem of not having a clear time-frame for basing decisions about direction of change. They noted that Country C had installed a democratic government a number of years ago and there have been many improvements in the labor standards environment since that change. Whether there have been positive changes in the last five years,

however, was much more difficult to determine.

There was also some sense that assessing direction of change suggested the idea of identifying a trajectory, making a judgment about whether things will be changing in the future.

3. The panel raised the issue of what should be the basis of assessing level of compliance. Although the panelists may not have explicitly decided on a comparative basis for decisions, it appeared that in some cases there was a *de facto* comparison with countries at similar levels of development ("not as bad as …").

4. There was some variation in the extent to which panelists allowed their general knowledge of a country's environment to impact their assessments. One panelist, in the absence of data, presumed that a particular kind of problem existed—based on his knowledge of the country.

5. The panel discussed the importance of understanding the sources and their traditions. For example, small changes in phrasing in reports of the ILO's Committee of Experts on the Application of Conventions and Recommendations (CEACR) can be very important. These might be overlooked by less-experienced assessors.

6. Some of the indicators include more than one topic which confuses assessment. For example, FoA A-16 asks about restrictions on "picketing or occupation of the workplace." The latter is a much different activity than the former. Similarly, FoA B-1 included job actions and crimes such as murder in a single indicator.

7. In the course of doing his individual assessments, one panelist personally contacted an expert in the Department of State in order to get up-to-date information. This raised the question about how such informal sources should be integrated with more established or freely-available sources.

8. As with other country panels, the Country C panelist found the outcome indicators (FoA C-1 to C-3) difficult to interpret. Although they assumed that the intent was that higher union density (C-1) is likely a sign of greater freedom of association, the indicator is a factual one and not one of "compliance." For number of strikes (C-2), it is unclear whether more strikes is a sign of an active and successful labor movement or of problems with the labor standards. In terms of outcomes, it was noted that it would be useful to have some indicator of the level of awareness among workers of their rights. There weren't suggestions, however, of how this would be measured.

9. The question was raised about the problem of moving from anecdotal evidence to an overall assessment of a country. It was noted that there may be certain sectors that get disproportionate attention in the western press. For example, apparel and toys, because of their western markets and consumer movements may skew the impression of the country as a whole.

10. There was often little information for assessing the indicators for "Discrimination and Equality." The range of bases for discrimination (gender, ethnicity, caste, etc.) also made it complex. It was noted that there were few indicators for assessing discrimination within a country against internal migrants or indigenous people.

11. In terms of the matrix approach generally, the panelists seemed to see it as a useful organization of information. It might be that the "C" indicators (outcomes), rather than items for assessing "compliance," might be used to provide an introductory context—a portrait of the labor conditions—for assessing the legal framework and the government performance. Understanding the national context should also involve some indications of the state of the economy and the social situation.

Appendix H

Matrixes for Assessors on the Three Panels

Country A   Freedom of Association and Collective Bargaining

Assessor: a		Direction of Change			
		Improving	Steady state	Worsening	(no assessment)
Level of Compliance	Some problems		A1(3), A9(3), A17(1), A21(3), B5(2), B13(2)		
	More extensive problems	A4(2), A12(2)	A3(2), A11(3), A13(2), A19(2), A20(2), B1(2), B2(1), B4(2), B8(3)		
	Severe problems	A2(1), A14(2)	A6(2), A7(2), A8(2), A10(3), A15(2), A16(3), B3(2), B6(2)		
	(no assessment)				A5(2), A18(), B7(), B9(), B10(), B11(), B12(), C1(3), C2(2), C3(2), C4()

Assessor: b		Direction of Change			
		Improving	Steady state	Worsening	(no assessment)
Level of Compliance	Some problems	A4(2), A5(2), A7(2), A21(2), B6(3)	A8(3), A10(2), A12(2), A13(2), B2(2), B4(2), B11(2)		
	More extensive problems	A3(3)	A2(3), B3(2)		
	Severe problems	B1(2)	A6(3), A11(2), A14(3), A15(3), A16(3), C4(2)		
	(no assessment)		C3(2)	C1(2)	A1(3), A9(3), A17(3), A18(), A19(3), A20(2), B5(2), B7(), B8(), B9(), B10(), B12(), B13(), C2(2)

Assessor: c		Direction of Change			
		Improving	Steady state	Worsening	(no assessment)
Level of Compliance	Some problems		A1(3), A2(3), A3(3), A8(3), A9(3), A11(3), A12(2), A13(3), A17(3), A18(2), A19(3), A20(2), A21(3), B4(2), B5(3), B7(1), B10(2), B12(3), B13(3), C2(2)		
	More extensive problems		A5(3), A6(3), A10(3), A14(3), A16(3), B1(3), B3(3), B6(3), B11(2)		
	Severe problems		A4(3), A7(3), A15(3), B2(3), B9(2), C4(3)	B8(3), C1(3), C3(3)	
	(no assessment)				

Confidence in Assessment:
(1) = low confidence
(2) = moderate confidence
(3) = confident
() = no response

