



The European Commission's science and knowledge service

Joint Research Centre

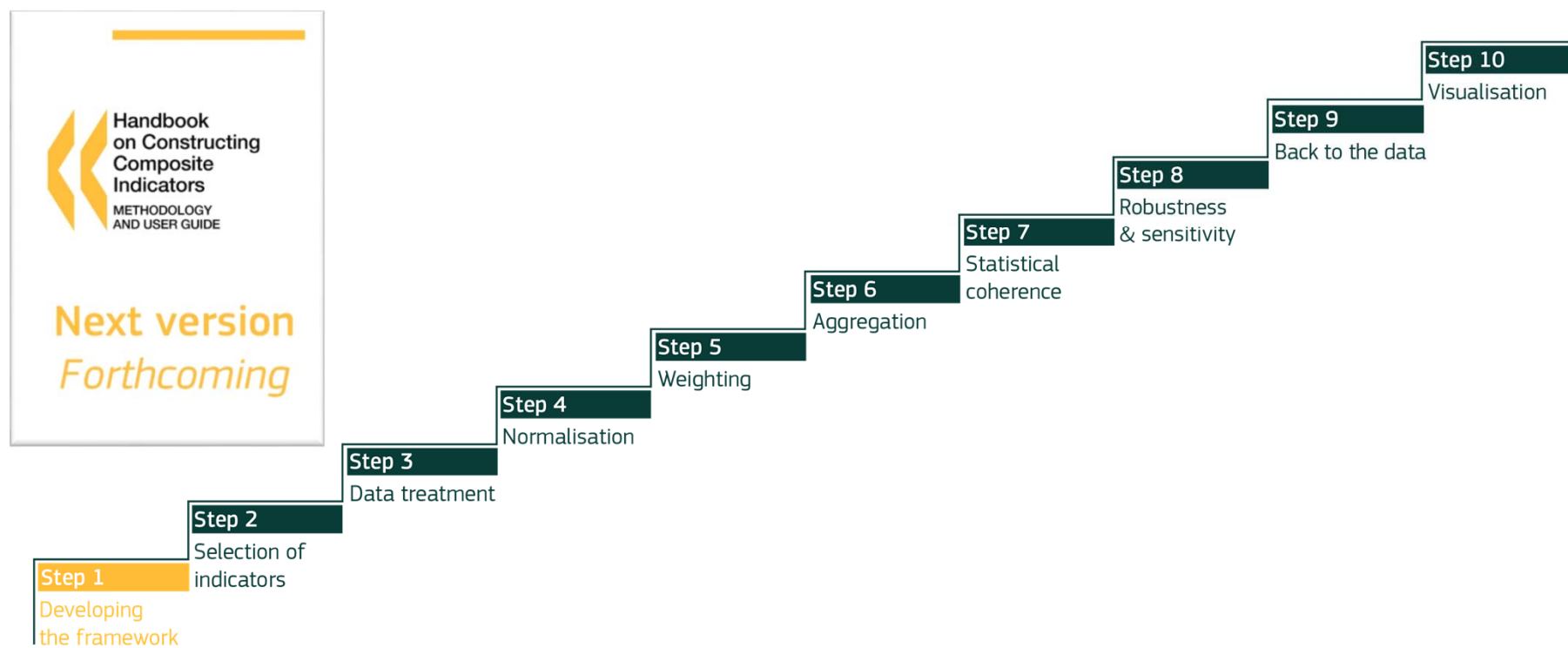
Step 1 & 2

Frameworks and Indicators

Dániel Vértessy

COIN 2019 - 17th JRC Annual Training on Composite Indicators & Scoreboards
04-06/11/2019, Ispra (IT)

Ten steps





We need an Index
A.S.A.P.

We need to revise
our Index A.S.A.P.

3 questions to clarify



What do we want to measure?

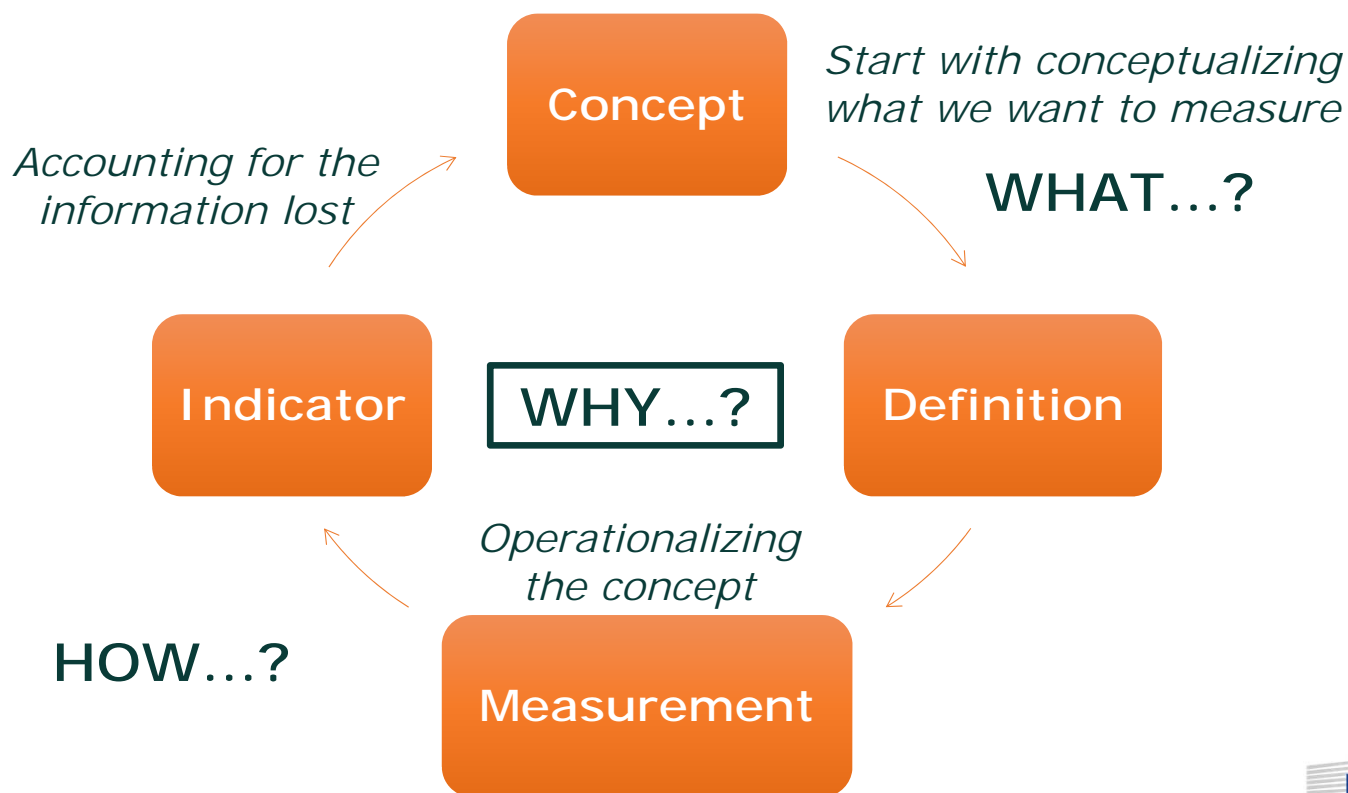


Why do we want to measure it?



How do we want to measure it?

Operationalizing the Concept



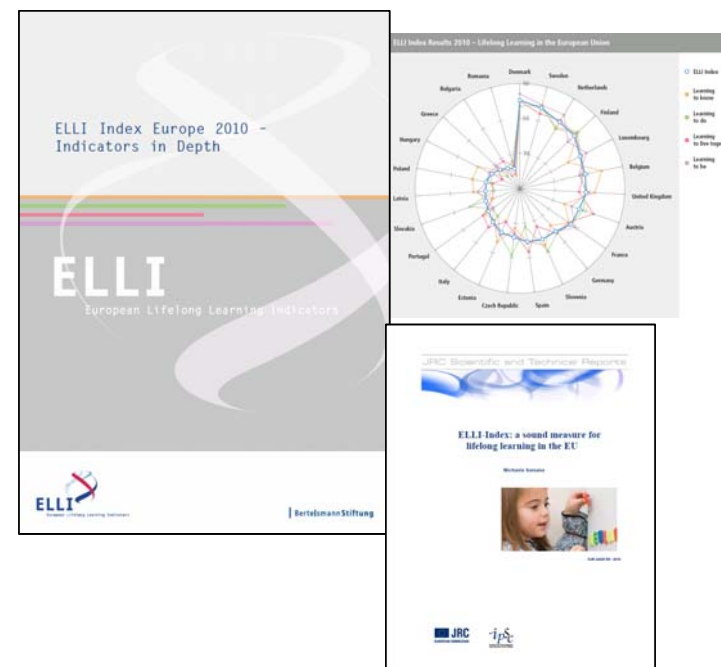
Composite Indicators of Lifelong Learning

CLI (2007, 2010)



Saisana, 2008

ELLI (2010)

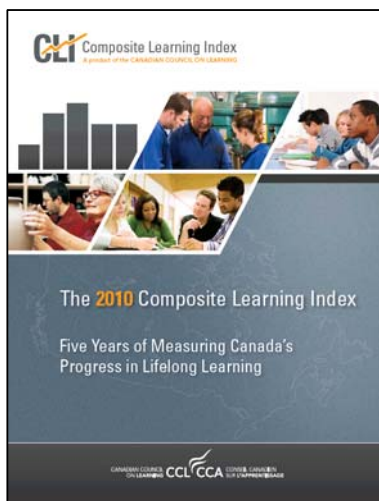


Saisana, 2010

Composite Indicators of Lifelong Learning

The **Composite Learning Index (CLI)** is a product of the Canadian Council on Learning. [It] is Canada's annual measure of progress in lifelong learning.

