



*Improving Learning*

Australian Council *for* Educational Research

# Key Quality Concepts of Learning Assessment


Regional TALENT Workshop

Ursula Schwantner & Andriy Dubovyk,  
Australian Council for Educational Research

4 December 2019

# Overview



- Introduction to ACER and Centre for Global Education Monitoring (ACER-GEM)
  - Key quality concepts for learning assessments
  - 14 Key areas of a robust assessment program
  - Applying the key quality concepts: The MTEG Model in Afghanistan
  - Group work on the 14 key areas
- 

# Australian Council for Educational Research

- Australian Council *for* Educational Research
- Established in 1930 as a not-for-profit international research organisation
- *ACER's mission:* Creating and promoting research-based knowledge, products and services to improve learning

# ACER research



Early childhood education



School education



Higher education



Vocational, adult and workplace education



Education and development



Indigenous education

<https://www.acer.edu.au/research>



# Centre for Global Education Monitoring (ACER-GEM)

- Collaboration between ACER and DFAT
  - to support the monitoring of educational outcomes worldwide to inform evidence-based education policy aimed at improving educational progress for all learners
- 

# ACER-GEM work program



System strengthening



Education 2030



Reviews and analysis



# Large-scale Learning Assessments







# Large-scale learning assessments

- focus on defined learning domains, based on an assessment framework, may be referenced to a national curriculum;
- can be international, regional or national in scope;
- focus on a particular population (sample-based or census).

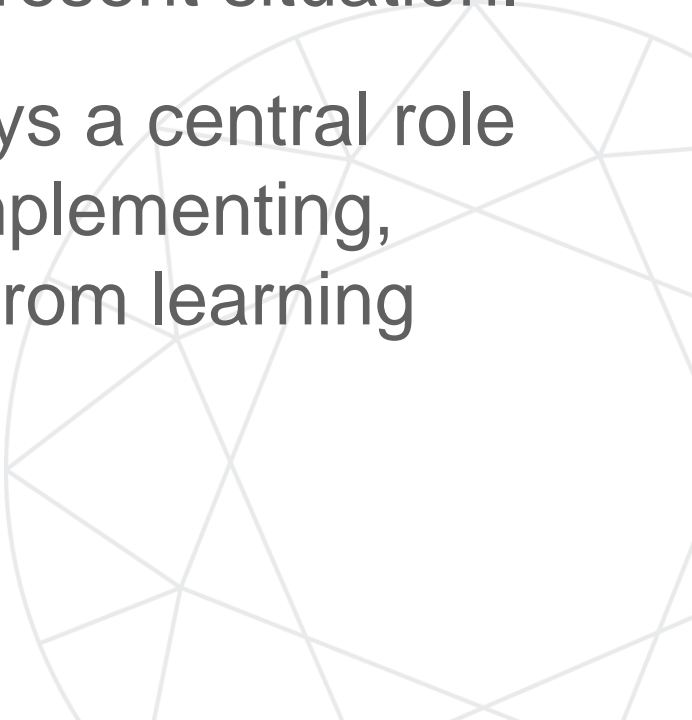


# Purposes of large-scale assessments

- to establish achievement levels of a particular population in a learning domain
- to monitor progress in learning outcomes over time, between grades
- to investigate associations between achievement and contexts in which learning takes place, and to quantify differences between sub-populations

# Data quality and accuracy



- To be effective, large-scale learning assessments need to gather data that provide an accurate reflection of the present situation.
  - Data quality and accuracy plays a central role in all phases of developing, implementing, analyzing and using the data from learning assessments.
- 



# Key Quality Concepts for Learning Assessments





# Key quality concepts

1. Clarity and consistency of purpose
2. Fitness for purpose
3. Objectivity and independence
4. Transparency and accountability
5. Technical rigour
6. Ethicality and fairness