Country A   Acceptable Conditions of Work

Assessor: a		Direction of Change			
		Improving	Steady state	Worsening	(no assessment)
Level of Compliance	Some problems		A2(3), A4(1), A6(3), A8(3), A9(3), A10(3), A12(3), A16(3),		
	More extensive problems	A3(2)		B5(1), B6(2)	
	Severe problems		A1(3), A5(3), A11(3), A13(3), A14(3)		
	(no assessment)	C3(1)		C4(1), C5(2)	A7(), A15(), B8(), B9(), B10(), B11(), B12(), B13(), C1(), C2()

Assessor: b		Direction of Change			
		Improving	Steady state	Worsening	(no assessment)
Level of Compliance	Some problems		A3(2), A7(2), A12(2), B1(2)		
	More extensive problems		B2(2)		
	Severe problems		B3(2)	B5(2), B6(2), B9(1), C4(3), C5(2)	
	(no assessment)			C1(2), C3(3)	A1(3), A2(3), A4(3), A5(3), A6(), A8(3), A9(3), A10(3), A11(3), A13(3), A14(3), A15(3), A16(3), A17(3), A18(3), A19(3), B4(3), B7(), B8(), B10(), B11(), B12(), B13(), C2()

Assessor: c		Direction of Change			
		Improving	Steady state	Worsening	(no assessment)
Level of Compliance	Some problems		A1(3), A2(3), A3(3), A4(2), A9(3), A10(3), A15(3), A16(3), A17(3), A18(3), A19(3), B1(3), B2(), B3(3), B4(3), B8(3), B12(1)	C3(2)	
	More extensive problems		A5(3), A6(3), A12(2), A13(3), A14(3), B6(2), B7(2), B9(2), B10(2), B11(1), C5(1)	A7(3), A8(3)	
	Severe problems		A11(3), B13(3), C1(1), C2(2), C4(2)	B5(3)	
	(no assessment)				

Country B   Freedom of Association and Collective Bargaining

Assessor: a		Direction of Change			
		Improving	Steady state	Worsening	(no assessment)
Level of Compliance	Some problems	A2(2), A3(3), A4(3), A8(2), B2(2)	A5(2), A7(1), A10(3), A11(2), A13(3), A16(1), A19(2), A21(3), B3(1),		
	More extensive problems	A12(2)	A6(3), A14(3), A17(), A18(2), A20(3), B1(2), B6(1), B8(1), B9(2), B10(2), B11(2), B12(1), B13(2), C1(3), C3(1), C4(1)		
	Severe problems		A1(3), A9(3), A15(3)		
	(no assessment)				

Assessor: b		Direction of Change			
		Improving	Steady state	Worsening	(no assessment)
Level of Compliance	Some problems	A2(3), A3(3), A4(3), A5(3), A6(3), A8(3), A10(3), A13(2), A20(3), B11(2), B13(2)	A7(2), A9(3), A11(3), A16(2), A18(2), A21(3), B4(2), B5(2), B6(2), B7(1), C2(1)		
	More extensive problems	A12(3), A14(3)	A15(2), A17(3), A19(2), B1(), B2(2), B12(2)		
	Severe problems		A1(3), B3(2), B8(2), B9(2), B10(2), C1(), C3(2), C4(2)		
	(no assessment)				

Assessor: c		Direction of Change			
		Improving	Steady state	Worsening	(no assessment)
Level of Compliance	Some problems	A3(), A6(2), A12(2), B9(3)	A2(3), A4(3), A5(2), B11(2)		
	More extensive problems		A13(), A14(3), B8(3)		
	Severe problems		A1(3)		
	(no assessment)	B10(3)	A7(), A8(3), A9(3), A11(3), A15(1), A17(3), A19(3), A21(3), B1(2), B5(2)		A10(3), A16(1), A18(), A20(1), B2(), B3(), B4(), B6(), B7(), B12(), B13(), C1(3), C2(), C3(), C4()

Country B   Forced or Compulsory Labor

Assessor: a		Direction of Change			
		Improving	Steady state	Worsening	(no assessment)
Level of Compliance	Some problems		A1(3), A2(3), A3(3), B8(1)		
	More extensive problems	A4(3), B4(2)	B2(2), B3(1), B5(1), B7(1)		
	Severe problems	B1(3), B6(1)	C1(1), C2(1), C3(1), C4(1)		
	(no assessment)				

Assessor: b		Direction of Change			
		Improving	Steady state	Worsening	(no assessment)
Level of Compliance	Some problems	A3(3), A4(3)	A1(3), A2(3), B3(2), B5(2)		
	More extensive problems	B4(2), B7(2), B8(2), C4(2)	B2(2), B6(1), C1(1), C2(1), C3(1)	B1(2)	
	Severe problems				
	(no assessment)				

Assessor: c		Direction of Change			
		Improving	Steady state	Worsening	(no assessment)
Level of Compliance	Some problems	A4(2)	B2(2), B3(3)		
	More extensive problems				B1(3)
	Severe problems				
	(no assessment)		A1(), A2(3), A3(), B7(3)		B4(), B5(), B6(), B8(), C1(2), C2(), C3(), C4()

