PISA for Development Overview

PISA for Development IAG Meeting

13 March 2015

Paris, France

Strand A: ETS

Ann Kennedy

Key Aspects of Project Implementation

- Overview of PISA for Development

- Assessment Instruments

- Translation, Adaptation, and Verification

- NPM Perspectives

- Target Population and Sampling

- Survey Operations

- Coding

- Data Management

- Analysis and Report Preparations

Overview

- Project overview

- Roles, responsibilities, and resources

- Project implementation

- Technical standards

- Questions and discussion

Project Overview

- Collaboration among OECD Secretariat, Contractors, and National Centres

- International Assessment

- Single Project Timeline

- International Technical Standards and Guidelines

Contractor Roles and Responsibilities

Contractor Roles and Responsibilities Strand A

- Development and Implementation of the Cognitive Instruments

  - Educational Testing Service (ETS): project management, survey design, data analysis and scaling

    - Pearson: Cognitive Frameworks and Subject Matter Expert Groups

    - Westat: Sampling and Survey Procedures and Operations

    - cApStAn: Linguistic Quality Assurance

Contractor Roles and Responsibilities Strand B

- Strand B: Development and Implementation of the Contextual Questionnaires

  - The Learning Bar: Contextual Questionnaire Frameworks, Questionnaire Expert Groups, Analysis of Questionnaire Data

NPM Roles and National Centre Resources

National Project Manager (NPM)

- Oversees all national tasks related to the development and implementation of PISA-D

- Monitors or implements activities in all areas, in particular:
  - Sampling
  - Instrument review
  - Translation and adaptation of instruments
  - Survey operations
  - Coding
  - Data management
  - National report planning and dissemination

Profile of NPM

- Oral and written communication skills in English

- Experience in planning, organizing and conducting large-scale surveys

- Familiarity with a variety of areas including statistics, sampling, survey procedures and data management, procedures and analysis

- Experience with Microsoft applications such as Word and Excel

- Knowledge of and experience dealing with government agencies, school principals, parents and teachers within their own countries
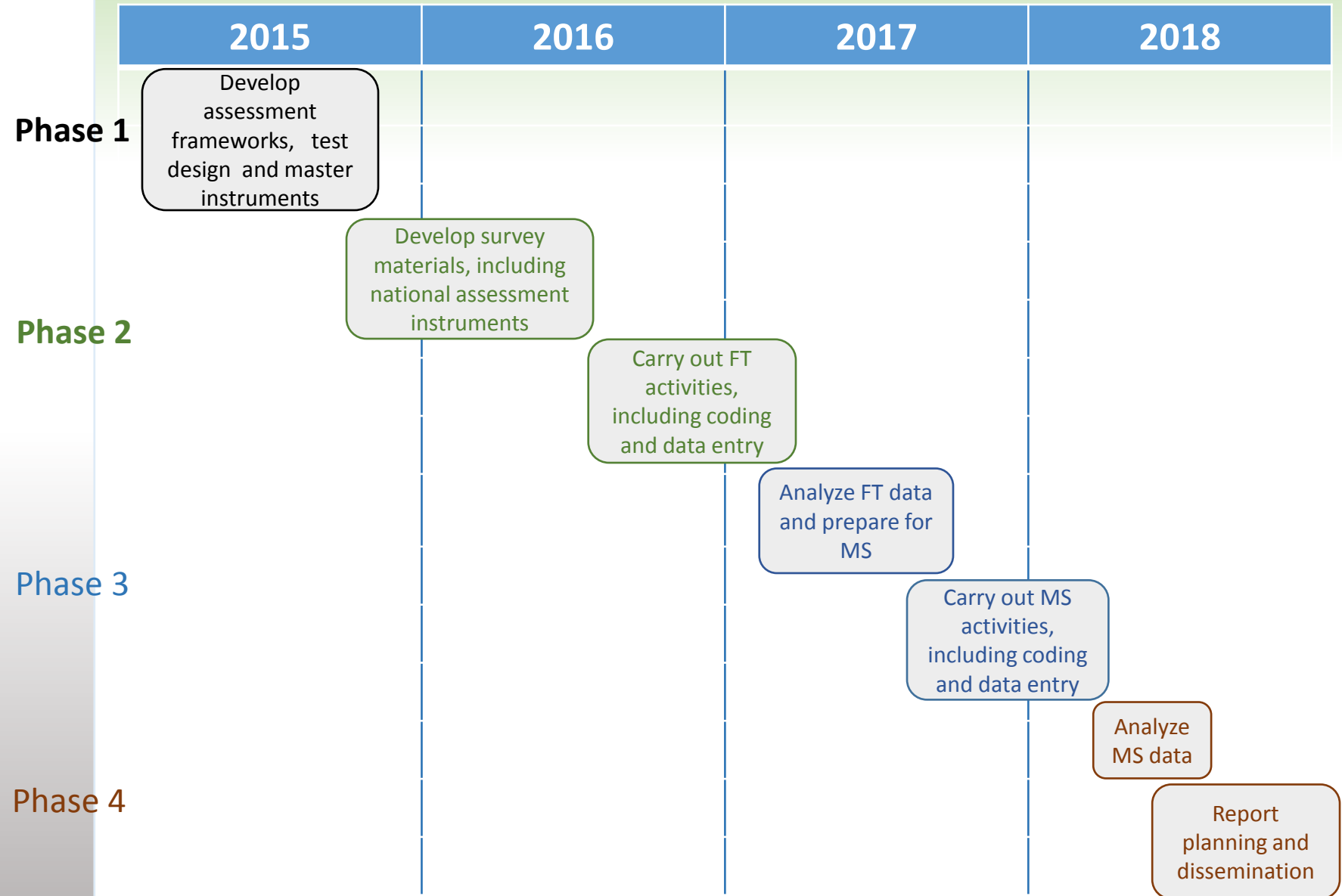
Additional Support at the National Centre

