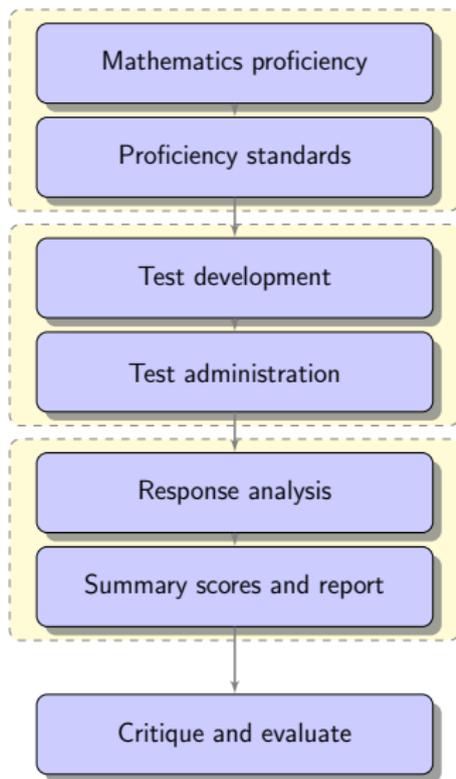# Assessment Process

## Dr. Kelvin D. Gregory

Education Consultant
Adjunct at Flinders University
kelvin.gregory@flinders.edu.au

4-6 NOVEMBER, 2015 / KATHMANDU, NEPAL

# A general process applied

# 1. Define and describe

An indicator needs to be clearly defined and important features identified and
described.
This is done by:

- Working with experts, using resources.

- Writing a succint definition of the indicator. No more than one paragraph.

- Then, in more detail, describing the features, traits or characteristics of that indicator. This is the hard part and will be done over time, with reflection and insight.

    ▶ What is meant by the "mathematics proficiency"?
    ▶ What is included in this continuum? And, looking at the other continua, what is not?
    ▶ What are the features of a school/classs/student with low or high mathematics proficiency?

# 1. Define and describe

This paragraph has:

- subject (content),
- **performance (cognition, attitude, or action)**
- sometimes conditions.

*"For the purposes of PISA 2015, mathematical literacy is defined as follows:*

*Mathematical literacy is an individual's capacity to* **formulate, employ, and interpret** *mathematics in a variety of contexts. It includes* **reasoning mathematically and using** *mathematical concepts, procedures, facts and tools to* **describe, explain and predict** *phenomena. It assists individuals to* **recognise** *the role that mathematics plays in the world and to* **make the well-founded judgments and decisions** *needed by constructive, engaged and reflective citizens."*

# 2. Describe the continuum

When describing indicator you quantify the components of the description. If you look at the communication component of PISA mathematics literacy:

- Level 6: formulate and precisely communicate their findings, interpretations and arguments.
- Level 5: reflect on their actions and formulate and communicate their interpretations and reasoning.
- Level 4: construct and communicate explanations and arguments.
- Level 3: develop short communications reporting these.
- Level 2: make literal interpretations of the results.
- Level 1: no statement

# The framework then typically describes in detail, the cognitive skills and content areas to be assessed. The PISA mathematics example:

- Begins with a detailed presentation of the mathematical thinking skills to be measured.
- It then references the traditional sub-domains of the mathematics curriculum.
- Finally, as a literacy assessment, it describes the context of the assessment items.
- This part of the framework is generally between 10 to 15 pages.

| Description | Page |
| --- | --- |
| Mathematical processes | 9 |
| Formulating situations mathematically | 10 |
| Interpreting, applying and evaluating mathematical outcomes | 12 |
| Fundamental mathematical capabilities underlying the mathematical processes | 12 |
| Mathematical content knowledge | 16 |
| Change and relationships | 17 |
| Space and shape | 18 |
| Quantity | 18 |
| Uncertainty and data | 19 |
| Content topics for guiding the assessment of mathematical literacy for 15 year old students | 19 |

# 2. Describe the continuum

PISA, TIMSS, PIRLS and other large assessments actually develop their descriptions **after** the assessment. They

- Measure students' overall performance.
- Obtain a measure of item difficulty
- Then using a standard setting procedure, describe what students can do on selected regions of the scale.