# Key Areas of a Robust Assessment Program

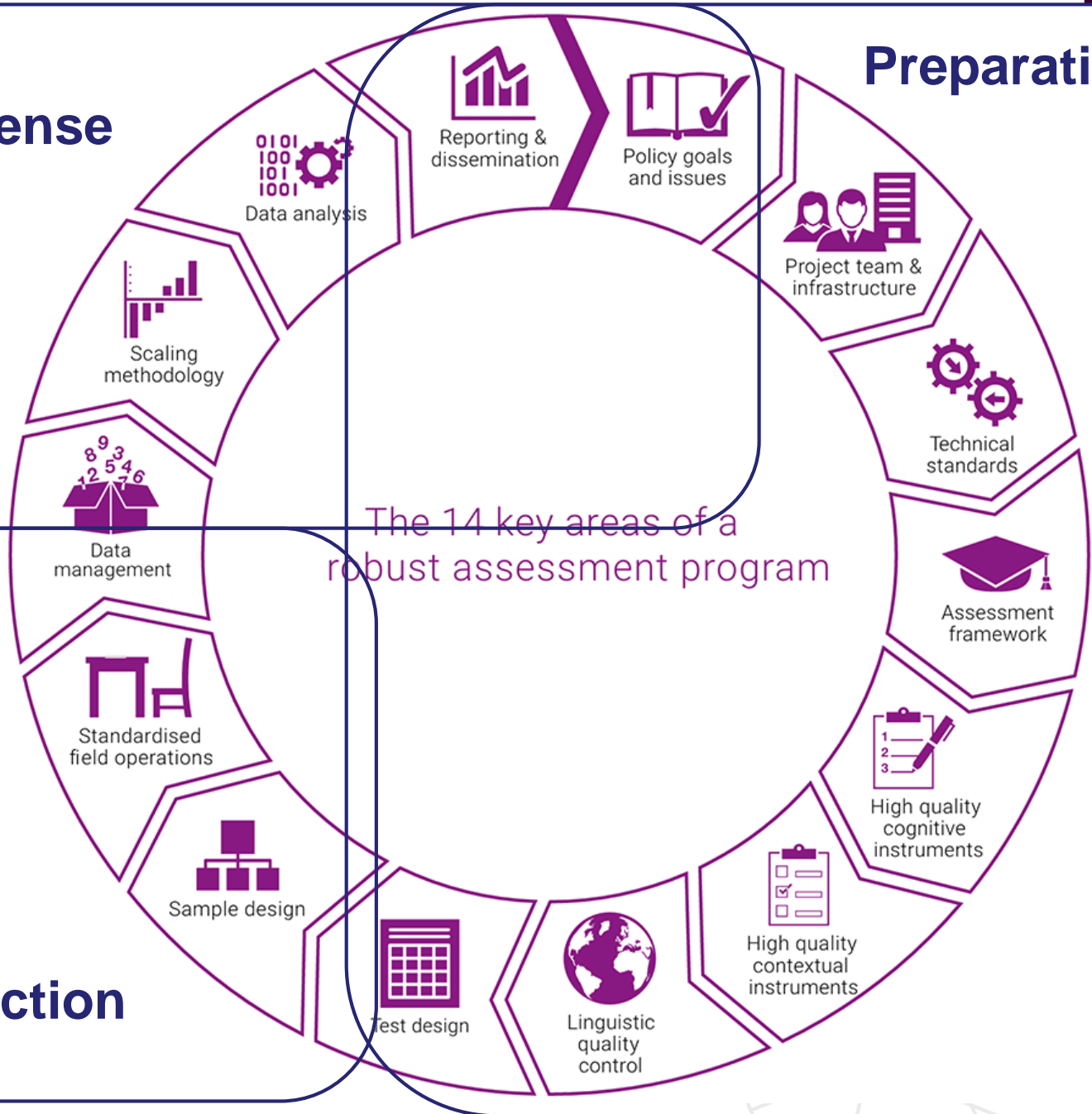


# Making sense of data

# Preparation

The 14 key areas of a robust assessment program

# Data collection



# **Key Area I: Formulating policy goals and priorities to be addressed with the learning assessment**

**Clearly articulated policy goals and measurement priorities that are relevant to key stakeholders and inform the content, design and scope of the assessment program.**

- **Identify education priorities.**
- **Engage stakeholders.**
- **Identify how results will be used.**
- **Evaluate the feasibility of implementing an assessment based on the policy goals and measurement priorities.**