- Administrative Activities

- Sampling

- Instrument review and finalization

- Translation and adaptation of instruments

- Survey operations

- Coding

- Data management

- National report planning and dissemination

Facilities

- PCs with Windows and Microsoft Office Professional applications

- Access to a reliable, high-bandwidth Internet connection and email facilities

- A secure space for conducting coding operations and data entry

- A secure space for storage of electronic files and paper materials

Project Implementation

Summary

2015	2016	2017	2018

Phase 1

Develop assessment frameworks, test design and master instruments

Phase 2

Develop survey materials, including national assessment instruments

Carry out FT activities, including coding and data entry

Phase 3

Analyze FT data and prepare for MS

Carry out MS activities, including coding and data entry

Phase 4

Analyze MS data

Report planning and dissemination

NPM Meetings

| Meeting	Description	Date
1	Review draft cognitive and contextual assessment frameworks and characteristics of the available item pools; capacity building activities around item development and selection	Sept. 2015
2	Adaptation, translation, and verification of all survey materials and sampling overview	Jan. 2016
3	Student sampling, FT survey operations, and test administrator training	April 2016
4	Scoring training and data management	July 2016
5	Analysis and interpretation of FT results, plans for analysis of contextual questionnaires, and preparation for MS data collection	May 2017
6	MS survey operations, MS student sampling, test administrator training, scoring, and data management	July 2017
7	Scaling methodology, and data analysis and reporting tools	Mar. 2018
8	Analysis and interpretation of MS results and preparations for reporting and dissemination	July 2018

Communication is KEY

- Day-to-day

  - Email, phone, videoconference

- PISA-D portal

  - Secure site for sharing project information, documents

- Progress reports

- Meetings (NPM, IAG)

- Additional training/workshops

Technical Standards

Technical Standards

- Standards for assessment preparations (sampling, translations and adaptations), data collection activities, and data management and submission

- Internationally recognized best practices

- Adherence to standards contributes to creating a quality international dataset that allows for making valid cross-national inferences

Technical Standards: Goals

Consistency	Precision	Generalisability
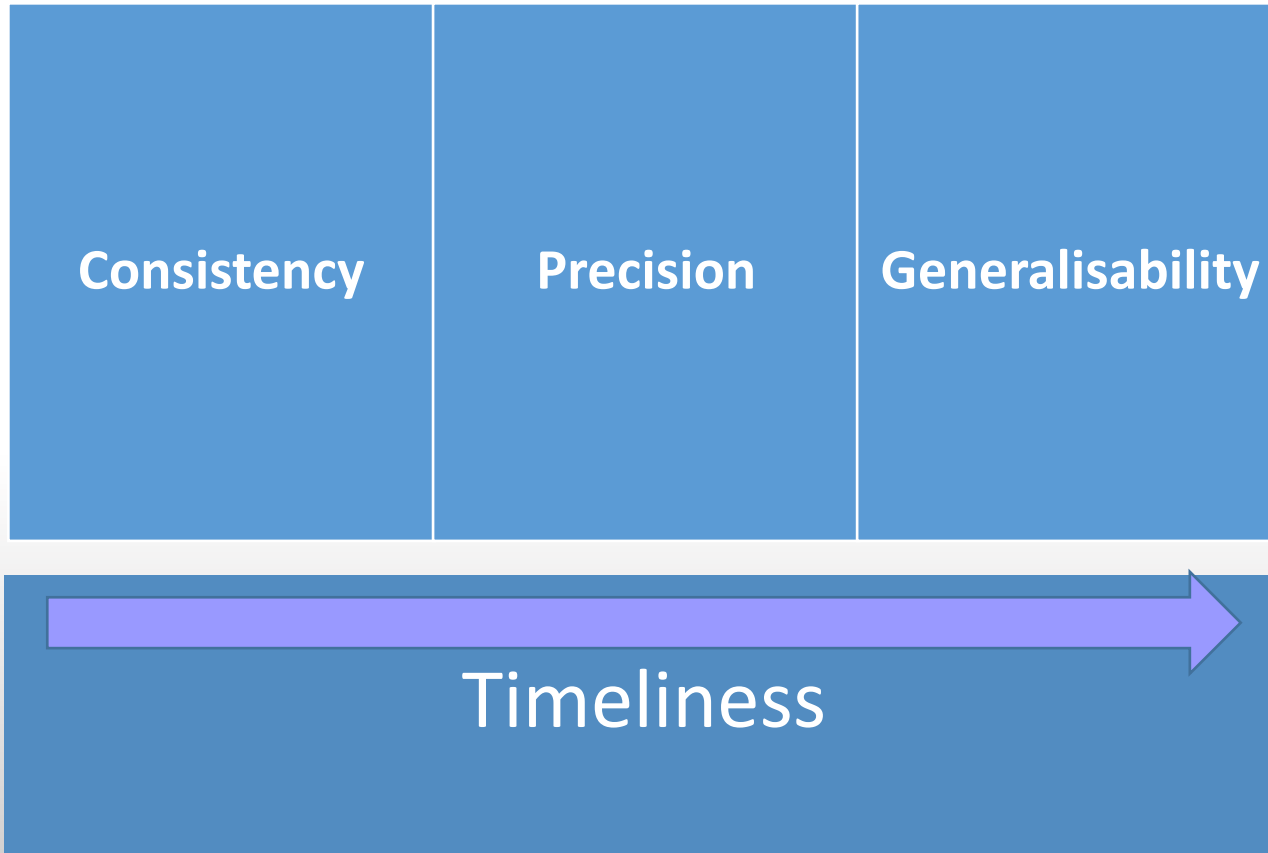
Timeliness