“It is based on a combination of statistical indicators that reflect the many ways Canadians learn, whether in school, in the home, at work or within the community.”



- takes a **holistic approach to learning** that encompasses more than what occurs in the classroom.
- LLL = a philosophy that involves the development of **knowledge, skills and values**,
- throughout **all stages of a person's life** —from early childhood through adulthood.”

3 questions to clarify



What do we want to measure?

- LIFELONG LEARNING: multi-dimensional in nature; no obvious single indicator
 - Knowledge, skills, values;



Why do we want to measure it?

- To *measure* progress over time
- To *benchmark* communities (cities) across Canada
- To *advocate* how learning is critical to the success of individuals, communities & the country



How do we want to measure it?

- Using the UNESCO / Delors Framework

The role of a Theoretical Framework

- Guides the choice of pillars (or dimensions); weights; aggregation methods, etc.
- Supports the interpretation of results (and potential trade-offs)
- Command stakeholder respect

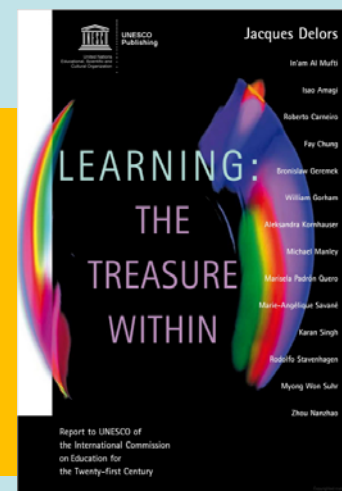
➔ In sum: helps answer the What/Why/How questions



A four-pillar framework on learning in the 21st century:

- | | |
|-----------------------------|---------------------------|
| • Learning to know | – formal education system |
| • Learning to do | – vocational learning |
| • Learning to live together | – social cohesion |
| • Learning to be | – personal development |

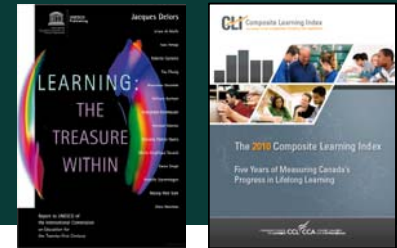
(different motivations, infrastructural needs, etc.)



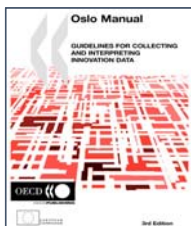
? HOW? – Typical questions to ask

- Why is LLL important for the individual / for businesses / for policy?
When should the state intervene?
- What are the differences vis-à-vis closely related, established concepts?
(= value added of a new index?)
 - i.e., Programme for the International Assessment of Adult Competencies (PIAAC)
- Measure learning... The inputs? The process? The outcomes?
- Links between individual and societal levels?
- How to compare learning happening in formal / informal / non-formal education?
- How to distinguish quantity vs. quality?
- What comprises 'good' performance, and how can it be measured?
- What kind of data to use? (consistent, internationally comparable)
- What countries & years to cover?

CLI [& ELLI] in a structural model



An example: What is innovation?



Oslo Manual Definition: “the implementation of a new or significantly improved product (good or service), or process, a new marketing method, or a new organizational method in business practices, workplace organisation or external relations”

<-> New combinations [of knowledge, capabilities, resources, skills]?

<-> What happens in Households?

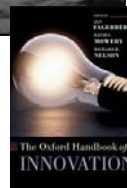
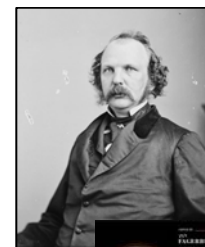
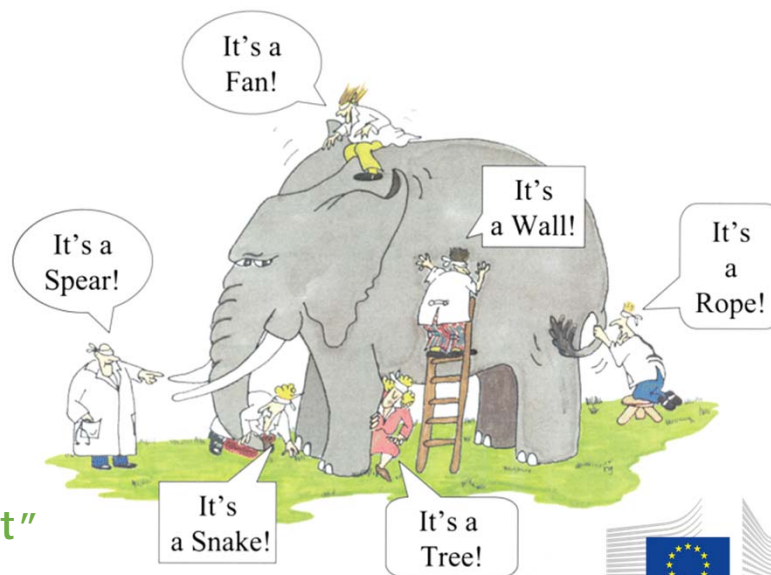
Degree of novelty?

- Radical Innovations?
- Incremental Innovations?
- Diffusion of innovations?

Who defines what is innovative?

- The producer?
- The consumer?
- The researcher/expert?

J.G. Saxe's “The Blind Men and the Elephant”



How do you need to measure innovation?

Measure the effort? [R&D target – 3%]

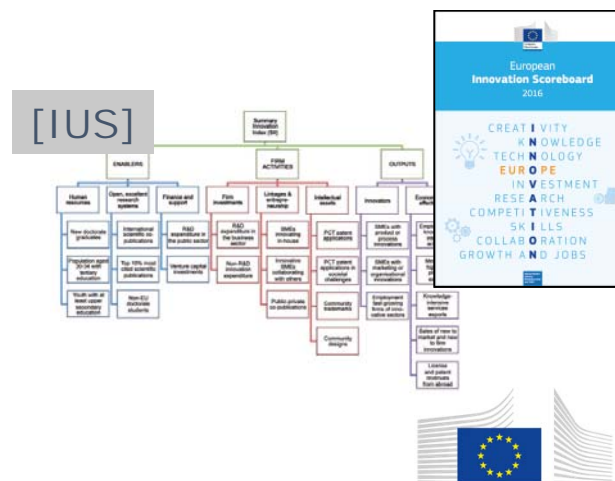
Measure the complexity? [GII]



- I.e., The GII [aims to] to better capture the richness of innovation in society and go beyond such traditional measures of innovation as the number of research articles and the level of R&D expenditures

Measure the performance of innovation systems? [IUS]

Measure the output? [IOI]



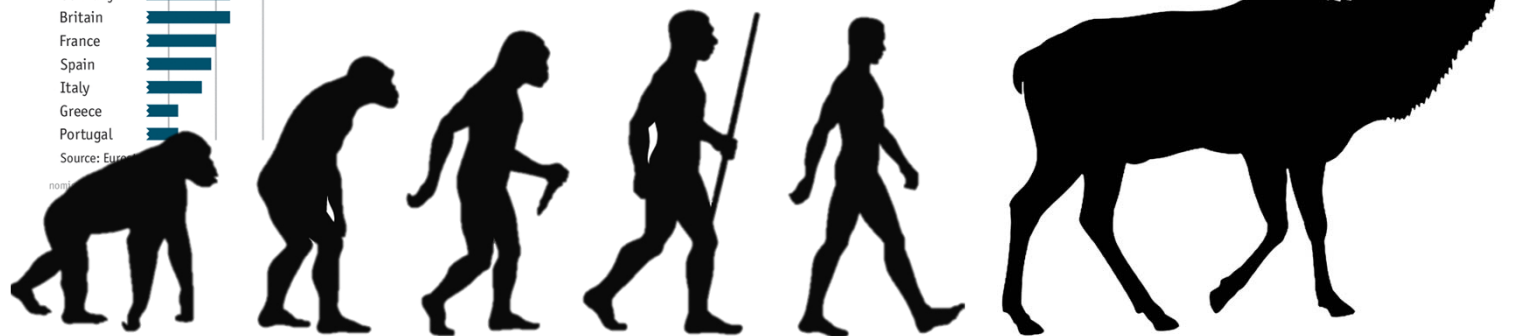
What's the added value of a new index?