Country C   Freedom of Association and Collective Bargaining

Assessor: a		Direction of Change			
		Improving	Steady state	Worsening	(no assessment)
Level of Compliance	Some problems		A1(3), A2(3), A3(3), A4(3), A5(3), A6(3), A7(3), A8(3), A9(3),		
	More extensive problems		A11(2), A15(2), A19(2), A21(3), B8(2), B9(1), B10(1), B13(2), C1(3), C3(2)		B11(1)
	Severe problems		A14(3), B1(3), B2(3), B3(3), C2(2), C4(2)		
	(no assessment)				

Assessor: b		Direction of Change			
		Improving	Steady state	Worsening	(no assessment)
Level of Compliance	Some problems	B5(2)	A1(3), A5(1), A9(3), A17(1), A18(1), A19(2), A20(2), B6(1), B7(1)		
	More extensive problems	A2(2)	A4(2), A7(2), A10(3), A11(2), A12(3), A13(2), A15(2), A21(3), B4(2), B9(1), B10(1), B11(1), B12(1), B13(2)		
	Severe problems	A14(3)	A3(3), A6(3), A8(3), A16(3), B1(3), B2(3), B3(3), B8(2)		
	(no assessment)	C1(2)			C2(1), C3(1), C4(1)

Assessor: c		Direction of Change			
		Improving	Steady state	Worsening	(no assessment)
Level of Compliance	Some problems		A1(3), A2(3), A5(2), A8(2), A9(3), A10(3), A11(2), A13(2), A18(1), A20(1), A21(3), B4(2), B6(2), B12(3), B13(2), C2(1)	C3(1)	
	More extensive problems		A3(3), A4(2), A6(3), A7(2), A12(3), A19(3), B1(3), B5(3), B7(1), B9(2), B10(2), B11(2), C1(1), C4(3)	A15(3)	
	Severe problems		A16(3), A17(3), B3(2), B8(3)	A14(3), B2(2)	
	(no assessment)				

Country C   Discrimination and Equality

Assessor: a		Direction of Change			
		Improving	Steady state	Worsening	(no assessment)
Level of Compliance	Some problems	A9(2), B1(2)	A1(3), A2(3), A3(3), A4(3), A8(3), B2(3), B10(2), C2(3), C7(2)		
	More extensive problems	A5(3)	A6(2), A7(2), A10(3), B5(2), C1(3), C5(2), C6(2), C8(2), C9(2)		
	Severe problems				
	(no assessment)				B3(), B4(), B6(), B7(), B8(), B9(), B11(), C3(), C4(), C10()

Assessor: b		Direction of Change			
		Improving	Steady state	Worsening	(no assessment)
Level of Compliance	Some problems	B9(2), B10(2)	A1(3), A2(3), A3(3), A4(3), B2(2)		
	More extensive problems	A5(3), C1(3)	A6(2), A7(3), A8(2), B1(2), B5(2), B6(2), B8(1), B11(2), C2(2), C4(2)		C10(2)
	Severe problems		A9(2), A10()	C6(2), C7(2), C9(3)	C8(2)
	(no assessment)				B3(), B4(), B7(), C3(), C5()

Assessor: c		Direction of Change			
		Improving	Steady state	Worsening	(no assessment)
Level of Compliance	Some problems		A1(3), A2(3), A3(3), A4(3), A9(2), B2(3), B6(2), B10(2)		
	More extensive problems	C1(3), C2(3)	A8(2), A10(2), B1(1), B5(1), B7(1), B8(1), B9(1), B11(1), C3(1), C6(3), C7(1), C8(3)	C9(2)	
	Severe problems	A5(2)	A6(2), A7(2), B3(3), B4(2), C4(3), C5(3), C10(3)		
	(no assessment)				

Appendix I
Revisions in Assessments following Discussion

Panel A

Indi-cator	Level of Compliance			Diff. Ind.		Confidence (in initial assessment)				Level of Compliance-Revision			Diff. Ind.
	Panelists					Panelists				Panelists			
	a	b	c			*a*	*b*	*c*		a	b	c	
Freedom of Association and Collective Bargaining													
A02	Severe	More	Some	4		*Low*	*High*	*High*		More	More	More	0
A04	More	Some	Severe	4		*Mod.*	*Mod.*	*High*		More	Some	Severe	4
A05		Some	More			*Mod.*	*Mod.*	*High*			Some	Some	
A07	Severe	Some	Severe	4		*Mod.*	*Mod.*	*High*		Severe	More	Severe	2
A08	Severe	Some	Some	4		*Mod.*	*High*	*High*		Severe	More	More	2
A10	Severe	Some	More	4		*High*	*Mod.*	*High*		More	Some	More	2
A11	More	Severe	Some	4		*High*	*Mod.*	*High*		More	Severe	Some	4
A16	Severe	Severe	More			*High*	*High*	*High*			Severe	More	
B02	More	Some	Severe	4		*Low*	*Mod.*	*High*		Some	Some	Severe	4
B06	Severe	Some	More	4		*Mod.*	*High*	*High*		More	Some	More	2
Acceptable Conditions of Work													
A02	Some		Some			*High*	*High*	*High*		More	More	More	
A04	Some		Some			*Low*	*High*	*Mod.*		More	Some	Severe	
A05	Severe		More			*High*	*High*	*High*			Some	Some	
A07		Some	More				*Mod.*	*High*		Severe	More	Severe	
A08	Some		More			*High*	*High*	*High*		Severe	More	More	
A10	Some		Some			*High*	*High*	*High*		More	Some	More	
A11	Severe		Severe			*High*	*High*	*High*		More	Severe	Some	
A16	Some		Some			*High*	*High*	*High*			Severe	More	
B02	Some	More	Some	2		*High*	*Mod.*			Some	Some	Severe	4
B06	More	Severe	More	2		*Mod.*	*Mod.*	*Mod.*		More	Some	More	2
Total *difference index*--panel A				36									26

