The European Commission’s science and knowledge service

Joint Research Centre
Ten Steps Guide for Indices & Scoreboards

Michaela Saisana

COIN 2018 - 16th JRC Annual Training on Composite Indicators & Scoreboards
05-07/11/2018, Ispra (IT)
Example of a scoreboard

Social Pillar Scoreboard for the European Pillar of Social Rights

• 3 dimensions & 12 areas
• 14 headline indicators
• 21 secondary indicators
• 93 indicators in total (gender/age)

Example of a scoreboard
Social Pillar Scoreboard for the European Pillar of Social Rights

- 3 dimensions & 12 areas
- 14 headline indicators
- 21 secondary indicators
- 93 indicators in total (gender/age)
Example of a composite indicator

Global Innovation Index

- 1 index
- 2 sub-indices
- 7 pillars
- 21 sub-pillars
- 80+ indicators
Example of a composite indicator

Global Innovation Index

- 1 index
- 2 sub-indices
- 7 pillars
- 21 sub-pillars
- 80+ indicators
Example of a composite indicator

Global Innovation Index

- 1 index
- 2 sub-indices
- 7 pillars
- 21 sub-pillars
- 80+ indicators
Need for indices and scoreboards

What are they?
com’-po-site: made of various parts or elements
in’-di-ca-tor: a device providing information on the state or condition

Why are they needed?
Globalization + Complexity + Data Revolution

Who needs them?
International organizations, European Commission, governments, public
Uptake of indices in the EC policies

1987 First CI proposed by the EC
Proposal for a Council Regulation establishing a Community system of aids to agricultural income [COM(87) 166 final]

1985 Second CI developed by a government
German gov. to select areas eligible for investment aid in the North Rhine-Westphalia [85/12/EEC]

1982 First CI developed by a government
Danish gov. to select areas eligible for regional aid [82/691/EEC]

1973 First reference within an EC document but ...
14th rep. on the activities of the Monetary Committee [OJ 9.11.73 No C94]
Uptake of indices in the EC policies

Search for «composite indicator» OR «synthetic indicator» OR «aggregate indicator»

118 documents (most of them by the EC)
Uptake of indices in the academia

More than ten-fold increase since 2000!
Uptake of indices in the media

Google search results for "human development index" showing articles and news about the Human Development Index (HDI) and its implications in the Caribbean.
Polarized audience

**Enthusiastic supporters**, mostly from advocacy groups developing their own indices to advance a cause

**Skeptical economists and official statisticians** concerned by the subjective nature of the selection of variables, weights and aggregation
Enthusiastic supporters, mostly from advocacy groups developing their own indices to advance a cause

Skeptical economists and official statisticians concerned by the subjective nature of the selection of variables, weights and aggregation

Polarized audience
Indicators are powerful advocacy tools

2010
- A. Forrest decided to take up the cause of combatting human trafficking → he established the Walk Free Foundation

2012
- Bill Gates gave him some advice: “use a quantifiable metric. [...] if you can’t measure it, it doesn’t exist.”

2013
- A. Forrest got Richard Branson on board, and by Dec. 2012, they had appealed to 25 large companies and governments to ban the use of forced labor

- One year later, one activity got more attention than all of Walk Free’s efforts combined...
Indicators are powerful advocacy tools

2013
- September: The Global Slavery Index

2013
- October: The index was covered in >100 newspaper stories around the world

2014
- July: In a global survey of NGOs that work on trafficking issues, over 40% had already heard of this new index
Polarized audience

Enthusiastic supporters, mostly from advocacy groups developing their own indices to advance a cause

Skeptical economists and official statisticians concerned by the subjective nature of the selection of variables, weights and aggregation
Unlike the Alcohol Policy Index, most composite indicators cannot be validated versus a ground truth.

...and on top of that indices have ... strong Political and Policy Implications

- The **Malaysian** Industrial Development Authority insisted that Malaysia aims to move from the 24th to top 10 on in the **World Bank's `Doing Business'** ranking list. "We continue to ask ourselves what it will take to reach the top 10, and are we willing to do what it takes to get there." [Asia in Focus, Jan. 8 2007]

- The Minister of the Economic Development in **Kyrgyzstan**, expressed a hope (in 2008) that his country shall rank among top 20 countries in the **Doing Business rating** in three years.