Happiness in Europe

Self-reported, 2013, 10=most happy



Source: Eurostat

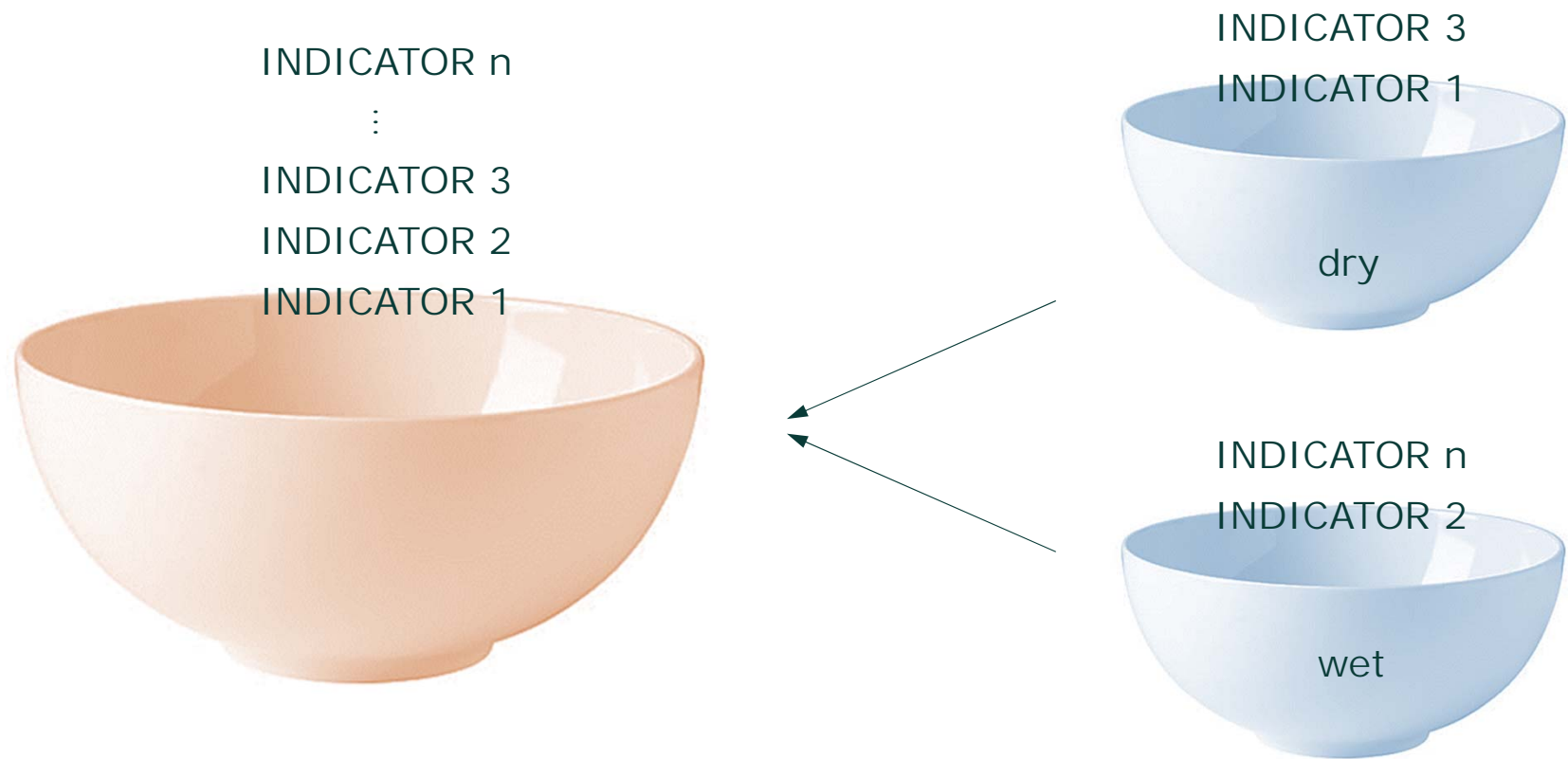


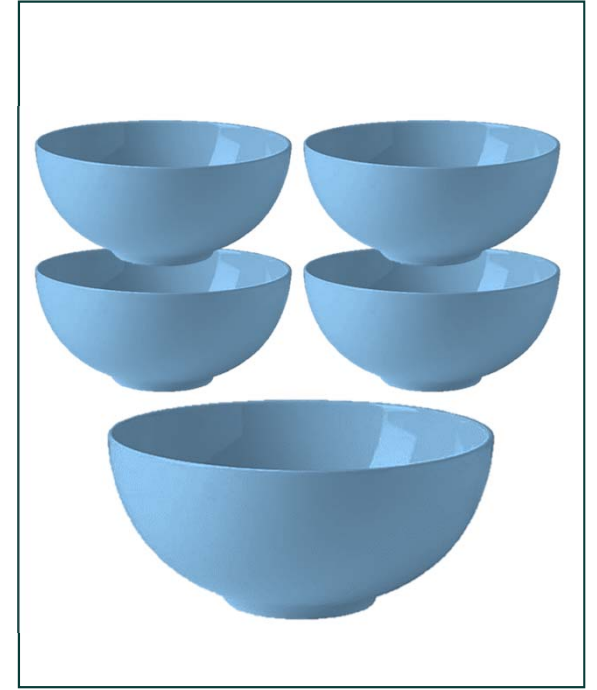
= How
Scandinavian
are you?

What GDP does not show...



from Conceptual Frameworks to Indicator Frameworks





How to find the most fitting framework?

EXAMPLE: Design an index to answer the following question:

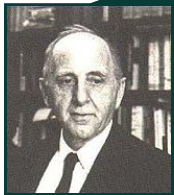
 Is the quality of our life improving?

The unavoidable discussions:

- How to define "quality of our life"? → Wealth? Or Well-being?
- What do we consider "improvement"? → Progress? Development of capabilities?
- What aspects or dimensions should we consider?

A very familiar discussion...

GDP



Measure wealth
(production)
single measure
(S. Kuznets)

Human Development



Human Development
Index (UNDP)

- 4 indicators,
- 3 dimensions



Capabilities and
choices: *no
single indicator!*
(A. Sen)

Well-being

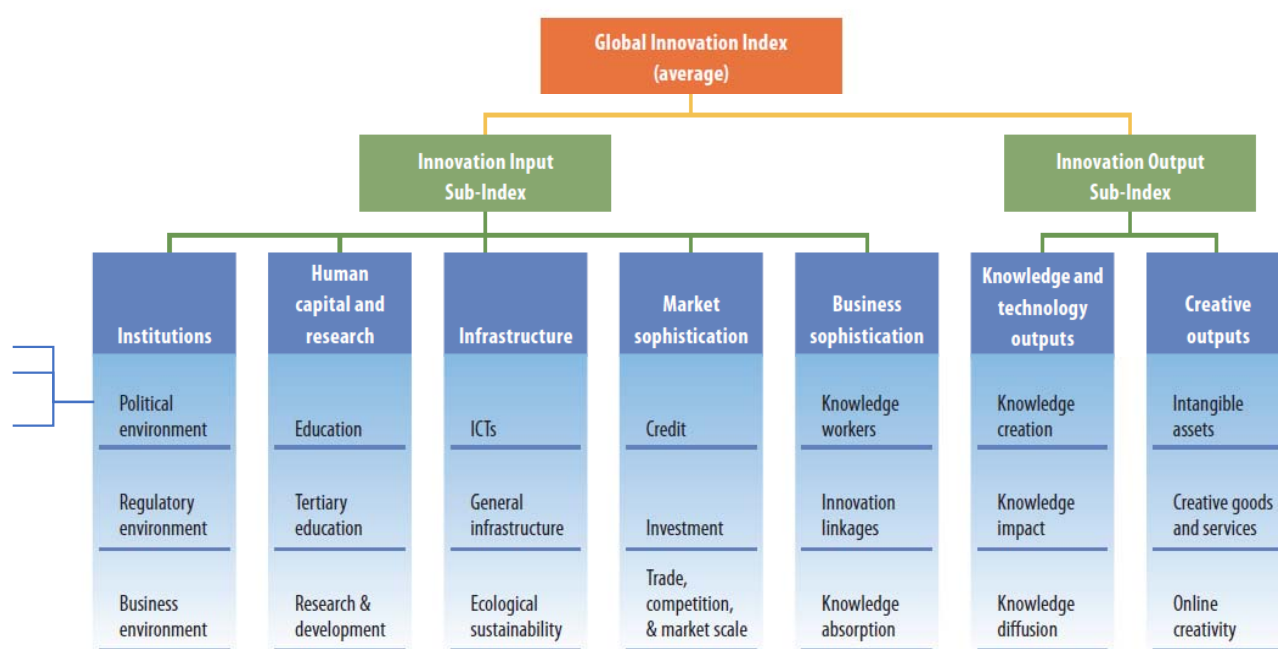
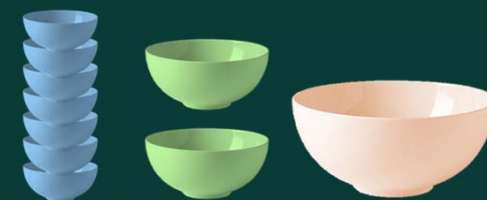


Better Life Index
(OECD)

- 50 indicators,
- 11 dimensions

"There is more to life than
the cold numbers of GDP
and economic statistics"

Multi-dimensional Frameworks



Global Innovation Index (WIPO, Cornell, INSEAD):

- 1 index
- 2 sub-indices
- 7 pillars
- 21 sub-pillars
- 82 indicators

Conceptually, do each point in the same “direction”? Does theory suggest any trade-offs?

➔ *Important for the “statistical coherence” step!*

Quality composite indicators

- Quality = accuracy?
[closeness of computation to the “exact true values”]
- Quality = fitness for use?
[depends on user needs, values, priorities]

- Quality assessment frameworks:

for Official Statistics



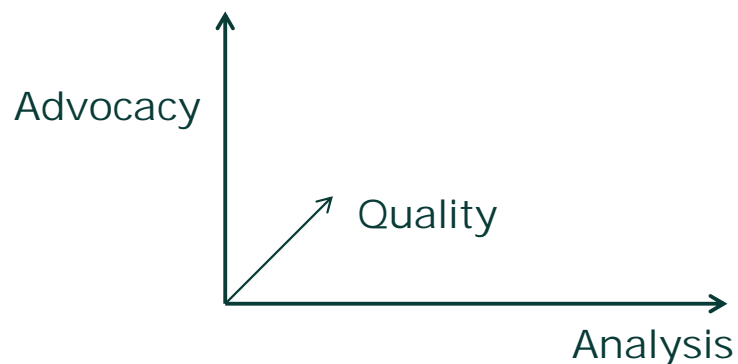
Figure 1. Quality Dimensions

for Composite Indicator Frameworks



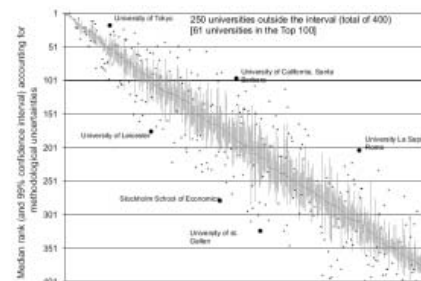
The Quality of Composite Indicators

“Composite indicators sit between **analysis** and **advocacy**, but **quality** discriminates the plausible from the rhetorical” (Saltelli, 2007)



Codified and continuously refined methodology

- The OECD-JRC Handbook (JRC-OECD, 2008)
- Audits – robustness and sensitivity analyses (i.e. Saisana et al, 2011; Paruolo et al, 2013)



Quality of Composite Indicators

- Statistical & conceptual – technical and normative aspects hard to separate
 - Audits help analyze developers' choices made with respect to:
 - the concept;
 - the operationalization process;
 - accounting for information loss...and thus help interpret results
 - The **time** dimension and the **paradox** of assessment:
 - Measurement should improve over time
 - Indicator development is a **dynamic** process (see Barré, 2010; JRC-OECD 2008; refinement of indicators)
- => Audits may need to be repeated in light of the evolution of an indicator!