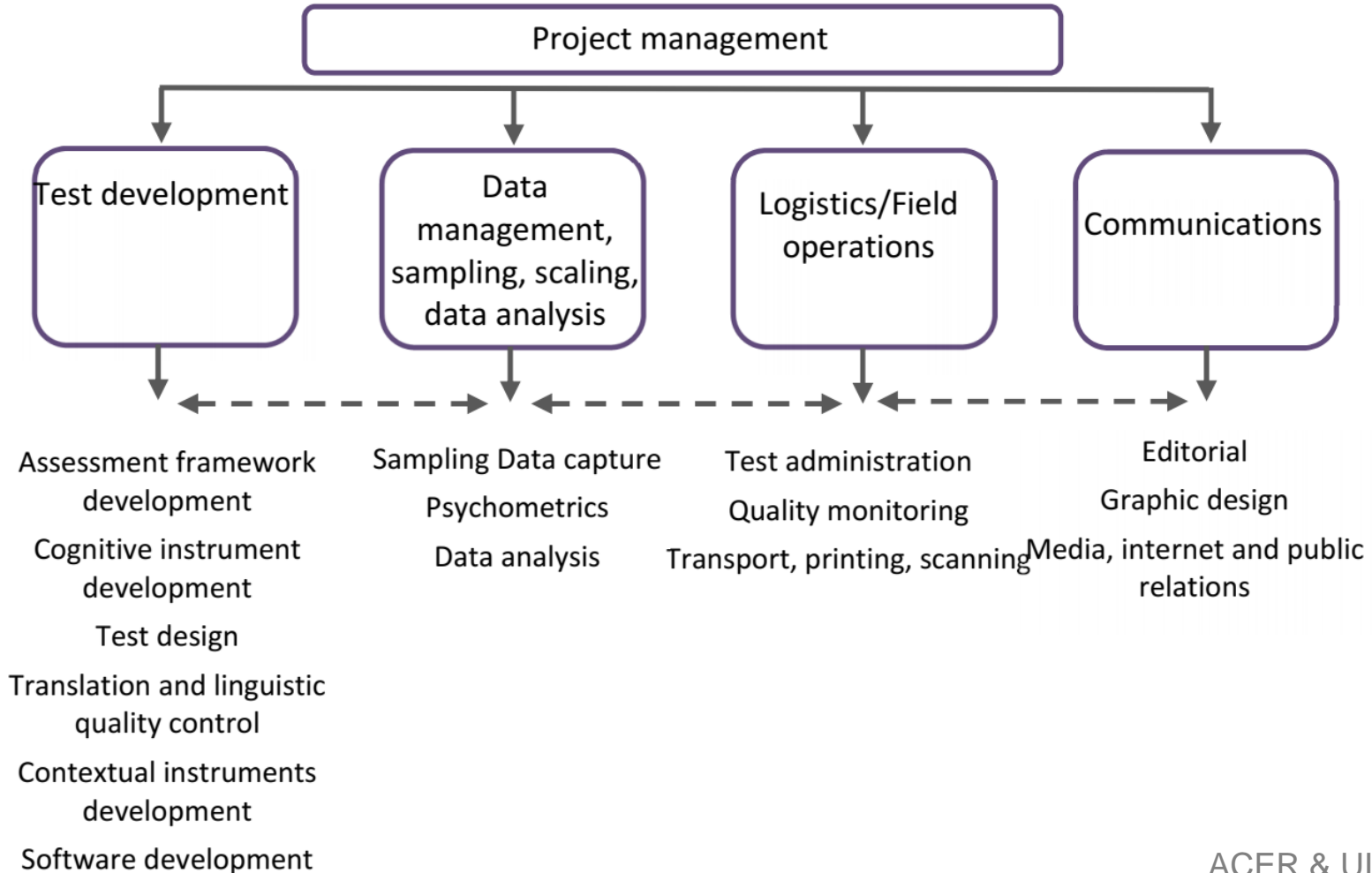


## **Key Area 2: Establishing and managing an assessment team responsible for designing and implementing the learning assessment**

**An assessment team with dedicated staff that is appropriately skilled and adequately resourced to respond to the diverse demands of designing, implementing, analysing, and disseminating the outcomes of the learning assessment.**

- **Establish an assessment team.**
- **Develop capacity where needed.**
- **Outsource when needed.**
- **Secure physical infrastructure.**