Panel B

Indi-cator	Level of Compliance			Diff. Ind.		Confidence *(in initial assessment)*				Level of Compliance-Revised			Diff. Ind.
	Panelists					Panelists				Panelists			
	a	b	c			a	b	c		a	b	c	
Freedom of Association and Collective Bargaining													
A02	Some	Some	Some	0		Mod.	High	High		Some	Some	Some	0
A04	Some	Some	Some	0		High	High	High		Some	Some	Some	0
A05	Some	Some	Some	0		Mod.	High	Mod.		Some	Some	Some	0
A07	Some	Some				Low	Mod.			Some	Some		
A08	Some	Some				Mod.	High	High		Some	Some		
A10	Some	Some				High	High	High		Some	Some		
A11	Some	Some				Mod.	High	High		Some	Some		
A16	Some	Some				Low	Mod.	Low		Some	Some		
B02	Some	More				Mod.	Mod.			Some	More		
B06	More	Some				Low	Mod.			More	Some		
Forced or Compulsory Labor													
B06	Severe	More				Low	Low			Severe	Severe	Some	4
B07	More	More				Low	Mod.	High		Some	More	Some	2
B08	Some	More				Low	Mod.			Some	More	Some	2
C01	Severe	More				Low	Low	Mod.		Severe	More		
C04	Severe	More				Low	Mod.			Severe	More	More	2
Total *difference index*--panel B				n/a									n/a

Panel C

Indi-cator	Level of Compliance			Diff. Ind.		Confidence (in initial assessment)				Level of Compliance--Revisions			Diff. Ind.
	Panelists					Panelists				Panelists			
	a	b	c			*a*	*b*	*c*		a	b	c	
Freedom of Association and Collective Bargaining													
A02	Some	More	Some	2		*High*	*Mod.*	*High*		Some	More	Some	2
A04	Some	More	More	2		*High*	*Mod.*	*Mod.*		Some	More	More	2
A05	Some	Some	Some	0		*High*	*Low*	*Mod.*		Some	Some	Some	0
A07	Some	More	More	2		*High*	*Mod.*	*Mod.*		Some	More	More	2
A08	Some	Severe	Some	4		*High*	*High*	*Mod.*		Some	More	More	2
A10	Some	More	Some	2		*High*	*High*	*High*		Some	More	Some	2
A11	More	More	Some	2		*Mod.*	*Mod.*	*Mod.*		More	More	Some	2
A16	Some	Severe	Severe	4		*Mod.*	*High*	*High*		Some	Severe	More	4
B02	Severe	Severe	Severe	0		*High*	*High*	*Mod.*		Severe	Severe	Severe	0
B06	Some	Some	Some	0		*Low*	*Low*	*Mod.*		Some	Some	Some	0
Discrimination and Equality													
A05	More	More	Severe	2		*High*	*High*	*Mod.*		More	More	More	0
A08	Some	More	More	2		*High*	*Mod.*	*Mod.*		Some	More	More	2
A09	Some	Severe	Some	4		*Mod.*	*Mod.*	*Mod.*		Some	More	Some	2
A10	More	Severe	More	2		*High*		*Mod.*		More	Severe	Some	4
B01	Some	More	More	2		*Mod.*	*Mod.*	*Low*		More	More	Some	2
B06		More	Some				*Mod.*	*Mod.*		More	More	Some	
B09		Some	More				*Mod.*	*Low*		Some	Some	More	
C07	Some	Severe	More	4		*Mod.*	*Mod.*	*Low*		More	More	Some	2
C10		More	Severe				*Mod.*	*High*		More	More	Severe	
Total *difference index*--panel C				34									28