The “Gisselquist framework”



Rachel M. Gisselquist

Blog

What Does Good Governance Mean?

Rachel M. Gisselquist



A framework of 10 questions to guide the development and evaluation of composite indexes

(Developed for governance indicators)

Focuses on:

- Concept, definition, operationalization
- + Data + Quality of methodology

Critical Questions:

1. What precisely does it aim to measure?
2. Content validity: does the operational definition capture the concept?
3. How good (reliable, valid and complete) are the data used?
4. Is the measure (including all of its sub-components) transparent and replicable?
5. How sensitive and robust is the measure to different data and design choices?
6. Does the measure allow the analyst to address key questions of interest?

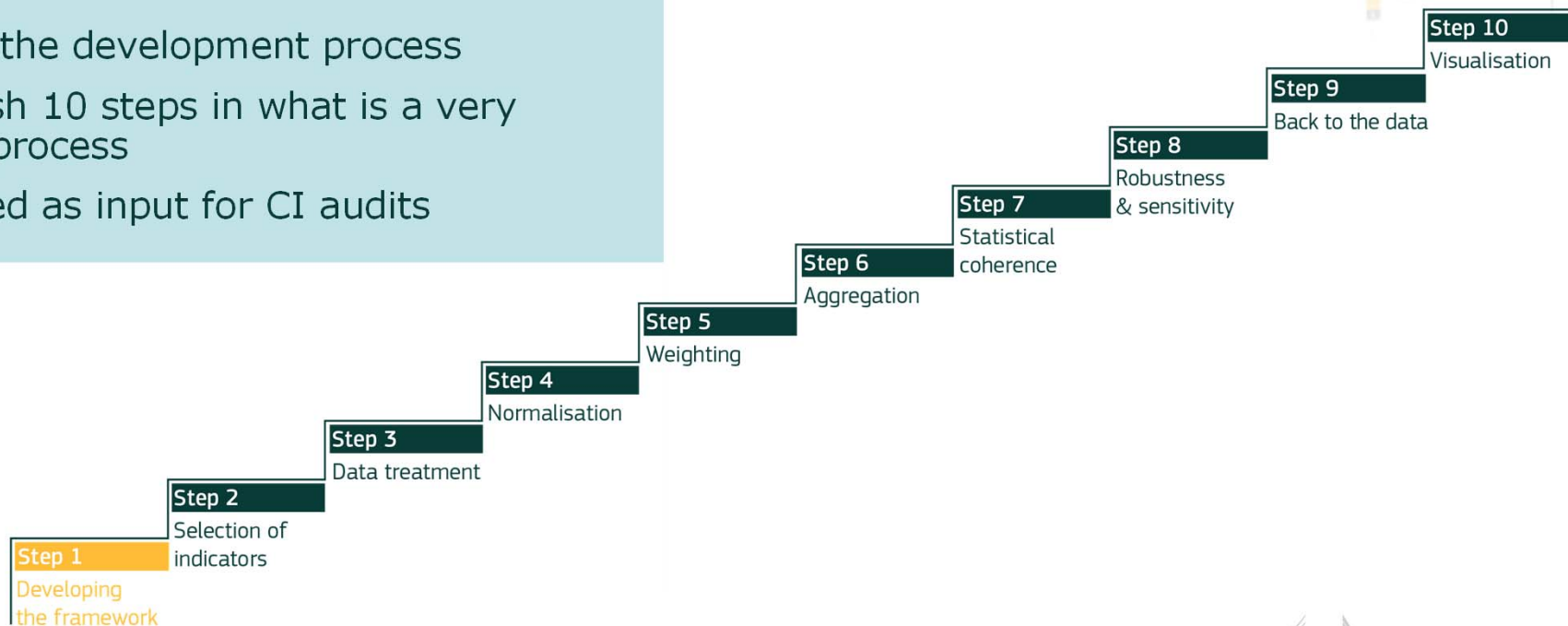
Less Critical Questions:

7. Does the measure fully capture [the concept of interest] in all its complexity? [descriptive compl.]
8. Does the measure behave as theory predicts? [theoretical fit]
9. How precise are index values and are confidence intervals specified? [precision of estimates]
10. Is the weighting ‘correct’?

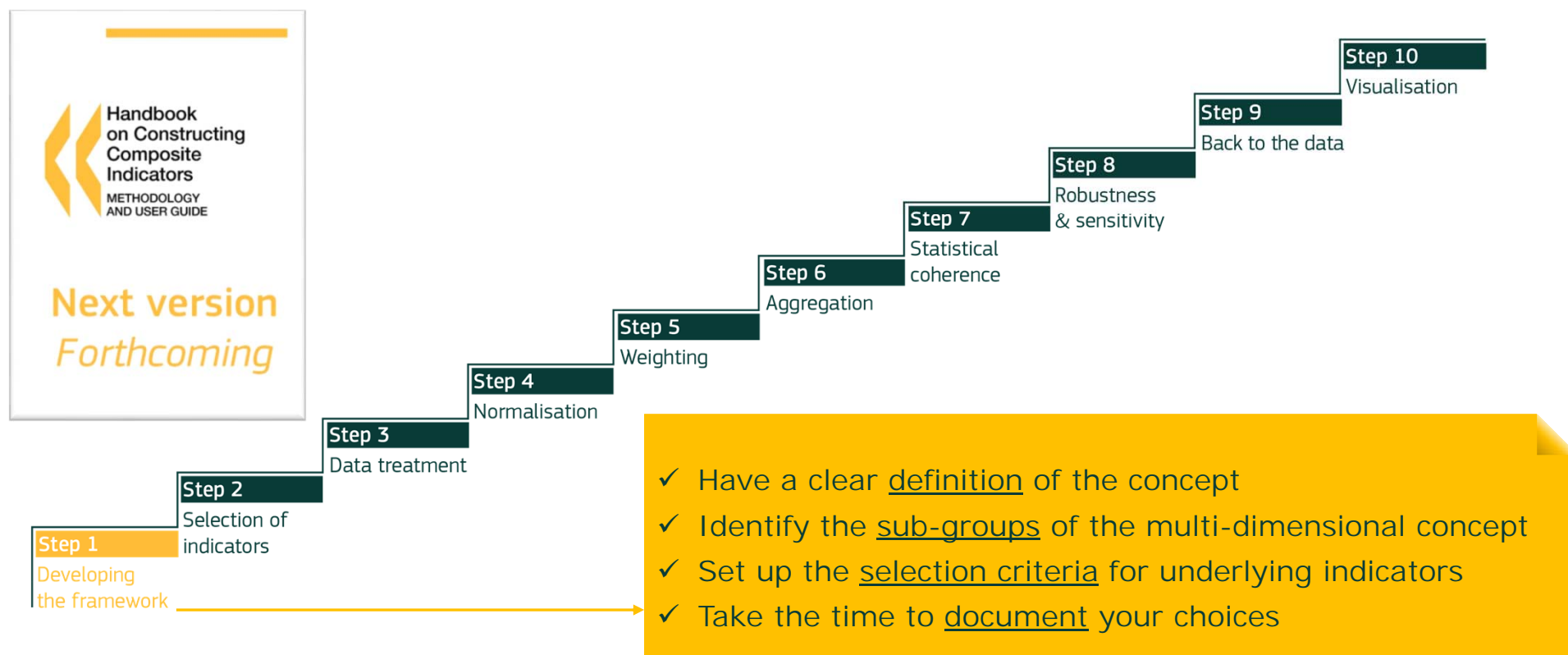
(Gisselquist, 2014)