# Key Area 2: Establishing and managing an assessment team responsible for designing and implementing the learning assessment



## **Key Area 3: Formulating and articulating technical standards to guide assessment activities**

**A key document that clearly describes standards of technical quality for all aspects of the learning assessment, and indicates how standards can be used as part of quality monitoring and reporting.**

- **Establish a technical advisory committee.**
- **Determine the details of the technical standards.**
  - **Sampling.**
  - **Data.**
  - **Psychometric.**

## **Key Area 4: Developing an assessment framework**

**A document that uses a consistent terminology to communicate the purpose and characteristics of the learning assessment to individuals/groups who are working on it and to a broader audience.**

- **Establish expert committees.**
- **Develop the framework.**
- **Consult stakeholders.**
- **Review the framework.**

## **Key Area 5: Developing high quality cognitive instruments**

**Cognitive instruments containing items with proven reliability, validity and fairness with regard to the population(s) of interest.**

- **Establish a team of test developers.**
- **Establish mechanisms for obtaining input from outside the test development team.**
- **Develop scoring guides.**
- **Develop instructions.**
- **Review, pilot and field trial test items.**

## **Key Area 6: Developing high quality contextual instruments**

**Contextual instruments containing items with proven reliability, validity and fairness with regard to the population(s) of interest.**

- **Establish a team of contextual instrument developers.**
- **Establish mechanisms for obtaining external input and feedback.**
- **Write contextual items.**
- **Review, pilot and field trial contextual items.**

## **Key Area 7: Linguistic quality control for translation of cognitive and contextual instruments**

**Cognitive and contextual instruments that are appropriate, linguistically equivalent, and psychometrically equivalent across multiple languages.**

- **Develop a translatable source version.**
- **Provide supporting documentation.**
- **Recruit and train translators.**
- **Implement a multi-step process.**
- **Manage the workflow.**
- **Analyse field trial statistics.**

## **Key Area 8: Designing the cognitive and contextual instruments**

**A design that ensures efficiency in sample sizes, balanced assessment content, appropriate assessment length, and stable measures over time.**

- **Consult psychometricians and item developers.**
- **Allocate items within and across test forms.**
- **Allocate items within and across questionnaire/interview forms.**
- **Allocate forms to participants.**
- **Layout and proofread the forms.**



## Key Area 9: Sampling

**A sample that, through the use of scientific sampling methods, helps to guarantee appropriate levels of statistical precision and validity in the interpretation of assessment results.**

- **Develop a sampling plan and select a sample methodology.**
- **Construct or obtain a comprehensive sample frame of the target population.**
- **Conduct the sampling.**
- **Maintain optimal response rates.**
- **Apply the proper weighting methodology to improve the accuracy of estimates and to obtain correct standard errors.**

## **Key Area 10: Standardised field operations**

**Field operations that are standardised, documented and monitored to ensure that the data are collected under the same conditions, independent from the administration context, in an efficient and secure manner.**

- **Develop an operation plan.**
- **Prepare field operations manuals and documentation.**
- **Contact sampled institutions and individuals.**
- **Recruit and train key personnel.**
- **Check assessment materials.**
- **Administer the assessment according to standardised procedures.**
- **Verify and register the return of assessment materials.**

## **Key Area II: Managing data**

**A final database that respects respondent anonymity, is free from discrepancies and errors, and is appropriately structured and documented for analysis and dissemination.**

- **Develop a data management plan.**
- **Design the codebook.**
- **Prepare data management and data capture software and hardware.**
- **Establish and train a data manager and data management team.**
- **Prepare protocols for data capture, data cleaning, verification and validation.**
- **Prepare data backup protocols.**
- **Prepare data documentation and transfer protocols.**

## **Key Area 12: Scaling cognitive and contextual data**

**Cognitive and contextual data that is scaled using well-developed analytical tools in order to support a range of useful comparisons and to communicate information that is meaningful to a range of users.**

- **Develop a data analysis plan.**
- **Choose analytical model.**
- **Identify scales and possible sub-scales.**
- **Analyse cognitive data.**
- **Describe the cognitive scales.**
- **Analyse contextual data.**

## Key Area 13: Analysing data

**Analytical results that are fully documented and reproducible, and that permit valid and useful inferences about the population(s) of interest**