Technical Standards: Types

- Data Standards
  - Quality of data, or assurance of that quality

- Management Standards
  - Operational objectives met in a timeline and coordinated manner

- National Involvement Standards
  - Ensure that internationally developed instruments are widely examined for cross-national, cross-cultural, and cross-linguistic validity

Key Standards: Data

- Target population and sampling

- Language of testing

- Field Trial participation

- Translation and adaptation of tests, questionnaires and manuals

- Test administration (including printing, security of materials, and quality monitoring)

- Coding

- Data submission

Key Standards: Management

- Communication with international contractors

- Schedule for submission of materials

- Management of data

- Archiving of materials

Key Standards: National Involvement

- Feedback to international contractors on the development of instruments

- Promotion of participation, effective implementation, and dissemination of results

Questions and Discussion

Frameworks and Expert Group

Prof. John de Jong

Programme Director, PISA, Pearson

Expert Groups

- Three expert groups:
  - Reading
  - Science
  - Mathematics

- Expert groups combine PISA knowledge, assessment expertise and local education system understanding

- Responsible for:
  - PISA Framework review and extension
  - Item review and selection

Frameworks

- Each expert group will review, redraft and present new and extended frameworks

- Aim is to ensure that the PISA-D survey links to the main PISA scales

- Some of the considerations and challenges are described in the following slides

Reading literacy

Challenges for Reading Literacy

- Reading literacy: reading for understanding and using texts in appropriate contexts to engage with information

- Students in developing countries may struggle more with basic reading processes
  - Fluency, literal interpretation, finding and extracting the main ideas, inferencing

- Poor performance on higher-level skills such as interpreting, reflecting and evaluating may be masked by difficulties with basic reading processes
  - Will not result in useful feedback to countries on critical skills

- Current framework not as sensitive to these basic processes

Reading Literacy Modifications

- Use 2009/2015 reporting subscales
  - Access and Retrieve, Integrate and Interpret, Reflect and Evaluate
  - Note: these subscales have been modified & renamed for PISA 2018, but map back to 2009/2015

- Extend definitions downward for each subscale on what it means to be at lower levels of skills

- Select tasks that focus more on literal meaning simple comprehension.

Access and Retrieve

Potential modifications include a focus on tasks in which student must

- Locate single piece of explicitly stated information in a prominent position in a short, syntactically simple text with a familiar context

- Locate one or more independent pieces of explicitly stated information

- Discriminating between two pieces of information with explicit differences

- Select pieces of text which require minor inferences

Integrate and Interpret

Potential modifications include a focus on tasks in which student must

- Read and understand the meaning of simple sentences

- Interpret literal expression in text

- Recognise the main idea, theme or author's purpose in a text

- Recognize relationships between two pieces of information in adjacent sentences

- Comparisons or contrasts based on a single feature in the text

Reflect and evaluate

Potential modifications include a focus on tasks in which student must

- Make a simple connection between information in the text and common, everyday knowledge with explicit prompting

- Evaluate or reflect on a text when student is explicitly directed to consider relevant factors

Scientific literacy

Scientific Literacy Considerations

Three Competencies:
- Explaining phenomena scientifically;
- Evaluate and design scientific enquiry; and
- Interpret data and evidence scientifically.

- Each of the three competencies represent competency/literacy in the area of science and should be adjusted first.

- Then consider appropriate knowledge, contexts, and depth of knowledge.

- What portion of each of those competencies would be appropriate for developing countries?

Explaining phenomena scientifically

- Full competence requires an understanding of how knowledge has been derived and a high level of confidence about any scientific claims.

- Modifications might include:
  - Recall appropriate scientific knowledge (but no application);
  - Identify explanatory models and representations (maybe some level of use but no generation);
  - Make a simple prediction but not justify it;
  - Offer explanatory hypotheses and Explain the potential implications of scientific knowledge for society would likely be too high level

Evaluate & design scientific enquiry

- Full competence requires recognition of the significance of previous research in judging the value of any given scientific enquiry. Such knowledge is needed to situate the work and judge the importance of any possible outcomes.

- Modifications might include:
  - Identify a question explored in a given scientific study;
  - Distinguish simple questions that are possible to investigate scientifically;
  - Propose exploring a simple question scientifically;
  - Evaluate ways of exploring a given question scientifically and Describe and evaluate a range of ways that scientists use to ensure the reliability of data and the objectivity and generalisability of explanations – these may not be appropriate for the lower end of the performance scale.

Interpret data & evidence scientifically

- Full competence requires ability to analyse and evaluate scientific data, claims and arguments in a variety of representations and draw appropriate conclusions.

- Modifications might include:
  - Transform data from one representation to another and Analyse and interpret data and draw appropriate conclusions – these can be done with simple data
  - Identify the assumptions, evidence and reasoning in some science-related texts;
  - Distinguish between arguments which are based on scientific evidence and theory and those based on other considerations and Evaluate scientific arguments and evidence from different sources (e.g. newspaper, internet, journals) – these might not be appropriate to assess.

Knowledge & Complexity

- Knowledge recommendations:
  - Continue with 2015 categories of content, procedural, and epistemic
  - Adjust knowledge tested especially around epistemic for lower end of performance scale

- Depth of Knowledge
  - Added for 2015, cognitive complexity of item sets can also be considered.
  - Need to make sure items are on the lower end of the complexity continuum.