The 10-step methodology

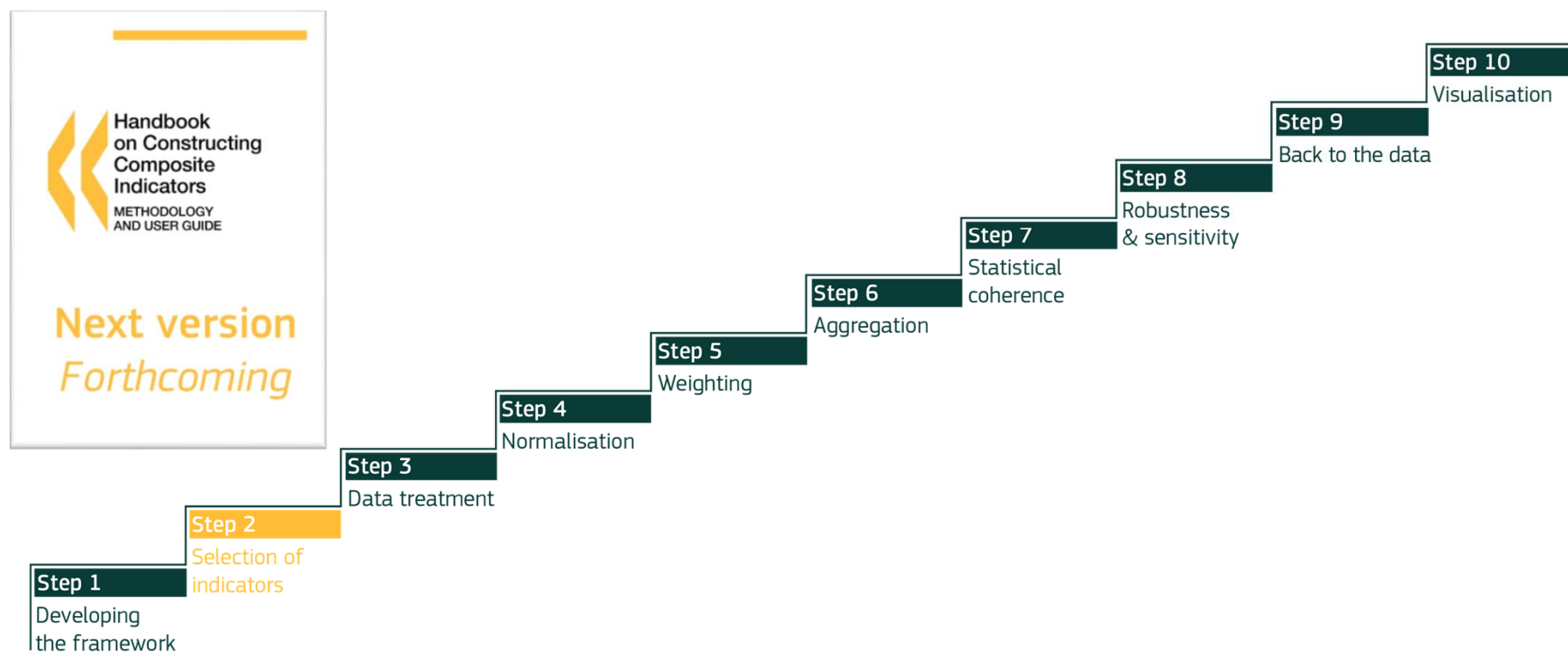
- Focus on the development process
- Distinguish 10 steps in what is a very iterative process
- Has served as input for CI audits



Step 1 Recap: Developing a Theoretical Framework



Step 2: Selection of indicators

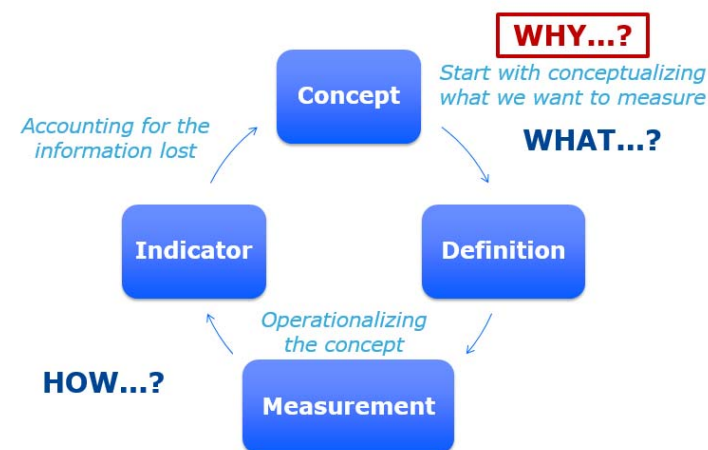


Indicators

How to find the right ones?

The Art of Quantification

- Quantification (modelling) involves making **normative choices** about...
- Normative choices affect:
 - the concept;
 - the operationalization;
 - accounting for information loss



⇒ Composite and stand-alone indicators are alike

Quantifying “essentially contested concepts”

- Example: develop an indicator of “research excellence”

complex framework,
(multiple dimensions,
Inclusion of interactions)



A simple framework,
(only top-level output
measures)

Debate with strong pro/con scientific arguments, until we realized:

- “Excellence” is similar to concepts such as “a work of art”, “fairness”, “social justice” —
 - groups of people disagree about the proper use
 - different uses serve different functions;
 - even after the variety of functions is disclosed, disputes continue — these are **essentially contested concepts!**
- There is “no one clearly definable general use” or “correct use”

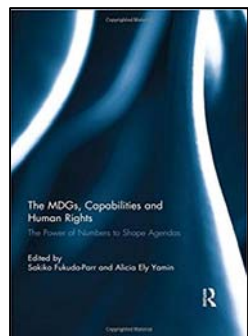


W. B. Gallie (1956)

Quantifying “essentially contested concepts” – Implications

- Recognize that a concept is essentially contested:
 - Accept that “rival uses” of the concept are not just likely, but represent a permanent critical challenge
 - Easy to underestimate opponents’ positions!
- Two sides of the coin:
 - Disputes improve the quality of arguments – and hence, improve measurement
 - False hope that one side will ultimately persuade and convert all opponents: leads to the *“...ruthless decision to cut the cackle, to damn the heretics and to exterminate the unwanted”* – and hence, hide important aspects & functions, and weakens measurement

The power of numbers



Numbers = targets: strong advocacy power
- i.e., **Millennium Development Goals (MDGs)**

Quantification:

- Makes global goals concrete... [SMART...]
- it also involves **simplification**, **reification** and **abstraction**, which have far-reaching implications for redefining priorities"
- All MDG goals/targets "led to **unintended consequences**
 - diverting attention from other important objectives and
 - reshaping development thinking



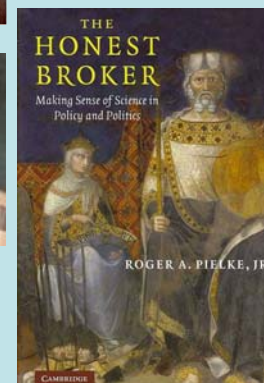
(Fukuda-Parr, S., Yamin, A.E., Greenstein, J., 2014)

Indicators are shaped by policy needs & discourse
Indicators, in turn, influence policy discourse



Indicators are embedded in a socio-political context

- Indicators & indicator frameworks are **value laden**; reflect policy discourse (Godin, 2002)
- Indicator developers & users should be aware of the consequences:
 - The “agora model” (Barré, 2001, 2010): indicators are debating devices – it’s the process that matters!
 - Be an “Honest broker” (Pielke, 2007)



Be an honest broker of Composite Indicators

Composite indicators

- Are very widely used
- Fill a demand for which there is no other alternative

Therefore we should use available tools to **increase robustness and credibility**:

1. **Transparency** — detailed description of methodology, data sources, assumptions
2. **Statistical soundness** — analysis of correlations, data structure, effects of weights, etc.
3. **Uncertainty and sensitivity analysis** — check effect of alternative but plausible assumptions. Honestly acknowledge uncertainty.

Selecting Indicators

Selecting indicators

- Assess the quality of available indicators
 - Quality in general vs fitness for purpose?
 - Trade-offs are central to the development of every indicator → see NQAF
- Discuss strength and weakness of selected indicators
- Provide a summary table of key characteristics
 - Coverage (across time & space)
 - Source
 - Type (hard or soft measures; input / process / output?)

The UN's National Quality Assurance Framework

- Refers to development process & results

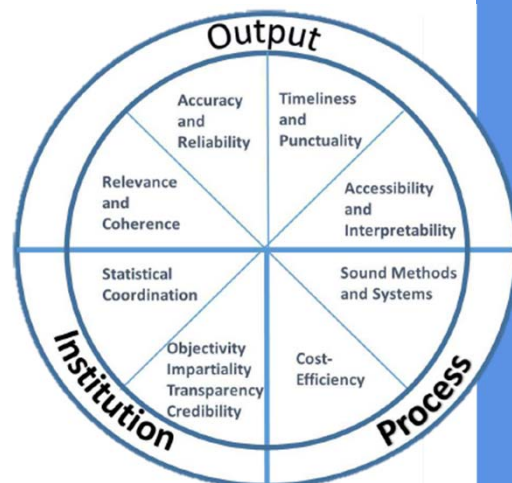


Figure 1. Quality Dimensions

See also: European Statistics Code of Practice (CoP) & ESS Quality Assurance Framework (QAF)



Statistics Division
Towards a National Quality Assurance Framework: the UN Statistical Commission Initiative



Managing statistical processes

- Assuring methodological soundness
- Assuring cost-effectiveness
- Assuring soundness of implementation
- Managing the respondent burden

Managing statistical outputs

- Assuring relevance
- Assuring accuracy and reliability
- Assuring timeliness and punctuality
- Assuring accessibility and clarity
- Assuring coherence and comparability
- Managing metadata