Appendix J
Commentary by Panelists on Specific Indicators

Issues arising in application of the indicators in the pilot test

<u>Freedom of Association</u>

A-01. whether the country has ratified ILO convention No. 87	Suggests binary response: yes or no--not some problems, more extensive problems, or severe problems
A-02. whether there are legal provisions that entitle workers or employers to establish and join organizations of their own choosing without previous authorization	Double-barreled question: "without prior authorization" confuses question (if referring to "registration" of unions, should say this); "registering"; combines workers joining unions with separate issue of employers forming associations Wording and intent: Government's discretion to certify or not extends the indicator beyond legal provisions
A-03. the extent to which there are legal restrictions on the ability of certain categories of workers to organize (such as civil servants, teachers, or workers without contracts of employment), and numbers of workers in each category	Wording and intent: "numbers of workers in each category" extends the indicator beyond legal framework to impact of its application
A-04. whether there are certain sectors where there is no right to organize (such as in export processing zones or in other tradable sectors or in agricultural and informal sectors)	Wording and intent: appears to extend beyond legal framework to application of law to particular workplaces/industries
A-05. whether there are other forms of distinction or discrimination in right to organize, such as race, nationality, sex, opinion, political affiliation, or citizenship (for example, excluding legal immigrants)	Suggests binary response: yes or no, although may imply gradations of compliance depending upon extent of limitations Wording and intent: Discrimination may extend beyond labor law, such as banning publications in foreign languages
A-06. whether there are legal provisions that permit the government to interfere in freedom of association by workers or employers	Wording and intent: appears to require going beyond the legal framework to look at how the law is applied; question about what is meant by "interfere in freedom of association"

89

A-07. whether there are legal restrictions on political activities of unions and employers' organization, either by establishing a close relationship between trade union organization and political parties, or by prohibiting all political activities for trade unions (including making financial contributions to a political party or candidate)	
A-08. whether national laws protect workers from discrimination if they join a union or participate in union activities	Suggests binary response: "Yes, national law protects workers" Wording and intent: appears to require going beyond the legal framework to look at how the law is applied
A-09. whether the country has ratified Convention No. 98 (collective bargaining)	Suggests binary response
A-10. whether collective bargaining is protected in law	Answer set problem: needs "no problems" as possible response
A-11. whether there are legal restrictions on the mechanism of collective bargaining (such as the mandatory exclusion of some issues or the ability of employers to refuse to bargain with a recognized union)	Wording and intent: appears to require going beyond the legal framework to look at how the law is applied--practical restrictions may not be in the law, but in its application Double-barreled question: Difficult to determine level of compliance when parenthetical examples of restrictions are qualitatively different
A-12. the extent to which there are categories of workers not permitted to negotiate a collective bargaining agreement (such as civil servants, teachers, or workers in special activities), and numbers of workers in each such category	Wording and intent: appears to require going beyond the legal framework to look at how the law is applied
A-13. whether there are certain sectors where there is no right to negotiate a collective bargaining agreement (such as in export processing zones or in other tradable sectors, or in agricultural and informal sectors	Wording and intent: appears to require going beyond the legal framework to look at how the law is applied; parenthetical examples of informal sector, export processing zones, agriculture are qualitatively different

A-14. whether the principle of a strike as a means of action of organizations is generally recognized	Wording and intent: "generally recognized" implies going beyond the legal framework to look at how the law is applied
A-15. the extent to which the government can forbid certain kinds of strikes (such as "protest strikes," "sympathy strikes," "go slow" strikes, or "work-to-rule" slowdown), or strikes under certain conditions (such as economic or political "crisis"), or strikes requiring a majority of workers involved to authorize a strike	Double-barreled question: single indicators encompasses qualitatively different situations (types of strikes and types of conditions) Wording and intent: WebMILS explanation of this indicator suggests that indicator should read: "…government can forbid…strikes that did <u>not</u> have approval from a majority of the workers involved."
A-16. the extent to which the government can limit picketing or occupation of the workplace	Wording and intent: appears to go beyond legal framework of industrial relations because actions against picketing and occupation of the workplace are usually taken under other parts of law (e.g., maintaining public order) Double-barreled question: picketing and occupying the workplace are qualitatively different
A-17. whether legal regulations ban employer lockouts	Suggests binary response: could be re-written with "to what extent" instead of "whether" they exist in law
A-18. whether workers are legally able to ratify or reject agreements reached by union leaders	Compliance question: unclear whether this corresponds to an international standard; union democracy internally is usually is regulated in a union's constitution and bylaws in national law but
A-19. whether legal regulations permit employers to dismiss striking workers, or permit hiring of permanent strike replacement workers	Wording and intent: appears to go beyond legal framework to practice
A-20. assessment of extent to which "right to work" laws or other "free rider" provisions undermine the ability of workers to organize and the extent to which laws require workers to join a given trade union as a condition of employment or that new workers be hired through a given trade union	Wording and intent: part of labor-relations discussion in U.S., but less universally applicable

A-21. whether workers' and employers' organizations may legally affiliate with international bodies	Wording and intent: some question about meaning of "affiliate" and whether this extends to received financial assistance from international groups
B-01. the extent to which union organizers are jailed or exiled by the government, or fired, injured, or murdered without prompt and effective prosecution on the part of the government	Double-barreled question: combines actions with very different levels of severity and sanctions; perhaps a question is needed about the effectiveness of the judicial process in such instances
B-02. the extent to which strikers suffer retaliation without prompt and effective prosecution on the part of the government	Double-barreled question: combines a question about employer actions (retaliation) with governmental response in case of such retaliation
B-03. whether there are defects in the government's complaint process, such as excessive delays or expenses, light penalties, or nonpunishment of offenders	Little or no information
B-04. the extent to which the government interferes in freedom of association (e.g., by declaration of martial law, "state of crisis," or by suspension or dissolution of associations by administrative authority)	Wording and intent: unclear whether the examples are intended to restrict indicator to discretionary governmental actions and/or emergency decrees); otherwise, overlaps with B01-B03
B-05. the independence of trade unions from political control	
B-06. the ability of trade unions to provide support for political parties and candidates	Wording and intent: unclear what is included in "support"
B-07. government actions to combat labor-related corruption (such as control of unions by criminal figures for use as a protection racket or for financial skimming) without prompt and effective prosecution	No information found by any of the panels
B-08. the adequacy of personnel and budgets of labor regulation departments compared to number of workplaces, the frequency and adequacy of labor inspections, the caseloads of labor administrative bodies and labor court, and whether bribes are paid to labor inspectors by employers without effective prosecution	Double-barreled question: combines a number of different questions; also, difficult to define threshold for "adequacy" Little or no information