Contexts

- Contexts recommendations:
  - Continue with contexts at the personal, local/national, and global.
  - Choose contexts appropriate to experiences in developing countries
  - Consider what exposure students have to global events

Mathematical literacy

Mathematics Challenges

- Students in developing countries may not have had the exposure to higher level mathematical concepts or may be struggling with the concepts.

- Understanding and application of mathematical concepts should be the minimum expectation

- Extending, analysing, and justifying will likely be much more challenging and should therefore be limited

- Presenting students with problems for which they lack skills will not provide usable data

Modifications

- Much of the framework needs little or no modification

- Some of the more challenging parts of the framework can be slightly modified to adapt to the challenges presented by students in developing countries

- Due to the richness of the framework, removing the challenging components can still result in a rigorous, valid assessment

- Some examples follow on the next slides

Modifications – Detail 1

Read, decode, and make sense of statements, questions, tasks, objects, images

Remove: ~~or animations (in computer-based assessment) in order to form a mental model of the situation~~

- All mathematics students should be able to show some degree of proficiency in reading, decoding, and understanding
- Forming a mental model is a higher level cognitive ability which would be especially challenging for students still working toward mastery of mathematical skills

Modifications – Detail 2

Use mathematical tools to ascertain the reasonableness of a mathematical solution

***Remove*: ~~and any limits and constraints on that solution, given the context of the problem~~**

- Determining reasonableness of a solution is a critical part of mathematics
- Understanding constraints can be a subtle nuance that eludes students who are still developing mathematical literacy

Modifications – Detail 3

Explain, defend or provide a justification for the processes and procedures used to determine a mathematical result or solution

Remove: ~~Connect pieces of information to arrive at a mathematical solution, make generalisations or create a multi-step argument~~

- Providing justification may be at the higher end of the skills for which students can be consistently successful
- Being able to synthesize and make generalisations with consistent success is unlikely for students still mastering mathematical skills

Call for experts

- We would like to suggest that each of the countries participating in PISA for Development nominate at least two people each from one of the domains:
  - Reading
  - Mathematics
  - Science

- These people should
  - have a good working command in English
  - Work in education in their country
    - Teachers
    - Curriculum developers
    - Teacher training
    - Ministry of Education
    - Etc.
  - Please write to: john.dejong&pearson.com

Questions and Discussion

Overview

- Goals

- Present PISA-D Contextual Questionnaire Framework outline

- Describe process to respond to each country's needs and emphasize timelines

Goals

- Provide each PISA-D country with data to inform policy

- Build capacity to use PISA-D data to inform policy in each country

PISA-D Contextual Questionnaire Themes	Modifications to previous PISA CQ Frameworks
- early learning opportunities - language at home and school - family and community support - quality of instruction - learning time **student socioeconomic status** **school resources**	- Add items on students' early learning experiences, - Add items on students' familiarity with the language of the test - Measures of parental involvement, social capital and cultural capital - Measures of role of other community members and of types of community - Enhanced measure of school attendance - Add items on participation in formal and informal labor market

Policy Framework
The Learning Bar Inc.

Policies about …	Key policy questions
Student Performance	What are the current levels of student performance? What are the current levels of student growth? What are the long-term trends in student performance and growth? To what extent do levels of performance and growth vary within and among schools?
Drivers of Performance	To what extent are levels of student performance and student growth related to school processes and practices?
Intervention Forecasting	Can we identify students that require specific types of interventions?
Resource Allocation	To what extent do schools vary in their levels of socioeconomic status? To what extent are students from certain sub-populations (e.g., low SES, visible minorities) concentrated in certain schools?
Equality	To what extent do performance levels and growth rates vary among students from certain sub-populations?
Equity	To what extent do students from differing sub-populations have access to key school resources and processes?
Successful Transitions	Are students acquiring the personal assets they need to make successful transitions at key stages of the schooling?

Policy Approach

- To determine the content for the student and school contextual questionnaires we can start with a set of policy questions, and discern what kinds of data are required to address these questions.

PISA policy products

PISA offers three types of policy-relevant products:

1) Indicators that monitor the functioning, productivity and equity of education systems;

2) Knowledge on factors that determine educational effectiveness;

3) A comparative database that allows researchers worldwide to study basic as well as policy-oriented questions.

PISA Policy Themes

- Policy issues that have been addressed by previous PISA cycles can be categorized into four themes:

  1) selecting and grouping students;

  2) resources invested in education;

  3) school governance, assessments and accountability; and

  4) the quality of learning environments.

Selecting and grouping students

PISA contextual questionnaires collect information on:

- children language and early schooling experiences;
- age of entry to school;
- grade repetition;
- special schools and school programs;
- selection into different types of schools;
- ability grouping within classrooms and within schools;
- parental choice of schools and school programs; and
- funding arrangements for private schools or charter schools.

Resources invested in education

- Policy questions concerning resources invested in education inquire about the effects of financial, human, material, and time resources.

- In PISA, most of the data on school resources are collected from the school principal or site administrator with the school questionnaire.

School governance

- Questions about school governance are concerned with how student performance is related to:
  - school autonomy;
  - school choice;
  - private and public schooling;
  - principal leadership; and
  - parental involvement.

- PISA framework also includes questions about the relationship between student performance and various approaches to assessment and accountability.

Quality of learning environments

- This theme pertains to truancy and school climate.

- The questions on school climate are based on a traditional 'school-effects' approach to analysis, with data collected on factors such as teacher-student relations and classroom disciplinary climate.