Selecting indicators

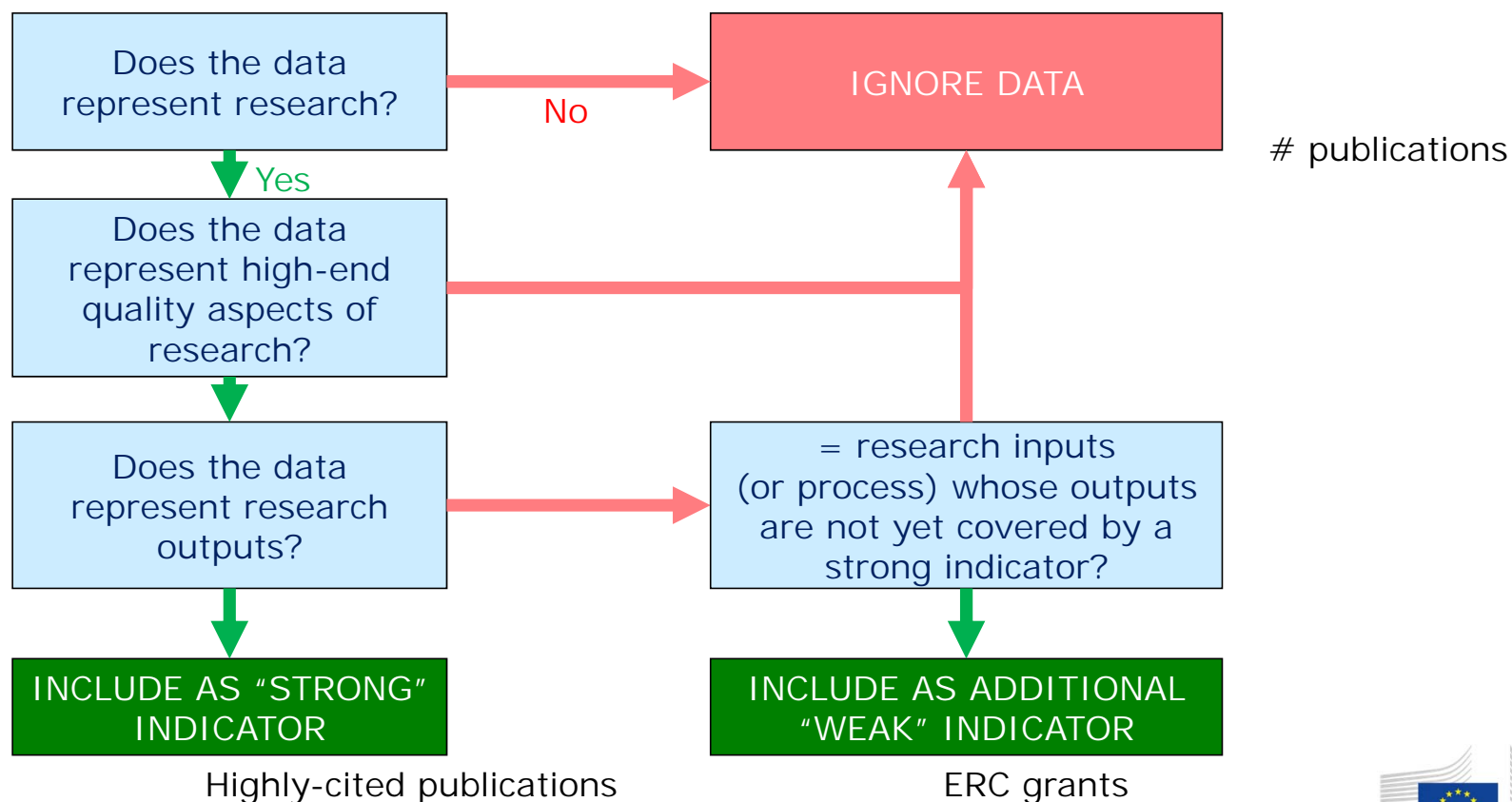
- Assess the quality of available indicators
 - Quality in general vs fitness for purpose?
 - Trade-offs are central to the development of every indicator → see NQAF
- Discuss strength and weakness of selected indicators
- Provide a summary table of key characteristics
 - Coverage (across time & space)
 - Source
 - Type (hard or soft measures; input / process / output?)

An example: selecting 'research excellence' indicators

- From an originally broad to a narrow framework
(see [Ferretti et al, 2017 forthcoming](#); [Hardeman and Vertesy, 2016](#) [Sorensen et al, 2015](#))
- Current use for monitoring progress to ERA:
 - how effectively do national research systems function?
- Research Excellence **Definition** = knowledge producing activities whose outputs are considered of high-end quality (based on [Tijssen, 2003](#))
 - Output oriented
 - Focus on quality, not quantity
 - Focus on top-end, not average
 - Distinguish Science and Technology (evaluated by peers vs. by market)



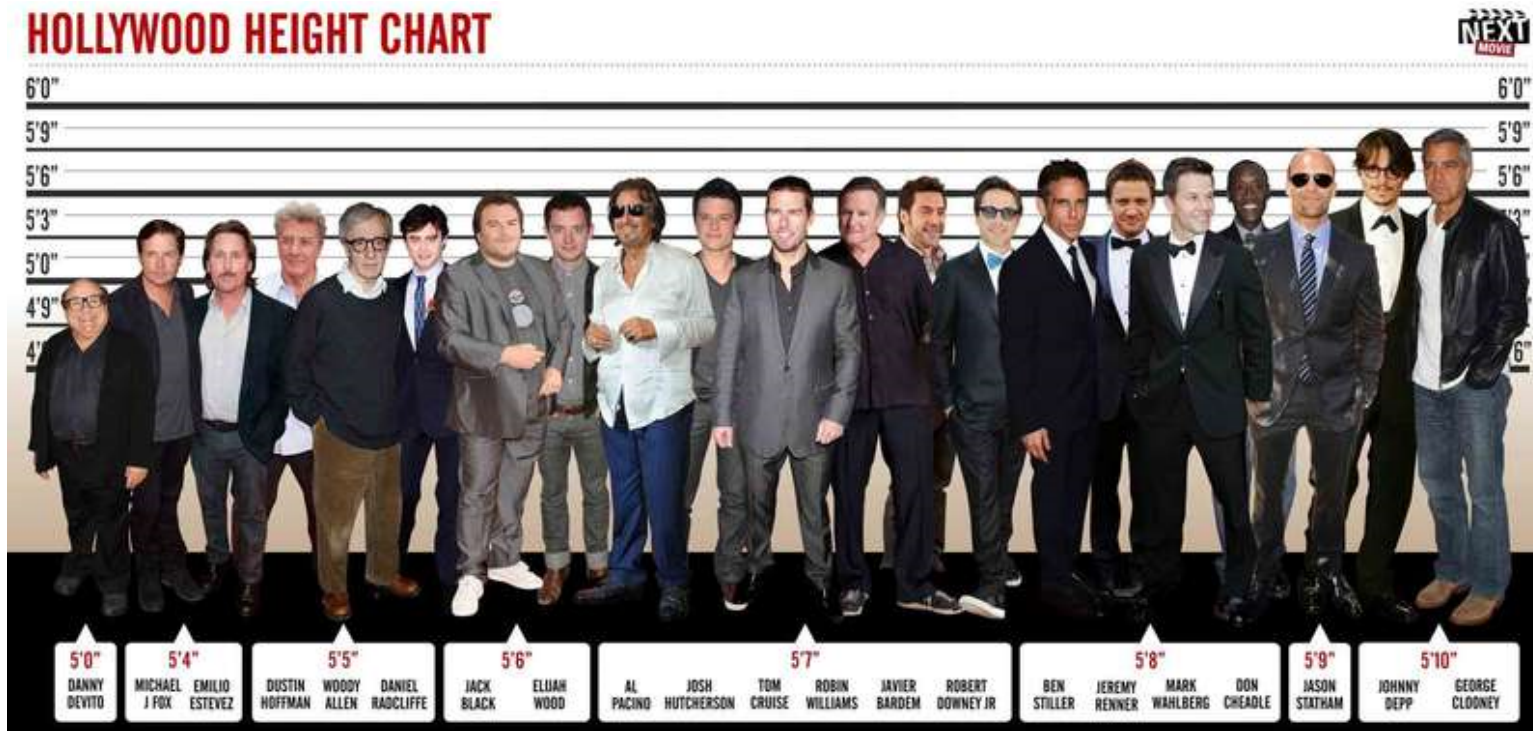
Flow chart for selecting indicators:



Selecting indicators: practical challenges

- A. Use available data (i.e., official statistics)
 - Was it collected for a similar purpose?
 - Was it collected for another purpose but is relevant?
- B. Generate new data
 - Run **surveys** => costs; coverage; replicability
 - Build from **microdata** => cost (also of replicability)
 - Exploit **Big Data** (or admin data)
 - => If desired indicators or desired granularity not available (cost, replicability)
- C. Combine different sources
 - Consider costs (€, time), ease of interpretation (intuitive?)
 - Spell out the desired quality for the indicators to collect!
 - Can distinguish country performance? Missing data acceptable?

Lamppost Bias: Select what is easy to measure?



Selecting meaningful indicators

| Country | Score | |
|---------|-------|---------|
| A | 3 | good |
| B | 2 | average |
| C | 2 | average |
| D | 2 | average |
| E | 2 | average |
| F | 2 | average |
| G | 2 | average |
| H | 2 | average |
| I | 2 | average |
| J | 2 | average |
| K | 2 | average |
| L | 1 | bad |

Cannot distinguish performance!

| Country | Score |
|---------|-------|
| A | 300.0 |
| B | 3.0 |
| C | 2.7 |
| D | 2.6 |
| E | 2.5 |
| F | 2.2 |
| G | 2.2 |
| H | 2.2 |
| I | 2.2 |
| J | 2.2 |
| K | 2.1 |
| L | 2.0 |

→ True value or a clerical error?

If true...
→ see
"outliers"

| Country | Score |
|---------|-------|
| A | 1.5 |
| B | 2 |
| C | n.a. |
| D | n.a. |
| E | n.a. |
| F | 3 |
| G | n.a. |
| H | n.a. |
| I | n.a. |
| J | 5 |
| K | 6 |
| L | 2 |

50% missing!