B-09 the efforts by national, state or provincial, and municipal authorities to educate workers about their rights and remedies and of the effectiveness of those efforts	Little or no information
B-10. the efforts on the part of national, state or provincial, and municipal authorities to engage in capacity-building for governmental officials with responsibility for labor matters and of the effectiveness of those efforts	Little or no information found
B-11. governmental actions in encouraging consultation and in facilitating dissemination of best practices in labor-management cooperation, through unions, employer groups, labor-management organizations, labor-oriented NGOs, and tripartite social dialogue	Little or no information found
B-12. government actions in hindering or facilitating the formation and functioning of labor-advocate NGOs and of independent labor inspection, monitoring, and certification organizations	Little or no information found
B-13. government actions in encouraging and enabling utilization of domestic and international channels about problems, difficulties, or violations of freedom of association and effective recognition of the right to collective bargaining	Little or no information found
C-01. union density	Factual question, not "compliance:" perhaps more useful as a descriptor of context; difficult to interpret in practice
C-02. frequency, length, and person-days of legal strikes	Factual question, not "compliance:" perhaps more useful as a descriptor of context; difficult to interpret in practice

C-03. percentage of workers covered by collective bargaining agreements	Factual question; not "compliance"--perhaps more useful as a descriptor of context; also little or dated information; P2D: International Yearbook on Labor Statistics (World Bank) is not available online; P3D: compliance is not the right term, BUT would be positive if a high % of workers are covered by agreements
C-04. incidents of discrimination against union organizers, unions, or employer associations	Double-barreled question: discrimination against employer associations is qualitatively different than against unions--may be more related to civil liberties than labor rights; also discrimination against union organizers is qualitatively different from discrimination against workers Little or no information found

Forced or Compulsory Labor

A-01. ratification of ILO Convention No. 29 (forced labor, 1930, in response to concerns about "native labor in colonial contexts)	Suggests binary response
A-02. ratification of ILO Convention No. 105 (forced labor, 1957, in response to forced labor camps as political punishment and discrimination)	Suggests binary response
A-03. ratification of ILO Convention No. 182 (worst forms of child labor, 1999)	Suggests binary response
A-04. constitutional provisions, statutes, or regulations prohibiting the principal forms of forced labor	
B-01. an effective system of labor inspection with responsibility for identifying, remedying, and enforcing forced labor prohibitions	
B-02. an independent judiciary to redress forced labor violations	Double-barreled question: combines question about whether there is there a judiciary that addresses forced labor with whether it is independent
B-03. courts with authority to order fines and penal sanctions, including incarceration, for forced labor violations	Wording and intent: asks about judiciary's authority, not whether the law authorizes incarceration for forced labor violations--part of ILO Convention 29
B-04. community awareness and use of local vigilance committee to identify and monitor forced labor situations	Wording and intent: term "vigilance committees" used primarily in India; is this intended to apply generally to community groups that assume similar responsibilities?
B-05. government sponsored credit institutions	Wording and intent: not a clear connection with forced labor--very relevant re places with extensive bonded labor-such as India--indicator could be prefaced with "If bonded labor is present,…" Alternative indicators could ask about government's efforts to prevent recruitment of people in forced labor situations; preventative steps the government is taking such as regulating labor recruiters and employment agencies

B-06. national, state, and local resources devoted to identification, remediation, and enforcement, measured by: the percentage of labor department budget devoted to enforcement of forced labor prohibition; the number of labor inspectors per 100,000 workers in the formal and informal sectors; and annual arrests, prosecutions, and penalties for crimes related to forced labor	Wording and intent: perhaps should be: "national, state, and local resources devoted to identification, remediation, and enforcement of forced labor provisions" with rest of original indicator either as examples for assessors; little information found
B-07. receipt of technical assistance from the ILO, U.S. government, or other sources for: establishing statutory and regulatory schemes; establishing a system of labor inspection; establishing rehabilitation systems	Double-barreled indicator: it could be possibly separated into three indicators reflecting: the statutory system, labor inspections, and rehabilitation; a question was also raised about whether "level of compliance" is appropriate--while it demonstrates a country's efforts, there is not ILO convention setting a standard for accepting technical assistance
B-08. government support of NGO activities related to forced labor, such as victim rehabilitation services, advocacy and awareness-raising programs, and microcredit lending practices	Little or no information
C-01. statistics and reports on the number of persons that have been taken out of forced labor	Little or no information
C-02. statistics and reports of persons who have been rehabilitated from forced labor	Little information and what exists is not current
C-03. statistics and reports of persons who have returned to forced labor	Little information found