PISA-D country involvement and timelines

- Work with The Learning Bar and others to identify policy questions unanswered by PISA-D contextual questionnaire data. (**June 2015**)

- Work with The Learning Bar to identify questions (or measures) to include in the PISA-D contextual questionnaires for that country. (**July 2015**)

- TLB presents suggestions to QEG (**August 2015**)

PISA-D country involvement and timelines

- Updated version presented to NPM's (**September 2015**)

- Final selection of country items in collaboration with TLB (**November 2015**)

- Beginning of the translation and verification process (**November 2015**)

QUESTIONS AND DISCUSSION

Robert Laurie

robert.laurie@unb.ca

Telephone +1 506 447 3435

Atlantic time zone (Canada) UTC -4:00

PISA for Development Assembling and Printing Paper Booklets

PISA for Development IAG Meeting

13 March 2015

Paris, France

Strand A: ETS

Ann Kennedy

Assembling and Printing Cognitive Booklets

Organization of Cognitive Assessment Materials

A unit is defined as a single stimulus and all its associated questions.

A cluster includes multiple units and is often based on estimated testing time (i.e., PISA uses 30-min clusters)

A testing booklet includes a cover page, general instructions and multiple clusters. For example, PISA is based on 120 minutes of testing time, so each booklets includes 4 clusters.

Unit
- Q 1
- Q 2

Cluster
- Unit X
- Unit Y

Booklet
- Cluster M
- Cluster N
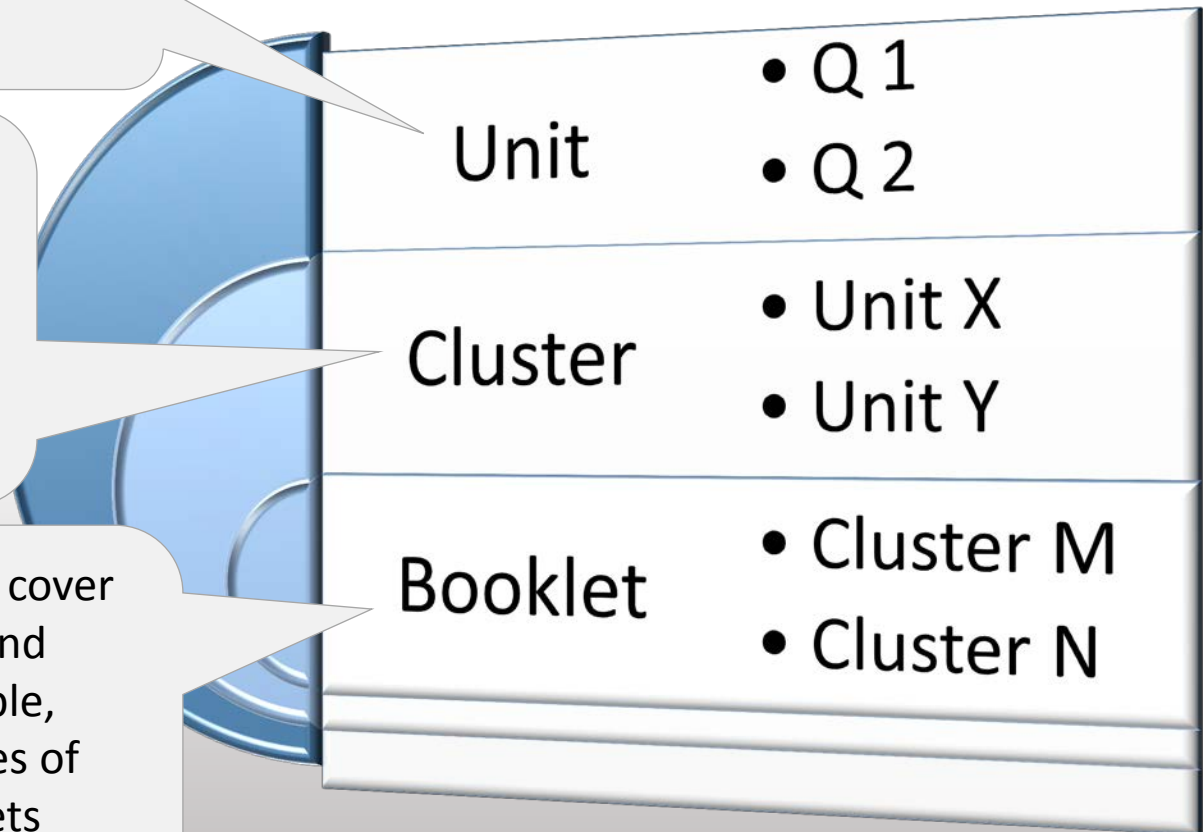
Assembling Cognitive Booklets

- Countries translate materials and general instructions, adapt cover pages and select paper size (i.e., A4 vs. Letter)

- Contractors centrally assemble testing booklets and submit to countries for review and approval

- Countries review and approve booklets

- Contractors finalize files and send final print-ready PDF versions of booklets

- Countries print booklets, following guidelines

Printing Cognitive Assessment Booklets

- All paper booklets must be printed according to the specifications provided by the Consortium.

- Countries will be asked to send a "pre-printing" version for approval

- The booklets should be printed exactly as they are at the time of the "pre-printing" approval.

- Countries are required to send a set of printed booklets to the contractor.