- **Develop a data analysis plan.**
- **Assign sample weights.**
- **Calculate the standard error.**
- **Analyse data.**
- **Analyse trends.**

## **Key Area 14: Reporting and dissemination**

**Appropriate products and approaches to reporting and dissemination that are tailored to the different stakeholder groups and promote appropriate and effective use of the assessment results by those groups.**

- **Identify different information needs of stakeholders.**
- **Develop a dissemination strategy.**
- **Develop dissemination products.**
- **Monitor how assessment data are used over time.**



**ACER**  
**Monitoring Trends in Educational Growth**  
MTEG in Afghanistan

# ACER Monitoring Trends in Educational Growth (MTEG)

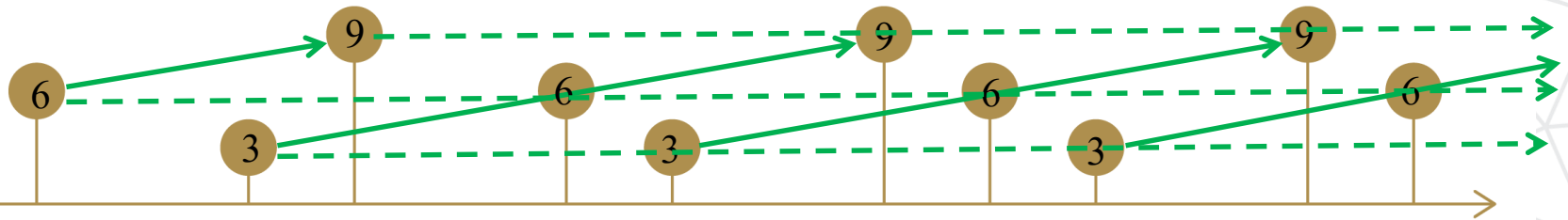
MTEG is a partnership model incorporating the key quality concepts and 14 key areas:

- Capacity building in learning assessment
- Measuring learning progression (grades 3–6–9)
- Reporting against SDG 4.1 targets and minimum proficiency levels



# MTEG Design

- Growth in cohort Grades 3 – 6 – 9
- Change in Grade performance over time



- Domains: reading, writing, mathematics
- Policy priorities: Learning levels and sub-groups of interest, e.g. gender, location, language of instruction

# MTEG in Afghanistan

## Grade 6

- Implemented in 2013
- Domains: reading, writing, mathematics
- Schools and student questionnaires
- Two languages: Dari and Pashto
- Sample size: 5979 students in 110 schools across 13 provinces
- Paper and pencil

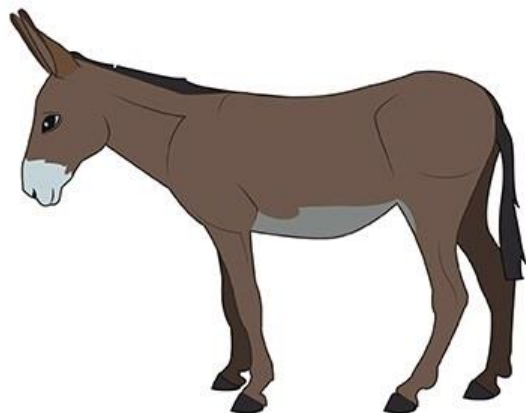
## Grade 3

- Implemented in 2015/16
- Domains: reading, mathematics
- Schools and student questionnaires
- Two languages: Dari and Pashto
- Sample size: 4936 students from 179 schools across 15 provinces
- Tablet-based assessment

# MTEG tablet based assessment – example item



Listen to each word.  
Select the correct word for the picture.



