

“READINESS TO LEARN” ASSESSMENT PROJECT

Abbie Raikes, UNESCO
a.raikes@unesco.org

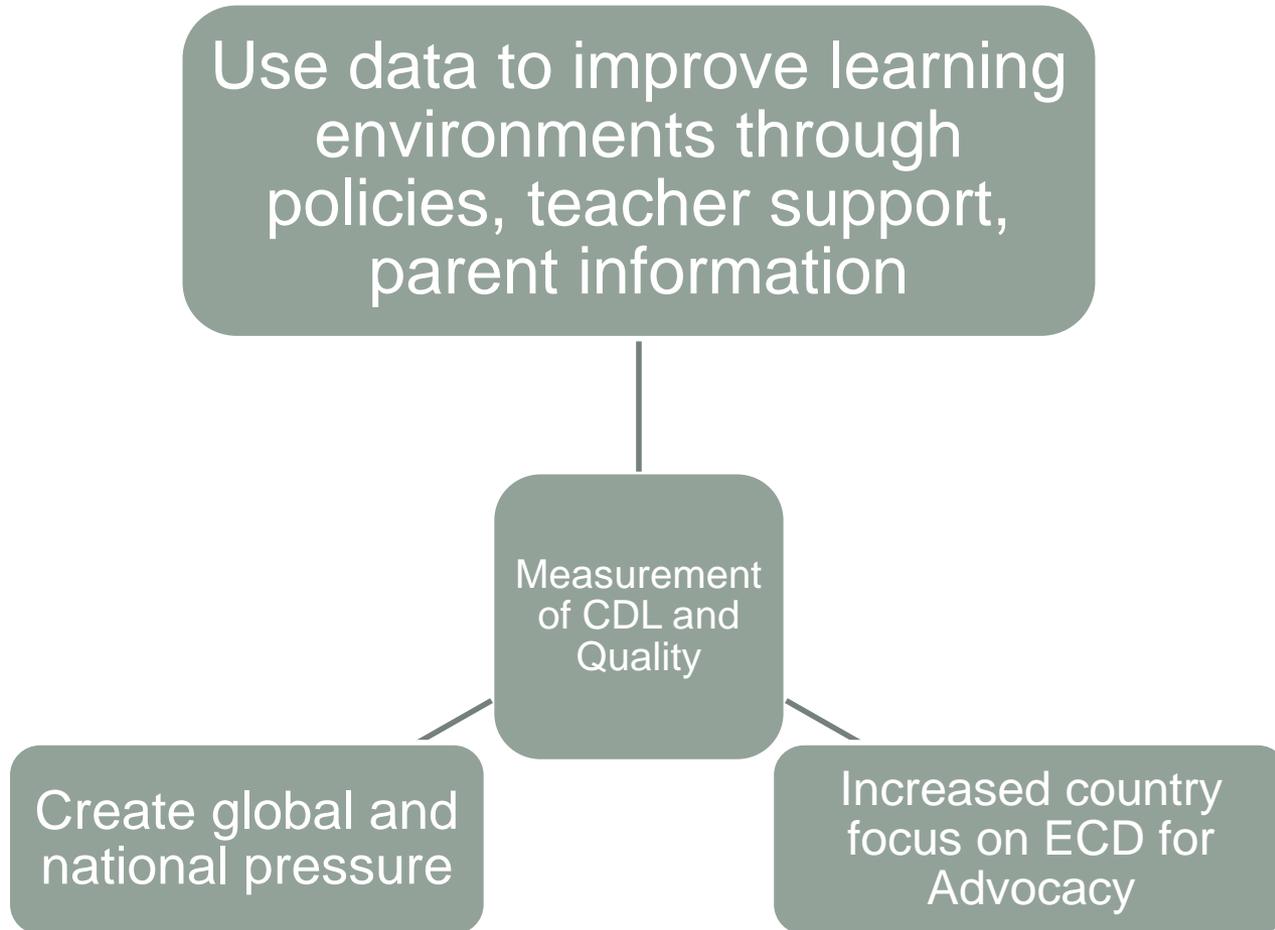
“Readiness to Learn” Assessment Project

- Develop prototypes for measurement of quality and child development/learning between ages 4 and 6, with emphasis on low-income countries
- Mandate from Learning Metrics Task Force
- Partnership with World Bank, Brookings Institution, UNICEF
- Work beginning now, field testing in early 2015 and completion by early 2016
 - Four field-testing countries: Africa and LAC