Assembling and Printing Contextual Questionnaires

Assembling Contextual Questionnaires

- Countries translate materials and general instructions, adapt cover pages and select paper size (i.e., A4 vs. Letter)

- Countries enter approved translations into questionnaire templates and submit to contractors for review

- Contractors review and approve booklets

- Countries finalize files and create final print-ready PDF versions of booklets

- Countries print booklets, following guidelines, and send to contractor for final optical check

Questions and Discussion

Summary

- Support in producing high quality instruments

- Linguistic [and functional] equivalence

- Existing resources

- Monitoring tools and documentation

- Verification and final check

- Trainer of trainers approach

Support to National Centres

- PISA National Centres and cApStAn
  work together since 1999 (PISA 2000 FT)

- to produce high quality assessment instruments

- as equivalent as possible across languages,
  across countries and across cultures

- Accumulated experience in identifying

  - Language-dependent meaning shifts

  - Culture-dependent perception shifts

  which may cause country/item interactions
  or language/item interactions

cApStAn's Track Record

- PISA 2000, 2003, 2006, 2009, 2012, 2015

- PISA 2015 = 74 countries, 46 languages, 95 national versions [185,000 words each, but 60,000 words trend material]

- But also: UNESCO/LAMP

- OECD/TALIS, OECD/PIAAC,

- STEP (WORLD BANK), PASEC,

- PIRLS, TIMSS

- National exams e.g. G6-G9 in Afghanistan

About 11 875 000 words to check last year

Linguistic Equivalence Functional Equivalence

- Studies that rely on translation assume that the same question can be asked in the different language versions produced (Willis et al, 2010)

- Translation/adaptation should retain the sense or meaning across languages

- quantity and quality of information should remain the same

- register is equivalent

- difficulty level of questions remains the same

Features that drive Equivalence

- Some correct translations may pose a problem due to reasons that are not immediately apparent

EXAMPLES:

- Numeracy skills are being measured, but the reading load changes across languages =>
measure a mix of reading skills and numeracy skills?

- Rendition is too literal (word-for-word translation): may fail to replicate measurement goals
-- although the syntax is reproduced

- Language-specific clues are added or removed

- Proportional length of key and distracters changes

- Literal matches between text and question are lost

Existing Resources (I)

- General Translation and Adaptation Guidelines

- Detailed item-per-item guidelines

- Tried and tested monitoring tools

- Communication protocols

Existing Resources (II)

- For all PISA items (test and QQ)
  - ENG and FRA source
  - Existing ESP national versions

- For test items from other sources
  - cApStAn develops FRA + ESP base versions

- For QQ items from other sources
  - cApStAn develops FRA + ESP base versions

Documentation

- In PISA, everything is documented

- Entire translation, adaptation and verification history of each item in each language

- In Excel Adaptation Forms

- Different processes for tests, questionnaires and coding guides

- Require ongoing interaction between National Centres and the Contractors

Different Procedures (I)

TEST MATERIALS

STRAND A

1. Adapted from FRA source
2. Adapted from ENG source
3. Adapted from ESP base version

ENGLISH SOURCE VERSION	CONSORTIUM RECOMMENDATION	Changes vs SPANISH BASE VERSION		
		SPANISH BASE VERSION	SPANISH TARGET VERSION	DESCRIPTION AND JUSTIFICATION OF CHANGE
Nicole	Adapt to a common name in your language	Nuria		
It is about 70% of...		Esta se sitúa en torno al 70%		

Different Procedures (II)

TEST MATERIALS

4. New language

STRAND A

Double Translation

❑ Two different translators each produce their version

Reconciling Process

❑ A senior translator with domain expertise merges two independent translations into a final version
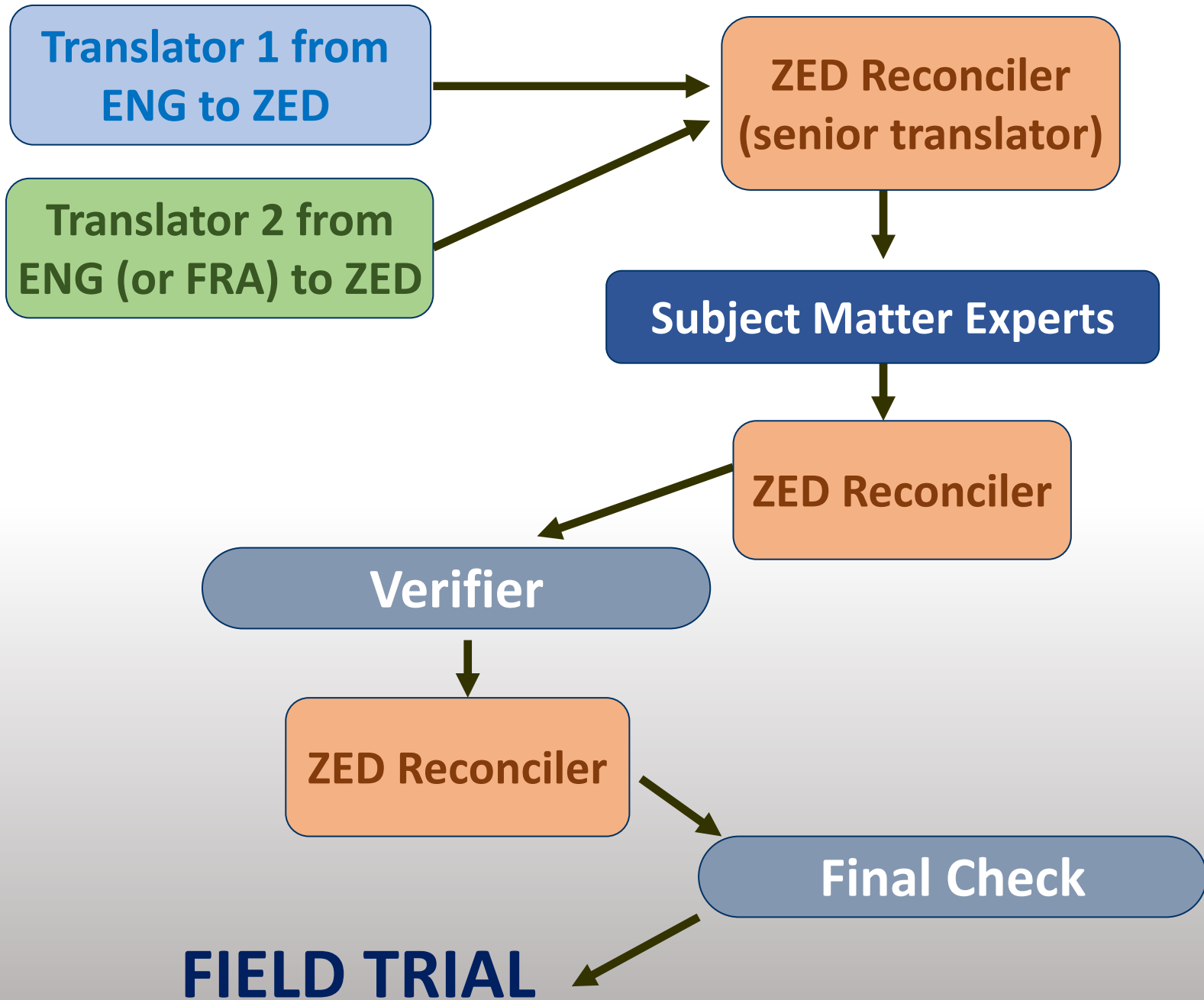
Different Procedures (III)