house

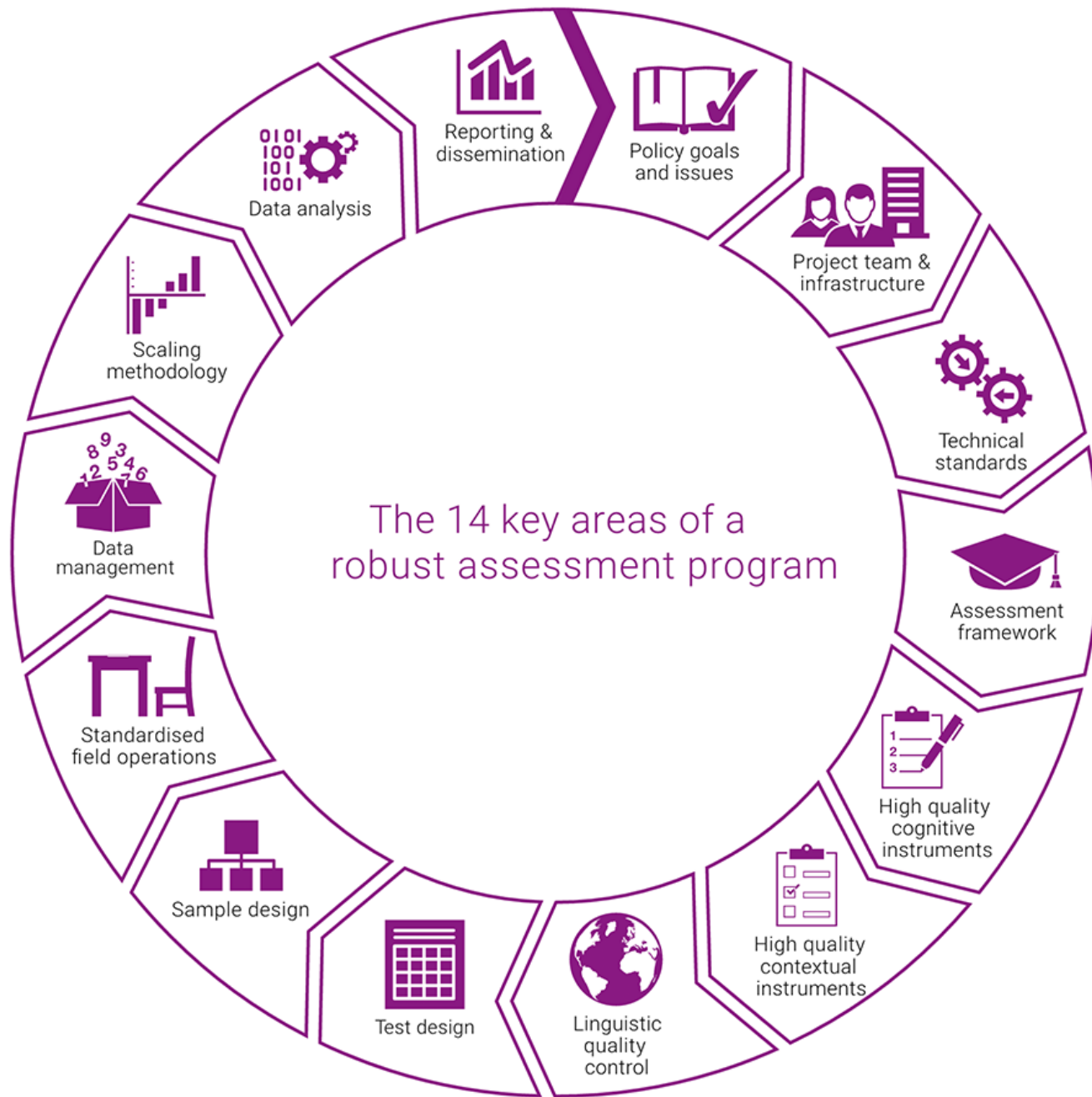


donkey



mountain

Next



# MTEG Partnership Model

Key area	ACER	Min Edu Afghanistan
<b>1. Policy goals and issues</b>	Advice	Stakeholder consultations
<b>2. Project team and infrastructure</b>	Advice	Resourcing and coordination
<b>3. Technical standards</b>	MTEG model	Review and adoption
<b>4. Assessment framework</b>	MTEG model	Review, contextualisation and adoption
<b>5. Cognitive instruments</b>	Item development	Review, contextualisation and adoption

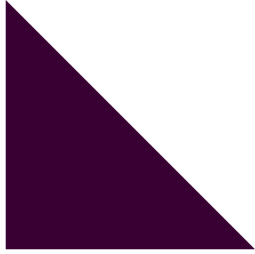
# MTEG Partnership Model

Key area	ACER	Min Edu Afghanistan
<b>6. Contextual instruments</b>	Questionnaire development	Advice on factors of policy interest; review
<b>7. Linguistic quality control</b>	Sourced linguistic verification	Translation
<b>8. Test design</b>	MTEG model for TD	
<b>9. Sample design</b>	Scientific SD, school sampling	Advice, sample frame
<b>10. Standardised Field Operations</b>	Advice, provision of manuals	Adapted manuals, organised and implemented test administration

