Performance Indices: Ranking the Rankings

By The Economist


International comparisons are popular, influential—and sometimes flawed.

EDUCATION ministers across the globe quake in the run-up to the publication, every three years, of the OECD’s Programme for International Student Assessment (PISA), which rates 15-year-olds’ academic performance in dozens of countries. Those that do well can expect glory; the first PISA ranking, published in 2001, surprised the world by putting unshowy Finland near the top in every subject and made it a mandatory stop-off for any self-respecting education policymaker. Germany’s poor showing, by contrast, led to national hand-wringing, school reforms and the creation of a €4 billion ($5 billion) federal education support programme.
Similarly influential is the yearly Ease of Doing Business Index from the World Bank. Government presentations to investors will always show the highlights (provided, that is, there are numbers worth boasting about). The Trafficking in Persons (TIP) report compiled by America’s State Department each year ranks governments on their perceived willingness to combat trafficking. A bad showing blackens a country’s name and can mean losing aid and investment.

Such performance indices, which rank social issues or policy outcomes in different countries by combining related measures into a single score for each, are enjoying a boom. Their number has soared over the past two decades (see chart). For many issues, rival indices must now battle it out. “Numbers, rating and ranking catch people’s attention and make information easy to process,” says Judith Kelley of Duke University, who studies the impact of global indicators on policy. Rankings spread like wildfire on the web: some have been cited online more than a million times.
The best indices are meticulous (PISA, for instance, combines dozens of carefully standardised sub-measures and raises statistical caveats). But others are based on shaky figures that are calculated differently in different countries. And choosing what to include often means pinning down slippery concepts and making subjective judgments.
An index of democracy, freedom or happiness means putting hard numbers to the fairness of elections, weighing civil liberties against economic rights, or deciding how much to rely on surveys.

However an index is calculated, voters tend to conclude that their country’s position is at least partly due to government policies—and governments agree, at least when they do well. Increasingly, though, the causality flows the other way. “Ratings and rankings can be powerful tools of both branding and influence,” says Ms Kelley. Together with Beth Simmons of Harvard University, she has found that a big reason for the boom in indices is their growing use by governments, NGOs and campaigners to shape new laws and get them passed.

**Numbers that count**

The researchers’ main case study is TIP, which was first published in 2001. That year’s annual report covered 79 countries; it now ranks almost 190. By placing a heavy weight on whether countries have laws against human trafficking, TIP has spurred a global move towards tackling the problem by introducing criminal penalties. Countries included in the ranking were more likely to go on to pass laws against
trafficking than those left out; those placed on its watchlist in one year were more likely to do so than those who were not.

The big reason appears to be that governments felt the heat. Media coverage of trafficking grew dramatically in countries covered by the index, but stayed flat elsewhere. The fear of international opprobrium counted, too. A 2010 press release from Pakistan’s Interior Ministry described “significant efforts” to get off the TIP watchlist—efforts it says “improved the stature of Pakistan before the world”.

All this makes TIP a fine example of the performance index as a tool of soft power. It also, for some, demonstrates the risk that indices oversimplify and go further than the data warrant. Its raw figures are second-hand, unreliable and not comparable from country to country, says Neil Howard of the European University Institute in Florence; those for prosecutions, for example, refer to quite different laws in different places. TIP’s influence, he says, “is out of all proportion to the quality of the data it is based on.” Other experts argue that the incentive it creates for countries to criminalise trafficking will not do much where law
enforcement is weak and the economic reasons to migrate for work are strong.

Trafficking is also an example of another trend: the proliferation of indices on similar topics. When Andrew Forrest, one of Australia’s richest men, decided to take on modern-day slavery, Bill Gates had some simple advice for him: find a way to quantify it, because “if you can’t measure it, it doesn’t exist”. The result was the Global Slavery Index, a ranking of over 160 countries based on the prevalence of slavery, broadly defined to include victims of trafficking, forced labourers and child brides.

This ranking received widespread attention, and its estimate of nearly 30m for the total number of people enslaved around the world made global headlines. But it has been heavily criticised. Like TIP, it is based on shaky data, making the decision to “name and shame” the ten worst performers particularly unfair, says Mr Howard. Among the ten is Benin, where his own field studies and interviews with presumed victims suggest that slavery is far less common than the index claims.

And for some countries where no one has tried to estimate the incidence of slavery, figures for others were used instead. Prevalence rates
for Britain were applied to Ireland and Iceland, for example, and those for America, to several western European nations, including Germany. Ronald Weitzer of George Washington University, who has picked through the methodology, describes these substitutions as “bizarre”. Such indices are a “merry-go-round of data that isn’t really data,” says Mr Howard. “The aims may be well-meaning, but sensationalism doesn’t help.”