QUESTIONNAIRES

STRAND B

1. Adapt from FRA source
2. Adapt from ENG source
3. Adapt from ESP base version
4. New language (double translation)

❑ In all cases: more extensive adaptation to local context and situation needed

❑ Subject to negotiation with Strand B Team

Questionnaire adaptation

1	2	4	5	6	9	10	11	12	13	14
Int. Question Number	International English Version	Int. Variable ID	Int. Code	Nat. Question Number	National version (text)	English translation of the national version	Justification for proposed changes; national centre comments	Consortium Comments	Agreement Status	Verifier
Q8	Some schools organise instruction differently for students with different abilities. What is your school's policy about this for students in <national modal grade for 15-year-olds>?			Q6	1학년(중학교의 경우 3학년)	first grade(third grade in case of middle school)	See Cell L: 59	See Cell: N59	Agreed	ok
	(Please tick one box in each row)				(각 항목에 대하여 한 곳에만 표시하십시오.)	literal translation	unified translation of same expression		Agreed	ok
a	Students are grouped by ability into different classes	SC08Q01			학생들을 능력에 따라 서로 다른 학급으로 편성한다	literal translation			Agreed	ok
b	Students are grouped by ability within their classes	SC08Q02			학생들을 학급 내에서 능력별 소집단으로 편성한다	literal translation			Agreed	ok
	For all subjects		1							ok
	For some subjects		2							ok
	Not for any subjects		3							ok
	SECTION B: STAFFING									ok
	<school reminder note>								Agreed	ok
Q9	How many of the following are on the staff of your school?									ok
	Include both full-time and part-time teachers. A full-time teacher is employed at least 90% of the time as a teacher for the full school year. All other teachers should be considered part-time.				정규 교사와 시간제 교사를 모두 포함시키하십시오. 정규 교사는 한 학년 수업의 최소 90% 이상 시간 동안 교사로 고용된 분을 말합니다. 그 밖의 모든 교사들은 시간제 교사로	ion				changed a typo: 시키하십시오-> 시키십시오. OK now.

National version + back translation + rationale

National adaptation required, indicated by <angle brackets>

Verification

- Segment by segment comparison of national version with source version

- Check compliance with translation and adaptation guidelines

- Report deviations and potential issues, using standardized intervention categories

- Suggest corrective action (propose alternative wording)

All in a cooperative spirit

Country: UAE
Language: Arabic

TEST ADAPTATION SPREADSHEET
CS635 SAVE THE FISH

LOCATION	SEGMENT ID	ENGLISH SOURCE VERSION	ITEM-SPECIFIC TRANSLATION/ADAPTATION GUIDELINE	COUNTRY COMMENT (ADAPTATIONS, DOUBTS,	VERIFIER INTERVENTION	VERIFIER COMMENT	CONSORTIUM REFEREE COMMENT	CORRECTION STATUS	COUNTRY POST-VERIF COMMENT	FINAL CHECK
Directions	82	Refer to … on the right. Use drag and drop to answer the question.	Recurring instructions Reference should be to subheading in stimulus	رجع إلى نص" دافقوا على الأسماك " على الجانب الأيسر استخدم السحب و الإفلات للاجابة عن السؤال	Inconsistency	not consistent versus other occurrences. changed by ver	Please **make sure** to keep the verifier correction 23/08/13 Thanks		ok, just small changes to be consistent	
Question										
Table	30 - 33	Herbivores Primary carnivores Secondary carnivores Producers	Use the correct terms as found in e.g. Biology textbooks "producer", in the context of biology, means "an organism, especially a green plant, that builds up its own tissues from simple inorganic compounds" - the translation should be consistent with unit CS607 Birds And Caterpillars	اكلت الأعشاب اكلت لحوم أولية اكلت لحوم ثانوية طائنات منتجة للغذاء	OK					
	2				Register/Wording	w of data to run]" was translated as w of the results to ther example".	Please keep the verifier correction	REQUIRES FOLLOW-UP		OK
		Graph								

