UNESCO Institute for Statistics

Issues in Measuring Learning

October 17th 2016

Abbie Raikes

UNESCO
United Nations
Educational, Scientific and
Cultural Organization

UNESCO
INSTITUTE
for
STATISTICS

UNESCO Institute for Statistics

Themes from Panel

- Measuring learning to lead to improvement
- Alignment with national priorities, existing goals, standards and culture
- Equity as organizing principle
- Ensure data are valuable and recognize country capacity

UNESCO Institute for Statistics

What's Ideal?

- Achieve expectations at global level for SDG monitoring by producing universal learning indicators for each target
  - Agreement on what counts as reaching minimum proficiency in global context
  - Define technical pathways to produce the indicators
  - Promote reliable national, regional and global measurement to produce necessary data
  - Equity in measurement

UNESCO Institute for Statistics

Big Issues to Resolve …

- While targets are focused on different ages and types of learning, there are common elements in each
- These include defining …
  - Comparability: What's globally comparable? How can multiple forms of data help define this?
  - Thresholds: What's "good enough" learning and development – minimum levels of proficiency, basic learning, developmentally on track
  - Periodicity: How frequently to measure – should be based on policy impact, sensitivity to interventions

Globally Comparable

- Two elements to start with …
  - Conceptual agreement on what should be relevant across contexts
  - Empirical support to demonstrate that measurement and more specifically, items function in similar-enough ways across contexts
- Possible and plausible to conclude some constructs are not globally comparable
  - Tricky balance in light of greater emphasis, political prioritization of areas that can be compared

How can comparability be addressed?

- Define common content through examination of existing measures, curricula framework
- Use psychometric models to order items in terms of difficulty
- Examine degree of similarity in how items function across countries
- Determine appropriate level at which to “compare”
- Other methods?

Options for Making Measurement Easier

1. One measure used everywhere: Same items, same administration everywhere, with some degree of adaptation
2. Common core of items: One set of items, part of larger and more culturally-adapted set
3. Common constructs, with items that may vary: May be able to "match" at level of construct, but with different items
4. Whatever country feels is appropriate (thresholds set by country measures that meet global standards)

UNESCO Institute for Statistics

Defining Thresholds

- What counts as
  - "minimum proficiency"
  - "fixed level of proficiency"
  - "developmentally on track"
- Absolute (*defined by reaching a set standard - criterion*) or relative (*defined in relation to rest of population – norm-referenced*)
- Is there a common level of functioning that is applicable to all people?
  - Should be considered within framework of the goal – focused on equity

UNESCO Institute for Statistics

Periodicity

- Frequency of measurement:
  - What's feasible? How frequently can data reasonably be produced?
  - How sensitive are data to changes in policies, investments, practice?
- Balance between feasibility and sensitivity
  - More feasible can also be less sensitive to change
- Country capacity for high-quality measurement, balance between measurement complexity vs attainable capacity building
- Showing change over time, as well as absolute scores

UNESCO Institute for Statistics

Comparability

- Pathways to defining “globally comparable” should be outlined for each target
  - Both conceptual knowledge and existing data should be used to generate recommendations for globally comparable
  - Multiple methodologies may be applicable and necessary
- “Ideal”: Some degree of comparability, respecting cultural differences and nature of learning in various constructs

UNESCO Institute for Statistics

Thresholds

- Articulating a threshold: core question of absolute vs. relative
  - How much do we know about the science of learning to define this?
  - Define what data and methodologies are required to produce absolute vs. relative data

- “Ideal”: Method most likely to reveal equity between and within countries … absolute or relative may vary by target?

Periodicity

- Consideration of
  - Likely sensitivity of the data to changes in policy and investment
  - Feasibility of large-scale data collection
  - Ability of data to show changes over time, to show progress

- "Ideal": Data collected on regular basis, sensitive to large shifts and able to detect progress

UNESCO Institute for Statistics

Main Points for Discussion

- How best to balance the technical challenges with political demands for data on learning
- What are the pros and cons of reaching for higher degrees of comparability across constructs and targets?
- How can the agenda on equity be expressed through measurement?