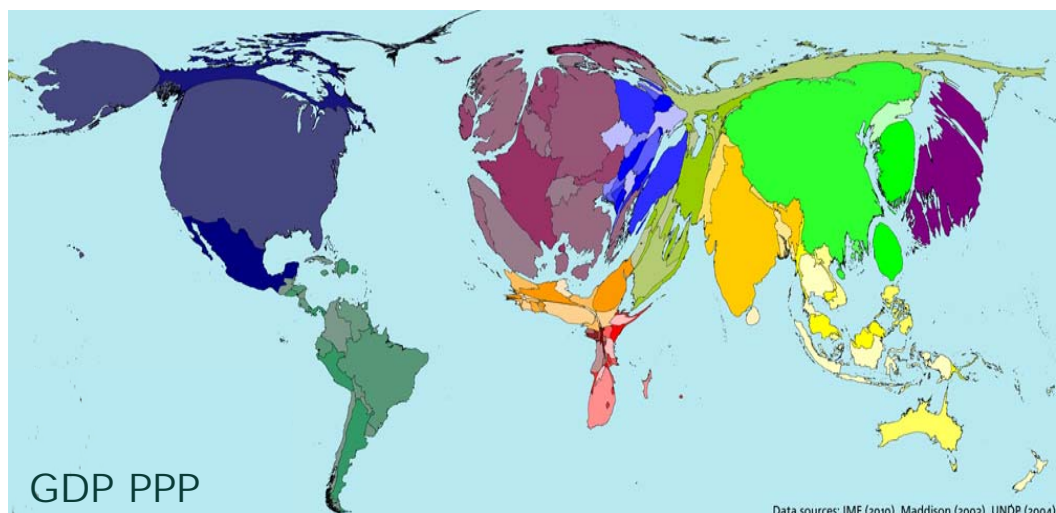
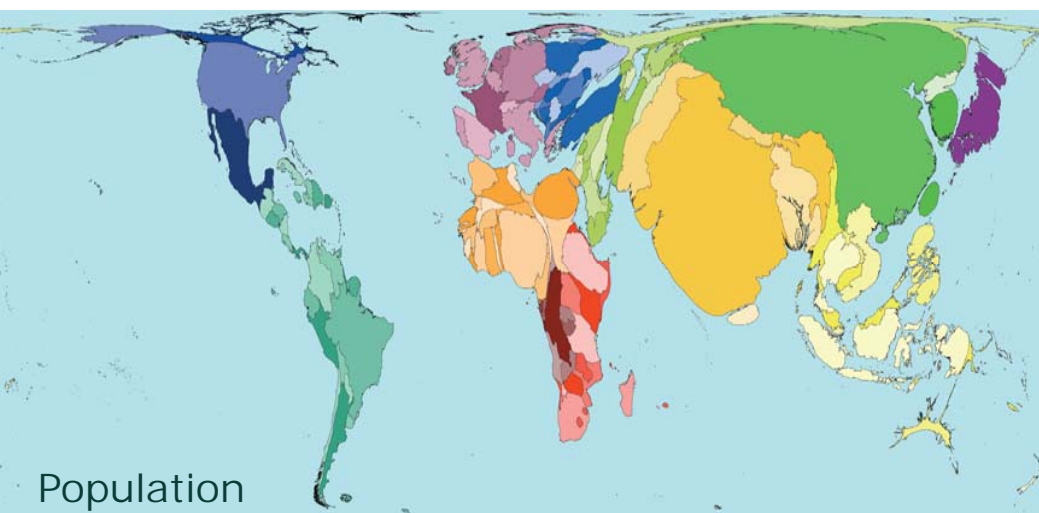
Typical trade-offs to tackle

- timeliness vs. completeness
- quality vs. breadth of coverage
- Novelty vs. acceptance
- Sophistication vs. intuition; reproducibility
- Changes in meaning at different levels of aggregation
- Dealing with units of different size:
 - Choice of denominator not straight forward!

| | Years | | | | | | | | | | | |
|-----------|-------|------|------|------|------|------|------|------|------|------|------|------|
| Indicator | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 |
| HICIT | x | x | x | x | x | x | x | x | x | | | |
| TOPINST | | | | | | ? | ? | ? | ? | ? | | |
| PCT | (x) | (x) | (x) | (x) | (x) | (x) | x | x | x | x | x | |
| ERC | | | | | | | | | (x) | x | x | x |
| | | | | | | | | | | | | |
| | | | | | | | | | | | | |
| | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 |
| HICIT | x | x | x | x | x | x | x | x | x | | | |
| TOPINST | | | | | | ? | ? | ? | ? | ? | | |
| PCT | (x) | (x) | (x) | (x) | (x) | (x) | x | x | x | x | x | |

The denominator problem

- How to avoid measuring 'size' when comparing units of different size?
 - i.e., EU vs. China?

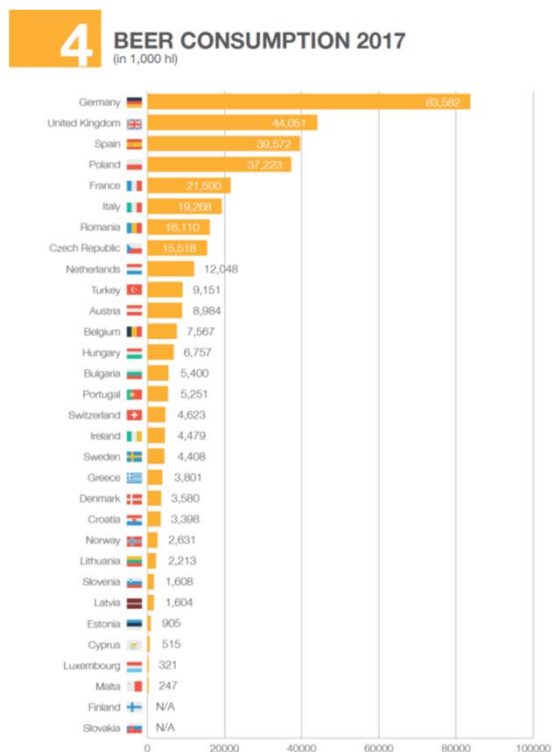


Meaning of the denominator → Reference population

The denominator problem

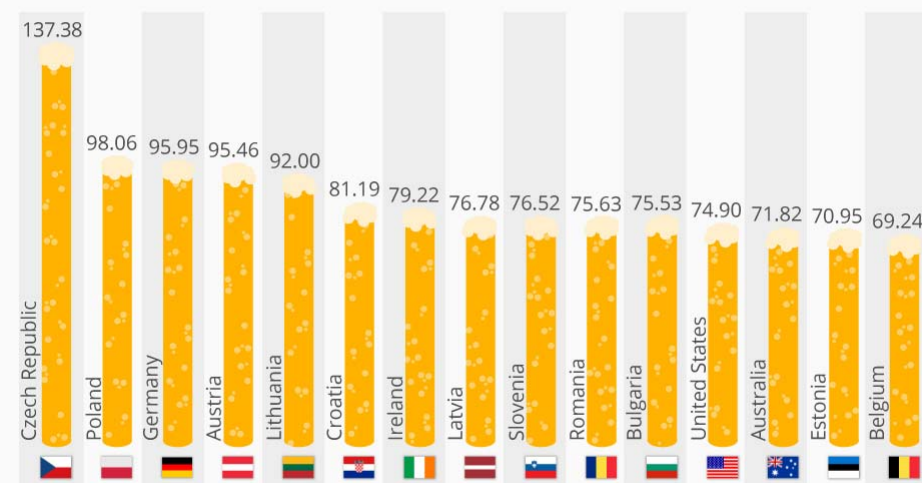
From Data

→ to Indicator



The Countries Drinking the Most Beer

Beer sales per capita in 2017 (in liters)



@StatistaCharts

Source: Statista Alcoholic Drinks Report 2018 - Beer

statista

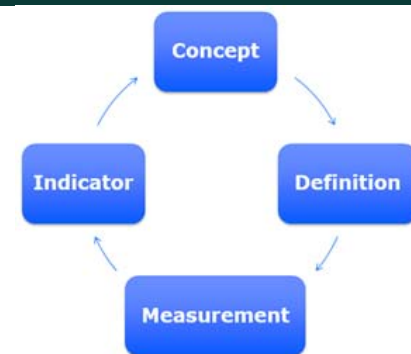
Populating the framework with indicators

An iterative process!

- ⇔ check against Conceptual framework
- ⇔ check against Statistical properties of the indicators

➔ *See sessions on Statistical Coherence*

- does the **correlation structure** reflect the conceptual framework?
 - If not → change indicator specification (denominator)
 - another proxy with better coverage?
 - latent dimension(s) confirming the conceptual structure?



| | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | |
|----|------|------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 1 | 1.00 | 0.46 | -0.58 | -0.06 | 0.23 | -0.39 | 0.34 | 0.49 | 0.15 | 0.02 | 0.17 | -0.07 | 0.37 | 0.44 | 0.45 | 0.46 | 0.43 | 0.45 | 0.46 |
| 2 | | 1.00 | -0.22 | 0.00 | -0.09 | -0.38 | 0.41 | 0.48 | 0.09 | 0.12 | 0.29 | 0.16 | 0.64 | 0.96 | 0.95 | 0.95 | 0.96 | 0.97 | 0.95 |
| 3 | | | 1.00 | -0.11 | -0.24 | 0.52 | -0.62 | -0.61 | -0.05 | -0.14 | -0.09 | 0.00 | -0.22 | -0.21 | -0.26 | -0.24 | -0.23 | -0.24 | -0.24 |
| 4 | | | | 1.00 | 0.37 | -0.32 | -0.27 | 0.36 | 0.16 | 0.06 | 0.04 | 0.18 | 0.02 | 0.05 | 0.03 | 0.00 | -0.01 | 0.02 | 0.01 |
| 5 | | | | | 1.00 | -0.19 | 0.77 | 0.57 | 0.08 | -0.01 | -0.02 | 0.30 | 0.14 | 0.14 | 0.12 | 0.11 | 0.11 | 0.10 | 0.15 |
| 6 | | | | | | 1.00 | -0.60 | -0.81 | 0.68 | -0.38 | -0.54 | 0.57 | -0.09 | -0.32 | -0.31 | -0.19 | -0.31 | -0.39 | -0.26 |
| 7 | | | | | | | 1.00 | 0.95 | -0.37 | 0.50 | 0.15 | -0.06 | 0.40 | 0.41 | 0.48 | 0.38 | 0.40 | 0.38 | 0.45 |
| 8 | | | | | | | | 1.00 | -0.73 | -0.19 | 0.56 | -0.57 | 0.30 | 0.43 | 0.51 | 0.31 | 0.44 | 0.49 | 0.42 |
| 9 | | | | | | | | | 1.00 | 0.11 | -0.02 | 0.31 | 0.14 | 0.09 | 0.10 | 0.15 | 0.11 | 0.09 | 0.12 |
| 10 | | | | | | | | | | 1.00 | 0.12 | 0.00 | 0.01 | 0.12 | 0.09 | 0.10 | 0.09 | 0.11 | 0.10 |
| 11 | | | | | | | | | | | 1.00 | 0.04 | 0.24 | 0.33 | 0.28 | 0.27 | 0.30 | 0.29 | 0.29 |
| 12 | | | | | | | | | | | | 1.00 | 0.12 | 0.15 | 0.12 | 0.14 | 0.13 | 0.11 | 0.14 |
| 13 | | | | | | | | | | | | | 1.00 | 0.89 | 0.97 | 0.92 | 0.92 | 0.88 | 0.97 |
| 14 | | | | | | | | | | | | | | 1.00 | 0.98 | 0.98 | 0.98 | 0.99 | 0.98 |
| 15 | | | | | | | | | | | | | | | 1.00 | 0.98 | 0.98 | 0.98 | 0.98 |
| 16 | | | | | | | | | | | | | | | | 1.00 | 0.98 | 0.98 | 0.98 |
| 17 | | | | | | | | | | | | | | | | | 1.00 | 0.99 | 0.99 |
| 18 | | | | | | | | | | | | | | | | | | 1.00 | 0.98 |
| 19 | | | | | | | | | | | | | | | | | | | 1.00 |

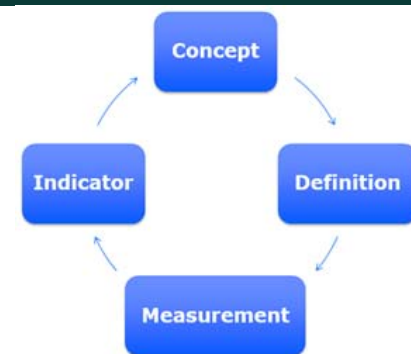
- ⇔ Interpretation of principal component analysis (PCA) outcomes vv. conceptual framework?

Populating the framework with indicators

...an iterative process
& a Learning process!

⇔ Does it meet the expectation of experts, analysts, policy users?

- ➔ Make development participatory!
 - articulate different perspectives
 - Compromises unavoidable

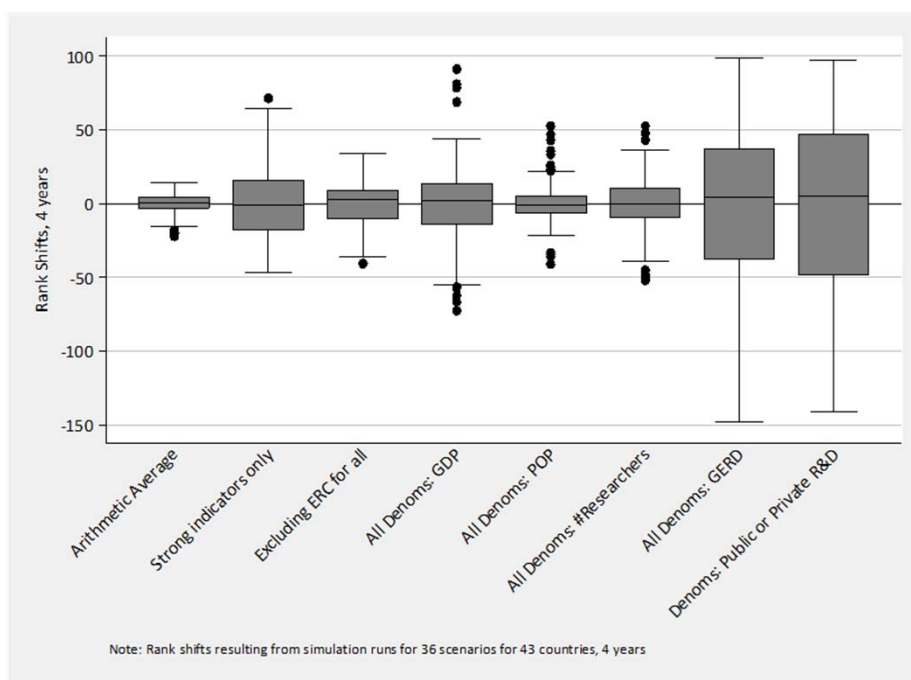


Assumptions influence outcome

The impact of conceptual and methodological choices can be quantified

→ *see sensitivity analyses*

- How important are underlying assumptions?
- Identify key modelling choices
- Test their impact on the final composite scores and rankings



(Source: Hardeman and Vertesy, 2015)

Choosing the IT platform



- Excel: → *see COIN Tool*
 - “WYSIWYG”: offers quick assessment of data quality profile; likely to share results (graphs); available everywhere
 - easy to lose track of manipulations – *make sure to use functions & keep dynamic links to original data sources*



MATLAB



- Statistical software, i.e. STATA, Matlab or R (or Excel VBA)
 - Less intuitive, high initial learning cost –
 - Easy to document choices in script languages (i.e., stata .do files)
 - Excel not ideal for some steps (PCA, simulations, etc...)
- Structuring data: *find layout most suitable for context*
- Downloads can be programmed (see **readSDMX** for R; **getdata** of STATA; etc.); bulk download preferable also in Excel
- *Keep track of numerators, denominators, different versions tested*

A final note

- Make your choices clear for ALL (including yourself!)
 - What, Why, How...?
 - = just as important: what *not*, why *not*, how *not*...?
- Make your indicator time-resistant:
 - socio-political context may change!
- Clear documentation is essential
 - Think of updates –
an edition is not the end of history!





THANK YOU

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

COIN in the EU Science Hub
<https://ec.europa.eu/jrc/en/coin>

COIN tools are available at:
<https://composite-indicators.jrc.ec.europa.eu/>

The European Commission's
Competence Centre on Composite
Indicators and Scoreboards



References / 1

- Barré, R. (2001). The Agora model of innovation systems: S&T indicators for a democratic knowledge society. *Research Evaluation*, 10(1), 13-18.
- Barré, R. (2010). Towards socially robust S&T indicators: Indicators as debatable devices, enabling collective learning. *Research Evaluation*, 19(3), 227-231.
- Davis, K. E., Fisher, A., Kingsbury, B. and Merry, S.E. (eds.) (2012) *Governance by Indicators: Global Power Through Quantification and Rankings*. Oxford: Oxford University Press.
- Ferretti, F., Guimaraes Pereira, A., Hardeman, S. & Vertesy, D. (2017) "Research Excellence Indicators: time to re-imagine the 'making of'?" Paper under revision for Science and Public Policy
- Fukuda-Parr, S., Yamin, A. E., & Greenstein, J. (2014). The power of numbers: a critical review of millennium development goal targets for human development and human rights. *Journal of Human Development and Capabilities*, 15(2-3), 105-117.
- Gisselquist, R. (2014) Developing and evaluating governance indexes: 10 questions. *Policy Studies* 35(5), 513-531
- Hardeman S., Van Roy V., Vértésy D. and Saisana M. (2013) *An analysis of national research systems (I) A Composite Indicator for Scientific and Technological Research Excellence* EUR 26093. Luxembourg: Publications Office of the EU; doi: 10.2788/95887
- Hardeman S. & Vértésy D. (2015) *An update of the research excellence indicator*. Luxembourg: Publications Office of the EU; doi: 10.2760/770103

References / 2

- Hardeman, S., & Vertesy, D. (2016). Evidence-based policy learning: the case of the research excellence indicator. Paper presented at the 2016 STI Conference in Valencia, Spain
- JRC-OECD (2008). *Handbook on constructing composite indicators: methodology and user guide*. Paris: OECD.
- Paruolo, P., Saisana M. and Saltelli A. (2013) "Ratings and rankings: voodoo or science?" *Journal of the Royal Statistical Society: Series A*. 176(3), 609–634
- Power, M. (1994). *The audit explosion* (No. 7). Demos Paper No. 7.
- Rottenburg, R., Merry, S.E., Park, S-J. and Mugler, J. (eds.) (2015) *The World of Indicators: The Making of Governmental Knowledge through Quantification*. Cambridge: Cambridge University Press
- Saltelli, A. (2007). Composite indicators between analysis and advocacy. *Social Indicators Research*, 81(1), 65-77.
- Saisana, M., 2007. Composite learning index: Robustness issues and critical assessment. European Commission JRC
- Saisana, M. 2010. ELLI-Index: a sound measure for lifelong learning in the EU. JRC. Luxembourg: Publications Office of the EC
- Saisana M., D'Hombres B., Saltelli A. (2011) "Rickety Numbers: Volatility of university rankings and policy implications." *Research Policy* 40, 165–177
- Saisana M., Saltelli A., Tarantola S. (2005) "Uncertainty and sensitivity analysis techniques as tools for the analysis and validation of composite indicators", *Journal of the Royal Statistical Society A* 168(2), 307-323.
- Tijssen, R. J. (2003). Scoreboards of research excellence. *Research Evaluation*, 12(2), 91-103
- United Nations (2012) *National Quality Assurance Frameworks*
[URL: <http://unstats.un.org/unsd/dnss/qualityNQAF/nqaf.aspx>]