The examples of students' work are selected by psychometric and subject specialists; the descriptions are written by subject specialists and approved by a governing committee.

# 2. The TIMSS grade 4 Mathematics continuum and its standards

The continuum describes a quantity of a quality.

**Advanced**

Students can apply their understanding and knowledge in a variety of relatively complex situations and explain their reasoning. They can solve a variety of multi-step word problems involving whole numbers, including proportions. Students at this level show an increasing understanding of fractions and decimals. Students can apply geometric knowledge of a range of two- and three-dimensional shapes in a variety of situations. They can draw a conclusion from data in a table and justify their conclusion.

**High**

Students can apply their knowledge and understanding to solve problems. Students can solve word problems involving operations with whole numbers. They can use division in a variety of problem situations. They can use their understanding of place value to solve problems. Students can extend patterns to find a later specified term. Students demonstrate understanding of line symmetry and geometric properties. Students can interpret and use data in tables and graphs to solve problems. They can use information in pictographs

# 2. The TIMSS grade 4 Mathematics continuum

**Intermediate**

Students can apply basic mathematical knowledge in straightforward situations. Students at this level demonstrate an understanding of whole numbers and some understanding of fractions. Students can visualize three-dimensional shapes from two-dimensional representations. They can interpret bar graphs, pictographs, and tables to solve simple problems.

**Low**

Students have some basic mathematical knowledge. Students can add and subtract whole numbers. They have some recognition of parallel and perpendicular lines, familiar geometric shapes, and coordinate maps. They can read and complete simple bar graphs and tables.

# 3. Design a way of gathering data.

The data you need to locate a person on the continuum is gathered using some system.

- How might you gather data to locate a person on the continuum?
  - ▶ Questionnaire or test
  - ▶ Interviews
  - ▶ Observation
  - ▶ Portfolio
  - ▶ And many other ways. Think of all the different types of assessment and research methods.
- What data will you seek?
  - ▶ What, in particular, are you looking for?
  - ▶ What questions will you ask?
  - ▶ What observations will you make?

# Test blueprint template

Typically, for cognitive assessments we specify the distribution of items by cognitive skill and content sub-domain.

| Content | Lower order | | | Higher order | | | Total |
|---------|-------------|---------------|----------|-----------|------------|----------|-------|
|         | Remembering | Understanding | Applying | Analyzing | Evaluating | Creating |       |
|         |             |               |          |           |            |          |       |
|         |             |               |          |           |            |          |       |
|         |             |               |          |           |            |          |       |
|         |             |               |          |           |            |          |       |
| Total 1 |             |               |          |           |            |          |       |
| Total 2 |             |               |          |           |            |          |       |

# Test blueprint template

Typically, for cognitive assessments we specify
the distribution of items by cognitive skill and content
sub-domain.

| Content | Lower order | | | Higher order | | | To |
|---|---|---|---|---|---|---|---|
| | Remembering | Understanding | Applying | Analyzing | Evaluating | Creating | |
| The Universe and its Stars | 2 | | 3 | | 2 | | |
| Earth and the Solar System | 1 | 2 | | | | 1 | |
| Earth Materials and Systems | | 2 | 2 | 2 | 2 | | |
| | | | | | | | |
| Total | 3 | 4 | 5 | 2 | 4 | 1 | |
| Total 2 | 12 | | | 7 | | | |

# The total number of items

The number of items used in a domain depends on the usage of the cognitive assessment.

- If it is a standardized test instrument used for high stakes purposes, then all students should complete the same test.
  - ▶ The number of items is limited by the concentration limits of the students.

- If the aim is to measure the entire curriculum domain then you need more items.
  - ▶ It is not important that all students complete the same test; they can complete different forms.
  - ▶ It is important that there are links between the forms.
  - ▶ Student responses can be aggregated to form a complete high-level understanding of the curriculum mastery.

# TIMSS student testing times

TIMSS achievement tests have two parts. The achievement tests are followed by a break and then a questionnaire.

| Activity | Grade 4 | Grade 8 |
|---|---|---|
| Student Achievement Booklet Part 1 | 36 minutes | 45 minutes |
| Break | | |
| Student Achievement Booklet Part 2 | 36 minutes | 45 minutes |
| Break | | |
| Student Questionnaire | 30 minutes | 30 minutes |

# TIMSS, PIRLS, PISA and similar assessments

These assessments use a very large number of items for each domain.