# MTEG Partnership Model

Key area	ACER	Min Edu Afghanistan
<b>11. Data management</b>	Data processing and management	Marking of open-ended responses
<b>12. Scaling methodology</b>	Scaling using IRT Described proficiency scales	Participation in capacity-building workshops
<b>13. Data analysis</b>	Conducted data analysis	Participation in capacity-building workshops
<b>14. Reporting and dissemination</b>	Drafting reports; policy seminars	Review of reports; participation in policy seminars

# MTEG Afghanistan capacity building



- Workshops in areas such as:
  - modern assessment theory and practice,
  - assessment frameworks,
  - item review,
  - linguistic quality control
  - sample design,
  - introduction to IRT,
  - open-response coding,
  - test administration using tablets,
- Consultations and hands-on support







# MTEG Afghanistan reports

- Class 6 proficiency
- Class 6 girls and boys
- Class 6 school factors
- Class 3 proficiency
- Monitoring Trends in Educational Growth Assessment Framework for Afghanistan

<https://www.acer.org/au/gem/key-areas/system-strengthening/mteg>



# Key Quality Concepts of Learning Assessment

Contact:

Dr Ursula Schwantner, Senior Research Fellow,  
ACER-GEM: [Ursula.Schwantner@acer.org](mailto:Ursula.Schwantner@acer.org)

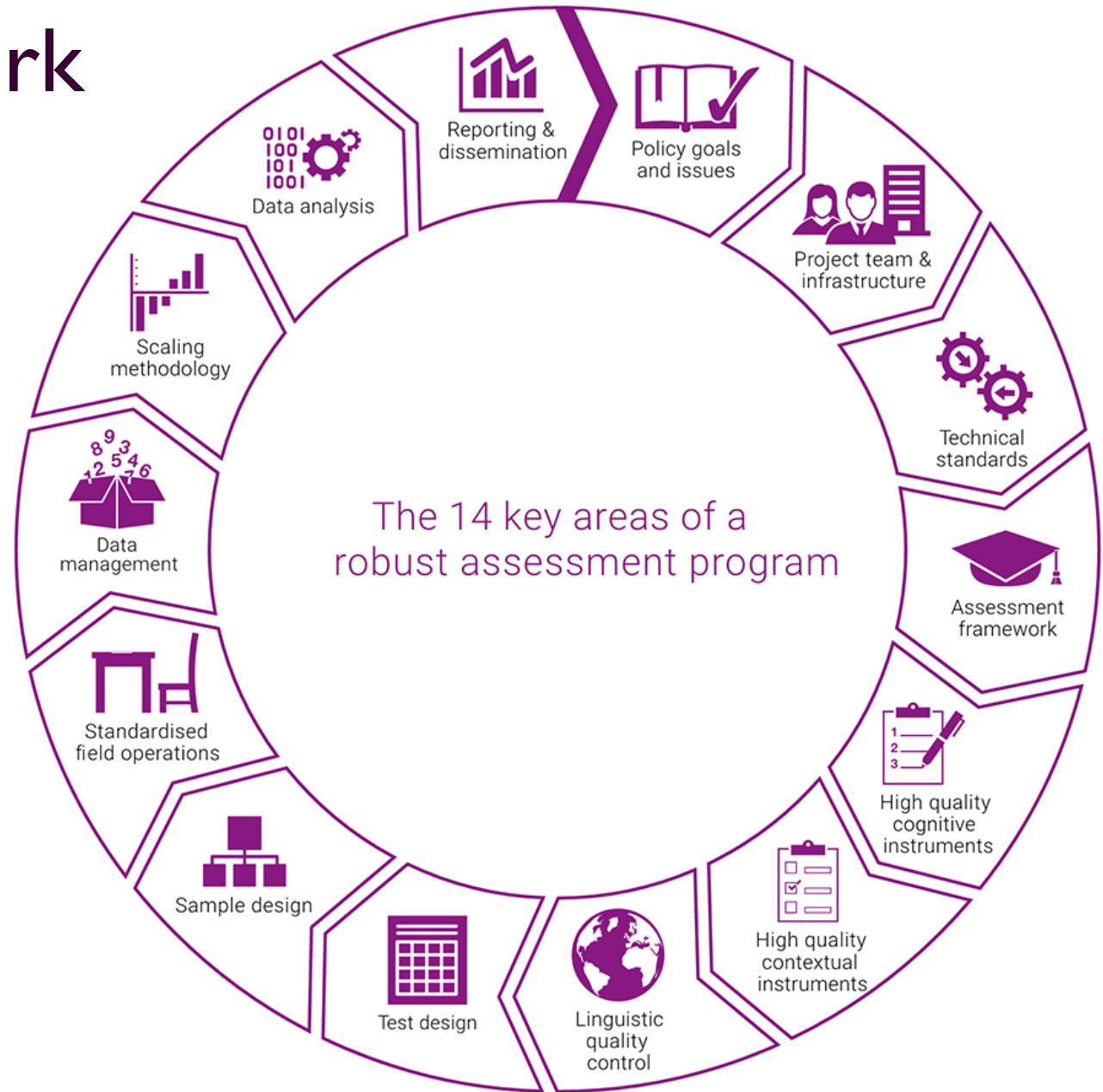
Dr Andriy Dubovyk, Manager School  
Assessment Services and International  
Development, ACER UK:  
[Andriy.Dubovyk@acer.org](mailto:Andriy.Dubovyk@acer.org)



