

The Social Life of Measurement

Sally Engle Merry¹

VERSTEEG, MILA, and TOM GINSBURG. 2017. Measuring the Rule of Law: A Comparison of Indicators. *Law & Social Inquiry* 42 (forthcoming). Available at <http://onlinelibrary.wiley.com/doi/10.1111/lsi.12175/abstract>.

Indicators promise to measure the world, to convert the buzzing confusion and complexity of everyday life into clean numbers that permit comparison and ranking. They succeed in ordering the world. But to some extent, the world also orders the indicators. In this fascinating article, Mila Versteeg and Tom Ginsburg note a striking similarity among rule of law indicators despite their varied definitions and measurement strategies. They argue that this similarity occurs because all these indicators tap into an underlying principle of impartial government, a claim well supported in the article. But they also suggest that this convergence is the result of the similarities in the use of expert opinion and overlapping data in the construction of all four indicators. In other words, the article hints at the ways in which indicators are social constructions shaped by experts and the constraints of knowledge production. Here I will take their second argument farther, suggesting several ways that indicators are creatures of the social context in which they are made.

The article shows very clearly that despite significant differences in definitions of the rule of law and measurement strategies, the four indicators of rule of law they examine have striking similarities. This means, I gather, that the indicators rank countries in a similar order of greater or lesser compliance with the standards of the rule of law. The article usefully points out the ambiguity of the concept itself, particularly the slippage between a narrowly procedural definition and one that incorporates substantive conceptions such as human rights. They conclude that the similarity they observe comes from an underlying principle that all are in some ways measuring, that of impartial government. But they raise a second possibility, more focused on the challenges of producing quantitative knowledge itself: that the experts are using similar data and assumptions to make their judgments. In other words, the similarity reflects shared expert judgments. The disparity they observe between the World Justice Project's expert opinion surveys and the popular surveys reinforces this argument.

I would like to expand this argument by discussing several other aspects of measuring the rule of law that show how the pressure of the surrounding social context, forms of expertise, and availability of data determines how a concept is measured. In this

Sally Merry Engle is Silver Professor of Anthropology at New York University. She may be contacted at sally.merry@nyu.edu.

sense, the world makes indicators. First, there are clearly interpretive issues in determining what criteria should be used to assess the rule of law. While the article usefully compares different approaches, it does not open up the question of how any project to define the rule of law selects one or another approach and set of factors (see Merry, Davis, and Kingsbury 2015; Merry 2016). It is helpful to distinguish between a procedural or substantive definition, but this does not explain which definition is chosen or what concepts are considered important to measure. There seems to be a clear assumption that the rule of law is what rich democracies have, so the categories chosen may be those valued in those societies, particularly when the creators come from those societies. Indeed, there is a strong Euro-American core to all of these projects. Thus, there is an initial interpretive move to define the rule of law according to Euro-American standards. The focus on human rights reinforces this cultural orientation, since this system is in its origins, although not in its present condition, a largely Euro-American project. Thus, the very conceptualization of rule of law is a cultural project shaped by the particular aspirations of some countries, often those promoting the measurement system.

Second, as the authors point out, there may be underlying similarities in expert opinions that contribute to the similarity in the indicators. I think this is an even more important point than they claim. The experts are using similar data in many cases, often drawn from the same underlying surveys such as those of Transparency International or Amnesty International's study of human rights violations. Moreover, they may well be relying on shared templates for organizing information and categorizing it. In building these indicators, as is the case with most indicators, the creators are constrained by the data available. When there have not been measures of precisely what the indicator tries to measure, creators are forced to turn to proxies, which often do not adequately measure the underlying concept. The World Justice Project is unusual in that it collects its own data rather than relying on preexisting data. However, it is still constrained by its reliance on expert opinions. When queried about the condition of some dimension of the rule of law in his or her country, experts may themselves lack adequate information and be forced to make a best estimate. Indeed, in few countries has there been substantial research on all the constituent dimensions of the rule of law. The creators have to do their best to put together the data that is available, usually making pragmatic compromises about what they know and what the terms of the indicator require.

Third, a closer look at the constituent elements of the rule of law underscores the data availability problem. The list of variables measured by the four rule of law indicators hints at some of the challenges: contract enforcement, property rights, corruption, judicial efficacy, etc. To convert any of these concepts into measurable phenomena is clearly awesomely complex. Does this mean measuring the rules or how they operate in practice? While it is difficult to determine whether there are laws governing these concepts, measuring the extent to which they are followed in practice is far more difficult. It is likely that they are followed in some regions more than others, in some areas of social life but not all, for some social classes but not others. In many countries, data on the range of variation in practices is likely to be missing, so that experts will be forced to offer educated guesses. Thus, at this level, constructing the indicator requires considerable interpretation by experts of situations that they may not know well.

Finally, with a few exceptions, indicators are effective if they appeal to audiences who use them. Some feed directly into decision-making, but those discussed here are primarily used to develop public opinion and provide knowledge that can pressure bad performers to reform. These indicators exist in a dense environment of alternative indicators with which they compete for attention and credibility. Those created by powerful institutions, such as the World Bank, have more credibility and reach than those sponsored by NGOs such as Freedom House or the World Justice Project. But all of them have to portray a vision of the world that seems reasonable and relatively familiar to publics who consume this information. If an indicator shows that Zimbabwe has a strong rule of law while Sweden ranks near the bottom, it is far more likely that the indicator will be ignored than that there will be a public reassessment of the commitment to the rule of law by these two countries. The countries that appear at the top of the rankings are typically those that readers would expect to find there; there may be small differences but the rankings are mostly predictable.

Producers of indicators need to generate familiar rankings in order to retain audiences. Some small changes indicate that there is variation in the world, that things are moving, but major deviations from general expectations simply produce doubt about the credibility and value of the indicator itself. In order to be used, therefore, an indicator has to reflect the world as it is already understood.

Since indicators need to tell a familiar story in order to garner an audience, their creators are pressured to tell that story. If the data shows something different, this is inconvenient at best and fatal at worst. Consequently, they have little incentive to provide too much information about how they choose their categories of analysis, gather data, assess and compare this data, and generate rankings. In my research on indicators of gender violence, human rights, and sex trafficking, I found a lack of detailed information about how decisions were made about what to measure, about the availability of data, about the pragmatic compromises necessary to convert social complexity into numbers, and about the uncertainties resulting from these challenges (Merry 2016). Social science research, in contrast, has less need to attract publics and more to establish credibility in the eyes of other social scientists. It typically offers far more information about data collection, data quality, and conceptual formation. But providing such information is a luxury that organizations creating indicators for public consumption cannot afford.

In sum, this fascinating article raises some very important questions about how indicators are produced and the kind of knowledge they provide. Versteeg and Ginsburg's discussion of the social dimensions of indicator production opens up a crucially important field for further research and examination. Following their path, I have suggested a number of other factors that enter into producing indicators. My analysis contributes another perspective on the striking finding of this article that there is a convergence of indicators despite a divergence of criteria defining them. The idea of an underlying common principle makes sense, but so does the idea that indicators are shaped by the world in which they are developed and disseminated. Even as indicators make the world knowable, the world also makes them.

REFERENCES

Merry, Sally Engle. 2016. *The Seductions of Quantification: Measuring Gender Violence, Human Rights, and Sex Trafficking*. Chicago: University of Chicago Press.

Merry, Sally Engle, Kevin Davis, and Benedict Kingsbury, eds. 2015. *The Quiet Power of Indicators: Measuring Governance, Corruption, and Rule of Law*. Cambridge: Cambridge University Press.