[The **WB Doing Business Report** has long been credited with bringing about reforms in countries – as many as 2000 distinct reforms since its 2003 launch (Source: The Economist 2013)]
This is WHY composite indicators should be …

developed sensibly and used responsibly

✓ ‘Sensible development’ of a composite indicator implies a quality control process based on both conceptual and statistical considerations.

✓ ‘Responsible use’ calls for care in drawing conclusions and recommendations without taking into account the conceptual context in which composite indicators were developed.
[...Refreshing key statistical terms...]
Variance

[...]Refreshing key statistical terms...]
[..., Refreshing key statistical terms...]

Percentiles
The correlation coefficient only assesses the strength of a linear relationship between two variables.

The correlation coefficient does not assess causation – that one event causes the other.
[...Refreshing key statistical terms...]

### Rule of Thumb for Interpreting the Size of a Correlation Coefficient

<table>
<thead>
<tr>
<th>Size of Correlation</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>.90 to 1.00 (−.90 to −1.00)</td>
<td>Very high positive (negative) correlation</td>
</tr>
<tr>
<td>.70 to .90 (−.70 to −.90)</td>
<td>High positive (negative) correlation</td>
</tr>
<tr>
<td>.50 to .70 (−.50 to −.70)</td>
<td>Moderate positive (negative) correlation</td>
</tr>
<tr>
<td>.30 to .50 (−.30 to −.50)</td>
<td>Low positive (negative) correlation</td>
</tr>
<tr>
<td>.00 to .30 (.00 to −.30)</td>
<td>negligible correlation</td>
</tr>
</tbody>
</table>
[...Refreshing key statistical terms...]

KPI-1

KPI-2
[...Refreshing key statistical terms...]
...Refreshing key statistical terms...

KPI-1 with KPI-2
\[ r = 0.768 \]

(high positive correlation)
[...Refreshing key statistical terms...]

KPI-1 with KPI-2

\[ r = 0.213 \]

(negligible correlation)
Ten Steps Guide for Composite Indicators

- Step 10. Presentation & dissemination
- Step 9. Association with other variables
- Step 8. Back to the indicators
- Step 7. Robustness & sensitivity
- Step 6. Weighting & aggregation
- Step 5. Normalisation of data
- Step 4. Multivariate analysis
- Step 3. Data treatment
- Step 2. Selection of indicators
- Step 1. Developing the framework

Finally endorsed (after 2 rounds of consultation) by the OECD high level statistical committee in March 2008.
Step 1
Theoretical/Conceptual framework

- Definition of the phenomenon
- Added-value of index/scoreboard
- Nested structure of the framework
- Involve experts and stakeholders

Table 5.2. Comparison between the conceptual and measurement frameworks of the Gender Equality Index and the Gender Gap Index

<table>
<thead>
<tr>
<th>Domains</th>
<th>Concepts measured</th>
<th>Concepts measured</th>
<th>Domains</th>
</tr>
</thead>
<tbody>
<tr>
<td>Power</td>
<td>Political decision-making</td>
<td>Minister-level positions</td>
<td>Gender ratio of years in executive office (prime minister or president) for the last 50 years.</td>
</tr>
<tr>
<td>Social</td>
<td>Political decision-making</td>
<td>Parliamentary level</td>
<td>Parliamentary positions</td>
</tr>
<tr>
<td>Economic</td>
<td>Economic decision-making</td>
<td>Members of central banks</td>
<td>Members on boards</td>
</tr>
<tr>
<td>Work</td>
<td>Segregation and quality of work</td>
<td>Work intensity</td>
<td>Advancement gap</td>
</tr>
<tr>
<td>Participation</td>
<td>Participation in the labour market</td>
<td>Economic participation and opportunity</td>
<td></td>
</tr>
<tr>
<td>Money</td>
<td>Financial resources</td>
<td>Earnings, income</td>
<td>Remuneration gap</td>
</tr>
<tr>
<td>Economic</td>
<td>Economic situation</td>
<td>Not at risk of poverty</td>
<td>Income distribution</td>
</tr>
<tr>
<td>Knowledge</td>
<td>Educational attainment and segregation</td>
<td>Literacy rate gap</td>
<td>Educational attainment</td>
</tr>
<tr>
<td>Time</td>
<td>Care activities</td>
<td>Domestic activities</td>
<td>Childcare activities</td>
</tr>
<tr>
<td>Social</td>
<td>Leisure and sport activities</td>
<td>Leisure and sport activities</td>
<td></td>
</tr>
<tr>
<td>Health</td>
<td>Status</td>
<td>Healthy life expectancy</td>
<td>Health and survival</td>
</tr>
<tr>
<td>Violence</td>
<td>Behaviour</td>
<td>Unmet medical needs</td>
<td>Unmet dental needs</td>
</tr>
</tbody>
</table>