C-04. statistics on the number of workers still in the prohibited forms of forced labor: slavery and abductions; compulsory participation in public works projects; coercive recruitment systems, particularly in rural areas; bonded labor, including bonded child labor; trafficking in persons; domestic workers in forced labor situations; prison labor and rehabilitation through work	Little information and what exists is out of date

Discrimination

A-01. ratification of ILO Convention 100 on equal remuneration	Suggests binary response
A-02. ratification of ILO Convention No 111 on discrimination (employment and occupation)	Suggests binary response
A-03. ratification of the U.N. International Convention on the Elimination of All Forms of Racial Discrimination	Suggests binary response
A-04. ratification of the U.N. Convention on the Elimination of Forms of Discrimination Against Women	Suggests binary response
A-05. whether there are laws that prohibit discrimination in employment on the grounds of race, color, sex (including sexual harassment), religion, political opinion, national extraction, and social origin	
A-06. whether there are laws that cover additional grounds on which employment discrimination is prohibited, for example disability or sexual orientation	Wording and intent: appears to require going beyond the legal framework to look at how the law is applied
A-07. whether there are laws that also protect migrant workers from discrimination in employment	Wording and intent: question about application to internal migrants; also question about whether indigenous peoples are protected; some ambiguity about whether the question also refers to protection of out-migrants from the country
A-08. whether there are laws that prohibit discrimination in access to and ownership of assets, including property ownership, inheritance, and access to other assets or credit	Wording and intent: extends beyond legal framework to implementation of the law (e.g., in cases of restrictions on rights of women and certain ethnic groups)

A-09. whether there are laws that make nondiscrimination in employment meaningful in practice with equal access to education, training, vocational guidance, and placement services; maternity protection; and parental leave	Wording and intent: "meaningful in practice" implies going beyond the letter of the law to its implementation
A-10. Whether there are laws or a legally established or recognized machinery for wage determination that ensure equal remuneration for work of equal value	Wording and intent: no provisions in law, but regulatory controls exist Little or no information
B-01. whether there is a national mechanism to promote equality and whether it has an employment focus or employment component in a broader policy	Little or no information
B-02. whether the country has a labor inspectorate	Suggests binary response
B-03 the breadth of labor inspections in the country, in terms of number of visits, frequency of visits, number of workers covered, etc.	Factual question, not "compliance;" unclear what standard to use to assess compliance
B-04. the level of resources devoted to the labor inspectorate in terms of number of personnel and budget, absolute or relative to number of workers or spending	Little or no information found
B-05. whether nondiscrimination issues are explicitly included in later inspections and inspectors trained in them	Little or no information found
B-06. whether there is a grievance mechanism for airing and investigating discrimination complaints, "whistle-blowers" or complainant are protected from retaliation, and the grievance mechanism is adequately supported and funded	Little or no information found; possible that if no information exists for this indicator, then that suggests no programs exist

B-07. measures of utility of grievance procedures, including number of complaints, length of process, penalties imposed in complaints upheld, and annual prosecutions and fines for violations of nondiscrimination laws	Little or no information found
B-08. where applicable, whether there are affirmative action programs aimed at promoting equality and not at elevating one group over another	Little or no information found
B-09. whether there are public education or sensitivity campaigns, both to education vulnerable citizens about their rights and to change cultural and traditional attitudes that contribute to discrimination	Suggests binary response Little or no information found
B-10. whether there are requests for or implementation of International technical assistance programs, which can signal improvements in government efforts and, possibly, effectiveness	Suggests binary response Wording and intent: phrase beginning with "which can signal…" explains the rationale for the indicator but it also extends it to assessing "effectiveness" of government programs
B-11. whether there are relevant supportive policies, such as child care	Little or no information found
C-01. differences in illiteracy rates between men and women and among different ethnic, racial, religious, or other groups	Little information on ethnic, racial and religious groups
C-02. differences in school enrollment rates (primary, secondary, and tertiary, or combined) between men and women and among different ethnic, racial, religious, or other groups	Little or no information found on ethnic and religious groups
C-03. differences in vocational training enrollment rates between men and women and among different ethnic, racial religious, or other vulnerable groups	Little or no information found
C-04. documented discriminatory practices, such as widespread sexual harassment or pregnancy testing for job applicants	Little information in WebMILS; some information found in reports of the Asian Development Bank and news sources

C-05. employment status—wage or salaried worker, self-employed, or contributing family member	Little information found; information that was found (e.g., women's labor force participation rates) was difficult to interpret and apply
C-06. distribution of women and other potentially vulnerable groups by sector	Wording and intent: unclear wording
C-07. distribution of employment in the urban informal sector	Wording and intent: unclear what is being asked; for example, does indicator ask about number of workers in the urban informal sector or their distribution within the informal sector?
C-08. distribution of women and other potentially vulnerable groups by occupation within sectors, including in administrative and management positions	
C-09. relative unemployment rates, especially changes over time and during adjustments	Wording and intent: what is meant by "during adjustments"
C-10. relative wage data could be an indirect indicator of occupational or sectoral discrimination that channels women and minorities into lower paid types of jobs, but they are not available for large numbers of countries or consistently over time	Wording and intent: unclear wording; descriptive of the rationale for an indicator rather than stating an indicator