# Group Work

Discuss:

1. What are the three key areas your country has strong capacity in?
2. What are the three key areas your country would require additional knowledge and support?
3. Report back



# References

- ACER Centre for Global Education Monitoring (ACER-GEM): <https://www.acer.org/gem>
- ACER centre for Global education Monitoring (ACER-GEM). (2016). Monitoring Trends in educational Growth. Assessment Framework for Afghanistan. Australian Council for Educational Research Ltd. Melbourne
- ACER & UIS (2017). Principles of Good Practice in Learning Assessment (GP-LA). Draft available under: <http://uis.unesco.org/sites/default/files/documents/principles-good-practice-learning-assessments-2017-en.pdf>
- American Educational Research Association. (2011). *Code of Ethics*.
- American Educational Research Association, American Psychological Association, National Council on Measurement in Education, & Joint Committee on Standards in Educational and Psychological Testing. (2014). *Standards for Educational and Psychological Testing*. Washington DC: AERA.
- Anzai D., Munro-Smith P., Robertson S, Walker, M. & Daraganov A. (2018). Monitoring Trends in educational Growth. Class 3 Proficiency in Afghanistan 2015-16. Australian Council for Educational Research Ltd. Melbourne
- Bordens, S. K., & Abbott, B. B. (1996). *Research design and methods: A process approach (3rd ed.)*. California: Mayfield Publishing Company.
- Floyd, J., & Fowler, J. (2009). *Survey research methods (4th ed.)*. Thousand Oaks, CA: Sage.
- Lumley, T., Mendelovits, J., Stanyon, R., Turner, R. & Walker, M. (). Monitoring Trends in educational Growth. Class 6 Proficiency in Afghanistan 2013. Australian Council for Educational Research Ltd. Melbourne
- Masters, G. (2014). *Assessment: Getting to the essence*. Melbourne: Australian Council for Educational Research.
- Masters, G. (2018). *A Commitment to Growth: Essays on Education*. ACER Press
- Murphy, M., & Schulz, W. (2006). *Sampling for national surveys in education*. Melbourne: ACER.
- Ross, K. N. (2005). Module 3: Sample design for educational survey research. In K. N. Ross (Ed.), *Quantitative Research Methods in Educational Planning*. Paris, France: UNESCO International Institute of Educational Planning. Retrieved from <http://www.iiep.unesco.org/en/library-resources/briefs-papers-tools>.
- United Nations Statistics Division. (2015). *United Nations Fundamental Principles of Official Statistics: Implementation Guidelines*. Retrieved from <http://unstats.un.org/unsd/dnss/gp/impguide.aspx>