- If all of the items were administered to a single student, he or she would take 3 to 6 hours to complete the test.
    - ▸ Instead the items are grouped into blocks.
    - ▸ Blocks may be breadth or depth blocks.
- Then tests are assembled so that each test consists of 4 to 6 blocks.
- The blocks are spiraled across the test forms so linkages can be made.

# TIMSS blocks are 18 minutes for Grade 4

There are 14 test booklets, each comprised on 4 blocks of items. The design to the first 8 booklets is:

|  | Part 1 | | Part 2 | |
|---|---|---|---|---|
| Booklet 1 | M01 | M02 | S01 | S02 |
| Booklet 2 | S02 | S03 | M02 | M03 |
| Booklet 3 | M03 | M04 | S03 | S04 |
| Booklet 4 | S04 | S05 | M04 | M05 |
| Booklet 5 | M05 | M06 | S05 | S06 |
| Booklet 6 | S06 | S07 | M06 | M07 |
| Booklet 7 | M07 | M08 | S07 | S08 |
| Booklet 8 | S08 | S09 | M08 | M09 |

# Sometimes a map is used to guide the assessment

There are 164 nodes in the diagram representing all topics on both the foundation and higher tier GCSE curriculum.
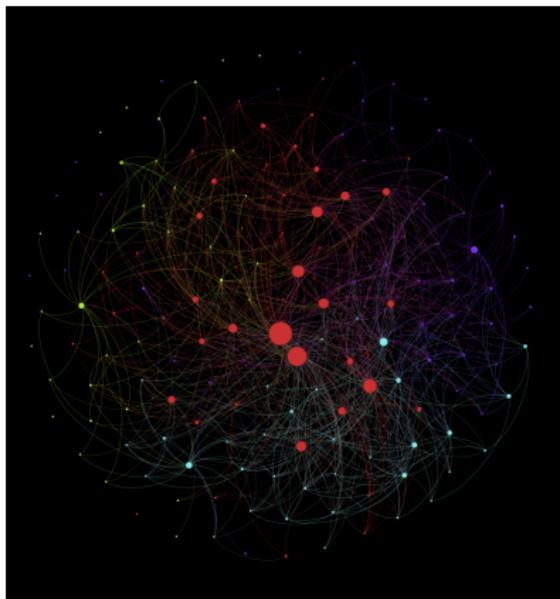
# 4. Gather the data

Now collect the data

- You will need to identify and minimize things in data collection that will invalidate the data.
- Inducements and leading questions are two such threats.
  - ▸ I will pay you 100,000 if you complete this questionnaire on motivation.
  - ▸ Have you always been depressed?
- If you need to compare people, then you will need to specify the conditions for data collection.

.

# 5. Summarize the data

Your data gathering system should have gathered
a lot of data.

- You need to summarise it so that the person is reasonably and justifiably located on the continuum.

- Some of the systems used to summarise data
  - ▸ Test or questionnaire scores.
  - ▸ Rubrics to interpret a response, an essay, a painting, performance or part work.
  - ▸ A qualitative data analysis system.
  - ▸ An expert judgement.

- You might also use a measurement theory
  - ▸ Item response theory like in PISA, TIMSS and PIRLS.
  - ▸ Or one of the other measurement theories.

# 6. Locate the person on the scale

The purpose of the process has been to locate each person onto the scale.

- Use your summary to locate the person on the scale.

- In some cases, that summary may be a number that places the person within a specific part of a region.

    - Once the person has been placed on the continuum,reflect: does the data really support the location of the person there?

# 7. Critique and evaluate

The indicator will not be perfect. There may be others ways of conceptualizing the indicator. The data gathering system you have used will have strengths and limitations. So will your data summary system.

- What are the strengths and limitations of each part of the process?

- You have made a number of assumptions during this process. What are they and to what extent are they reasonable?

- Is the measure affected,in some baised way, by gender, culture or some attribute which should not have this effect?