Step 2
Data selection

✓ Selection criteria for indicators
  ✓ input, output, outcome, process
  ✓ salience, credibility, legitimacy
  ✓ data coverage
  ✓ consider using proxy variables when official statistics are scarce

✓ Summary statistics
  ✓ Source & data availability (countries, time)
  ✓ type (hard, soft or input, output, process),
  ✓ descriptive statistics (mean, median, skewness, kurtosis, min, max, variance, histogram)

NEARLY HALF OF ALL CANADIANS VOLUNTEERED IN 2004.

According to the Canada Survey of Giving, Volunteering and Participating (CSGV/P), nearly 12 million Canadians over the age of 15 worked a total of nearly 2 billion volunteer hours in 2004. That amounts to 168 hours per volunteer, or the equivalent of 1 million full-time jobs.

Regionally, Saskatchewan residents were more likely to volunteer than residents of any other province—almost 10% more than the national average of 45% (see Chart 1).
Step 2
Data selection

- Discuss strengths and limitations of each selected indicator
- Involve experts and stakeholders
Step 3

Data treatment

- Make **scale adjustments** if necessary (e.g. divide by population, GDP, other)
- Check for **missing data and outliers**
- Treat outliers, if needed (so as to avoid that they become unintended benchmarks)
- Estimate **missing data**, if appropriate (and estimate confidence interval for each imputed value to assess the impact of imputation on the results)
Step 4
Normalisation

- Make directional adjustments (so that higher scores correspond to better performance in all indicators or vice versa)

- Select a suitable normalisation method that respects the conceptual framework and the data properties

Does a large group of foreign-born population decrease social cohesion? (yes/no → direction)

→ 5 seconds rule!
Step 4
Normalisation

- Make **directional adjustments** (so that higher scores correspond to better performance in all indicators or vice versa)

- Select a **suitable normalisation method** that respects the conceptual framework and the data properties
Step 5
Weighting

 ✓ Discuss if *correlation* among indicators should be accounted for in the weights
   ✓ If yes, then how (more correlated indicators more or less weight?)

 ✓ Select a *suitable weighting method* that respects the conceptual framework and the data properties
Step 6

Aggregation

- Discuss if compensability among indicators should be allowed (fully, partially, not at all)
- Discuss up to which level to aggregate
- Select a suitable aggregation method that respects the conceptual framework and the data properties

Outcome

8 + 2 \leq 5 + 5

Outranking matrix

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0</td>
<td>0.2</td>
<td>0.4</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>B</td>
<td>0.8</td>
<td>0</td>
<td>0.8</td>
<td>1</td>
<td>0.2</td>
</tr>
<tr>
<td>C</td>
<td>0.6</td>
<td>0.2</td>
<td>0</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>D</td>
<td>0.8</td>
<td>0</td>
<td>0.6</td>
<td>0</td>
<td>0.2</td>
</tr>
<tr>
<td>E</td>
<td>0.8</td>
<td>0.8</td>
<td>0.6</td>
<td>0.8</td>
<td>0</td>
</tr>
</tbody>
</table>
17.1) During the last 12 months, for how many months was your household’s main source of water sufficient to meet your household’s drinking, cooking, bathing and cleaning needs?

| Months: | ≤ 4 | ≥ 8 | Don’t remember (-1) |

17.2) How often do you worry there will not be enough water from your household’s main water source to satisfy your household’s drinking, cooking, bathing and cleaning needs?

| Never (1) | Rarely (2) | Sometimes (3) | Often (4) | Always (5) |

- 3 HHs Careless?
- 104 HHs Too worried?