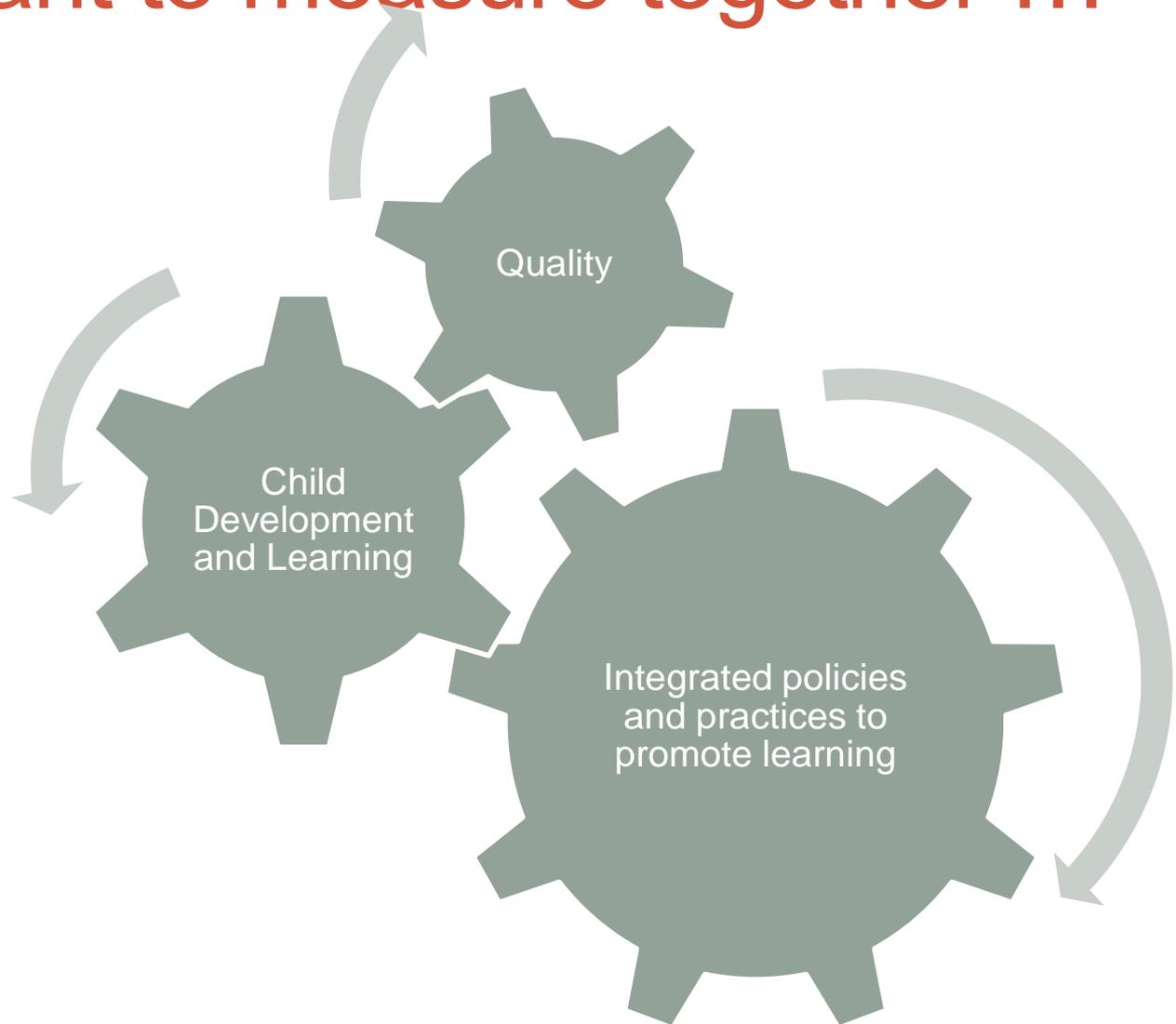
Why national measurement?

- National assessments can be longer, more frequent and adapted to local settings – *but don't contain the same set of items, so can't be considered globally comparable*
- Local data are important for policy and programmatic development
 - Design of curricula; connections with teacher training
 - Inform policies and funding decisions
 - More detailed information on inequities within countries
- May be possible to inform global tracking over time

Core Assumption: Measuring Contributes to Improvement



Important to measure together ...



Vision of what we will achieve ...

- Goal of free, open-source items to allow easy use and adaptation across countries
- Designed with users in mind from the start
- Intended for use in improving learning