Acceptable Conditions of Work

A-01. ratification of ILO Convention No. 131 on minimum wage fixing machinery	Suggests binary response
A-02. national constitution or law establishing one or more minimum wages	Suggests binary response
A-03. application of the minimum wage law or laws (to whom do the minimum wage law/laws apply? Do the legal minimums vary by geographic region, economic sector, and/or by establishment size?)	Double-barreled question: asks a number of different questions about the minimum wage
A-04. minimum wage level(s) specified in the law (in local currency and in US. Dollars and, if possible, adjusted using the World Bank's Purchasing Power Parity exchange rate for the local currency)	Factual question, not "compliance:" perhaps more useful as a descriptor of context; difficult to assess in terms of compliance
A-05. ratification of ILO Convention No. 1, 30, or 47, limiting the regular workweek to 48 hours or less	Suggests binary response; in case of the country under review, these ILO conventions were not ratified, but the national law meets the relevant standards
A-06. national constitutional or legal provision for a regular workweek of 48 hours or less	Suggests binary response
A-07. the coverage of laws limiting the regular workweek (e.g., only establishments of a certain size, only those in certain industrial sectors or regions)	Wording and intent: questions about the implementation of the law, particularly because of the high incidence of informal employment and unregistered workplaces
A-08. provision in national laws or regulations for overtime beyond the regular workweek	Suggests binary response
A-09. ratification of ILO Convention No. 14, providing a weekly day of rest	Suggests binary response
A-10. provision in the national constitution or law for a weekly day of rest	Suggests binary response

A-11. ratification of ILO Convention No. 101, providing paid holidays for agricultural workers	Suggests binary response
A-12. provision in the national constitution or laws for paid holidays for agricultural workers	Suggests binary response
A-13. ratification of ILO Convention No. 52, providing for an annual paid holiday of at least 6 days	Suggests binary response
A-14. ratification of ILO Convention No. 132, providing for an annual paid holiday of at least three weeks (supersedes ILO Convention 52)	Suggests binary response
A-15. provision in national laws/constitution for a specified number of paid holidays for employed workers	Suggests binary response; the number of paid holidays could provide a basis for rank-order response
A-16 ratification of ILO Convention No. 81, on labor inspection	Suggests binary response
A-17. provision in national laws for inspection of workplaces	Suggests binary response
A-18. ratification of ILO Convention No.155 on mechanism to provide health and safety	Suggests binary response
A-19. provision in the national constitution or laws for workplace health and safety	Suggests binary response
B-01. a mechanism for fixing minimum wages	Suggests binary response
B-02. an agency to promote and enforce laws governing hours of work	Suggests binary response; more information about this agency could be developed with questions comparable to those for the labor inspectorate (B-04 - B-07)
B-03. an agency to promote and enforce laws protecting occupational health and safety	Suggests binary response; no indicators about effectiveness of health and safety inspections; could be expanded with questions comparable to those for the labor inspectorate (B-04 - B-07)

B-04. a labor inspectorate	Suggests binary response
B-05. the breadth of labor inspections in the country (number and frequency of visits, geographic regions, or industry sectors inspected)	
B-06. the level of resources (e.g., personnel, pay, or budget) devoted to the labor inspectorate	
B-07. a labor inspectorate trained to and focused on wages, hours, and occupational safety and health standards	Little or no information
B-08. an administrative or judicial complaint mechanism	Little or no information
B-09. effectiveness of the complaint mechanism (in such terms as number of complaints brought compared with number of complaints heard, number of prosecutions, fines, or arrests, and length of time for complaint resolution	Little or no information; also combines a number of questions/indicators
B-10. government programs to combat problems in the areas of wages, hours, and occupation safety and health	Little or not information; unclear how this is distinguished from questions about a labor inspectorate
B-11. government-sponsored education programs focusing on wages, hours, and occupation safety and health	Little or no information
B-12. government receipt of international technical assistance in the areas of wages, hours, and occupational safety and health	Little or no information
B-13. government support for NGO activities designed to improve compliance with wage, hour, and occupational safety and health laws	Little or no information

C-01. average and median manufacturing wages compared with minimum wage(s) in local currency and in U.S. dollars	Factual question, not "compliance:" perhaps more useful as a descriptor of context; difficult to interpret in practice
C-02. average earnings in industries that export to the United States	Factual question, not "compliance:" perhaps more useful as a descriptor of context; difficult to interpret in practice; focusing on exports to the US does not contribute to an assessment of labor rights generally
C-3. average hours worked per week	Factual question, not "compliance;" unclear what standard would be used to assess average number of hours worked
C-04. number of work-related fatalities per 100,000 workers, both overall and by industry sectors	Factual question, not "compliance;" unclear would be a comparative standard
C-05. number of occupational injuries, both absolute and as a fraction of the total workforce and the workforce covered by health and safety laws	Factual question, not "compliance;" unclear would be a comparative standard