**Suggestion:** With survey data, given some unavoidable inconsistencies (in part due to the way the human mind works), use a (weighted) **arithmetic average** (rule of thumb: 5-10 indicators) within a subcomponent.
Step 7
Statistical coherence

- Assess if few indicators dominate the framework
  \((r_{\text{indicator,index}}>0.95)\)

- Assess if indicators behave as “noise” in the framework
  \((-0.3<r_{\text{indicator,index}}<0.3)\)

- Assess if indicators are negatively related to the index
  \((r_{\text{indicator,index}}<-0.3)\)

### Pearson Correlations

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Ecosystem Vitality</th>
<th>Environmental Health</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>EPI 2010</strong></td>
<td>0.29</td>
<td>0.90</td>
</tr>
<tr>
<td>Environmental Health</td>
<td>-0.08*</td>
<td></td>
</tr>
</tbody>
</table>

| Source | 117 | 615 |
Step 7
Statistical coherence

✓ ...
✓ Assess if indicators statistically fit better under different dimensions than those in the framework
✓ Assess if dimensions should be merged or split
✓ Assess if bias has been introduced in the index (e.g., due to population size, population density, GDP)
Step 8
Robustness & Sensitivity

- Identify the **sources of uncertainty** in the index development
- Assess the **impact of the uncertainties** to the index scores or ranks
  - complement scores/ranks with confidence intervals

- **Including/excluding variables**
- **Weights**
- **Normalisation**
- **Missing data**

Country 1 Country 2 Country 3

10 20 30 40 50 60

**JRC-COIN © | Ten Steps Guide for Indices & Scoreboards**
Step 8
Robustness & Sensitivity

✓ Identify the **sources of uncertainty** in the index development

✓ Assess the **impact of the uncertainties** to the index scores or ranks
  ✓ complement scores/ranks with confidence intervals
Step 8
Robustness & Sensitivity

✅ ... 
✅ Identify which uncertainties are more crucial in determining the final classification

- Explain why certain countries notably improve or deteriorate their relative position given the assumptions

Robustness ≠ Quality
Robustness ≠ Quality

ARWU 2008 ranking
(more robust, higher correlations)

THES 2008 ranking
(less robust, lower correlations)
Step 9

Look back into the data & vis-à-vis other measures

✓ Decompose performance at the indicator level (to reveal strengths and limitations for each country)

✓ Correlate the index with relevant measurable phenomena and explain similarities or differences

✓ Develop evidence based narratives

✓ Perform causality tests (if time series data or microdata are available)
Step 9

Look back into the data & vis-à-vis other measures

✓ Decompose performance at the indicator level (to reveal strengths and limitations for each country)
✓ Correlate the index with relevant measurable phenomena and explain similarities or differences
✓ Develop data-driven narratives on the results.
✓ Perform causality tests (if time series data or microdata are available)
Step 10

Visualisation & Communication

- Identify suitable visualisation tools for the targeted audience
- Select the visualisation technique which communicates the most information without hiding vital information
- Make your index/scoreboard EAST
  - Easy, Attractive, Social and Timely
Prerequisites for any index

- Carefully selected indicators
- Sound framework
- Sound model

Can help to depict reasonably reality
Prerequisites for any index

- Carefully selected indicators
- Sound framework
- Sound model

can only offer an imperfect mirror of reality
Powerful evidence based narratives supported by good statistical measures and good analytic work are a possibility which should not be left untried.

We need **relevant** and **sound**...
References and related reading

References and related reading

References and related reading

Welcome to email us at: jrc-coin@ec.europa.eu

**COIN in the EU Science Hub**

**COIN tools are available at:**
https://composite-indicators.jrc.ec.europa.eu/
Questionnaire – Post Day 1:

https://goo.gl/qy19yD