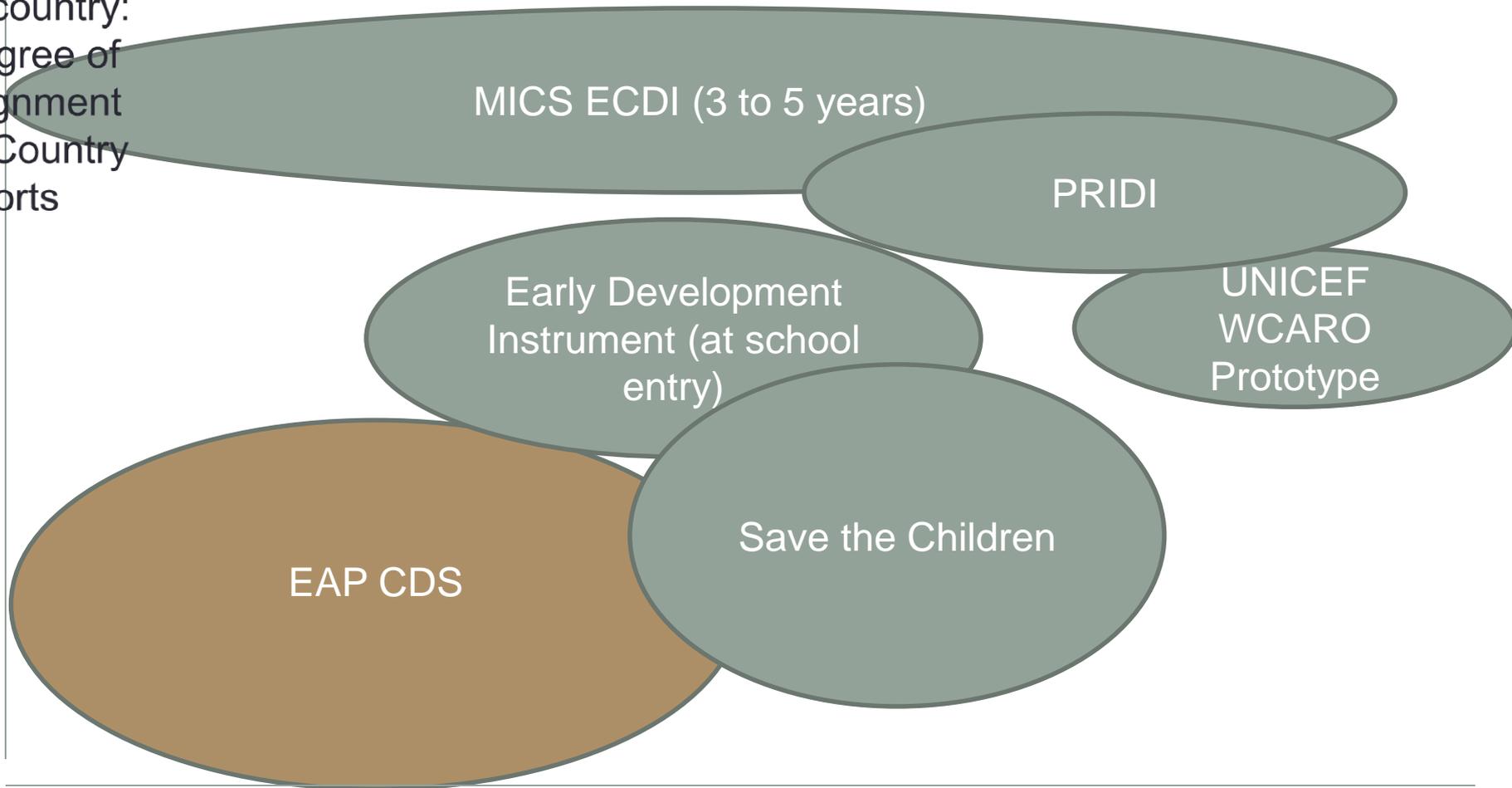
Flexible, feasible measures to provide on-going information about children's learning and development and quality of learning environments

What will it take to build: Two themes

- Content: What's in it?
 - Constructs that are scientifically valid and relevant
 - Item reliable and accurate; technically sound
 - Emphasis on “teachable” and “actionable” items
- Usability: How will it be designed?
 - How the data will be collected
 - How the results will be shared and used to improve practice – risks like high-stakes!

Existing Measures of Child Development and Learning

Population
to country:
Degree of
Alignment
to Country
Efforts



Number and range of countries included to date

Are we clear on what the construct is and why it's important?

Identify constructs to be measured and purpose of data

- ECD: Developmental domains established
- Quality: Not done

Has construct been defined and tested in a range of cultures? Are tests valid and reliable?

Review research literature; identify measurement options

- ECD: 70% done
- Quality: 20% done

Country feedback on goals and purpose of measurement

Is it possible to identify specific items to measure the construct that are feasible, reliable, relevant and useful?

How useful would the data resulting from these tests be for decision-making? How would the data be used and by whom?

Identify large set of items that could be especially useful

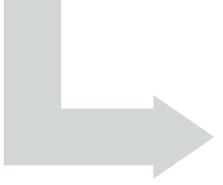
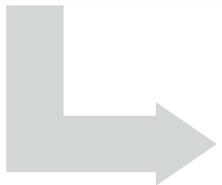
- ECD: Options already developed; need revision
- Quality: Not yet started

Field test to shorten list of items

Can items be grouped together to create a set of core and optional items, or does it seem that the tests are only valid when using many items?

How well do the responses from the smaller item sets predict outcomes of interest?

Validate



Field Testing and Validation Are Critical

Content and use both to be addressed in field testing:

- Technically strong data that are not useable won't lead to change
- Rough, inaccurate measures can't be basis for effective use
- Measures that align with country priorities will motivate users

Some of the anticipated issues

- Creating common core of items, while adjusting to local context
- Validating adequately – knowing that predictive validity, which is hardest to obtain, will take the longest
- Drawing on tremendous expertise already developed

In sum ...

- National efforts can complement work at global level
- Some indicators may be easier to develop first at the national level, which then can help inform global efforts
- Challenges around item comparability and ensuring proper use are important to recognize