Dropdown list:
OK
Added info
Missing info
Matches & Patterns
Inconsistency
Adaptation issue
Register/Wording
Grammar/Syntax
Mistranslation
Guideline not followed
Left in source language
Minor linguistic defect
Erratum or Update missed

REQUIRES FOLLOW-UP

Excel Adaptation Forms as Repositories for documentation

Typology of issues reported
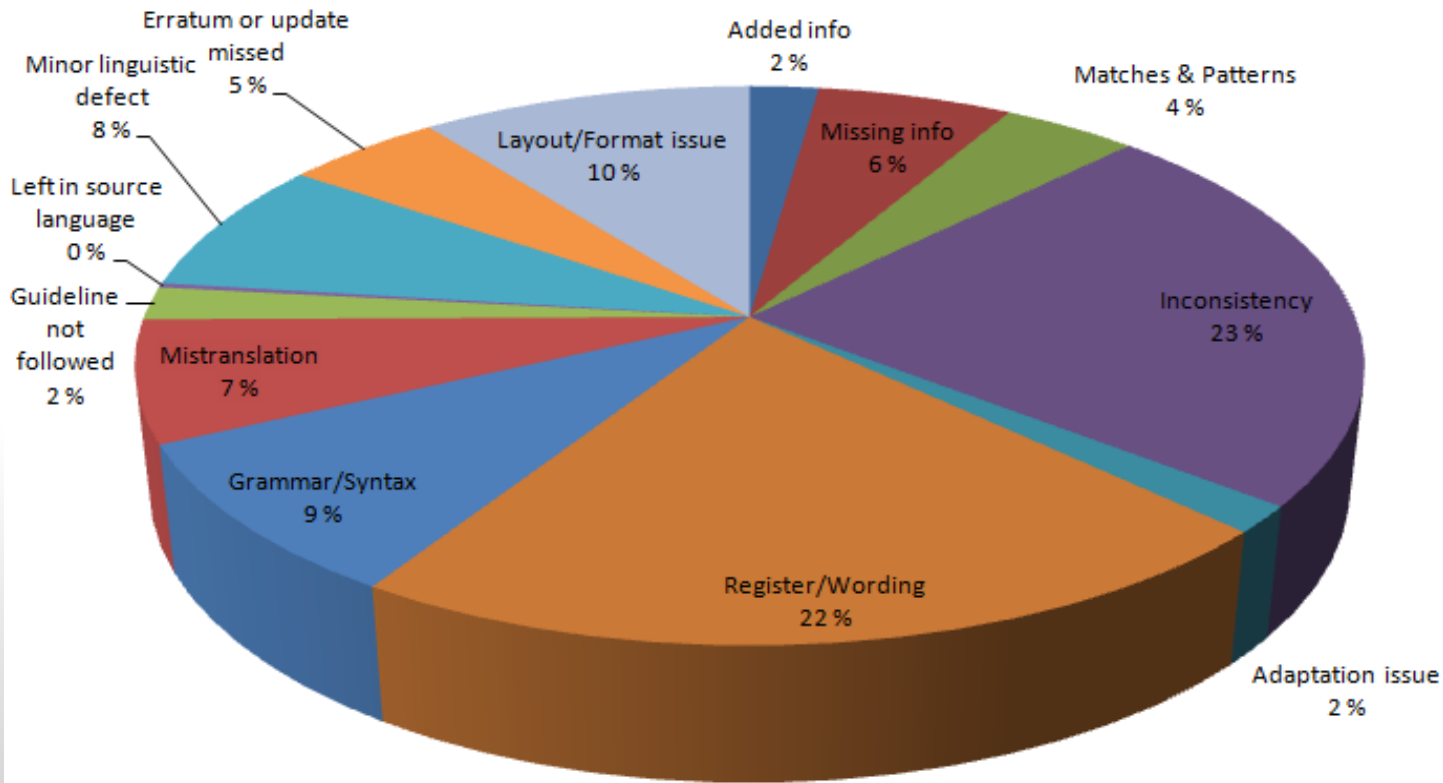
Per national version, per unit, per item, per domain, per language

Training the Trainer Approach
what capacities can you build?

- **Item writers** who gradually become more aware of potential translation hurdles in items they develop

- **Subject matter experts** who may gradually acquire more linguistic expertise

- **Dedicated linguists** who may gradually acquire more subject matter expertise

- **Project managers** who maintain centralised documentation

- **Researchers** who exploit this data

Training the trainer approach

- Decide who is best suited to attend the translation workshops in January 2016

- Most Senior Translator? Project Manager?

- Appoint a Coordinator of the multistep translation, adaptation and review design

- Responsible for centralised documentation, version management, timeline management

- Foster collective ownership of quality translation / quality adaptation

Questions welcome

pisa.verif@capstan.be

steve.dept@capstan.be

A Review of Sampling in PISA

PISA for Development IAG Meeting
13th March 2015

Keith Rust

Westat

Overview of Presentation

- Sampling Task Timeline

- Plan Main Survey (MS) first then Field Trial (FT)

- Why sample?

- How to sample?

- PISA target population

- PISA MS sample design

Overview of presentation

- Response rates

- Student sampling

- Field Trial

  - We will generally discuss procedures used in PISA, placed in the context of the PISA for Development timeline. These are not necessarily the exact procedures that will be implemented in PISA for Development

Sampling Task Timeline

- Field Trial Sampling activities

- 2015: Sampling Task 0 (ST0); new country education system questionnaire

- Jan. 2016: 2nd NPM meeting (FT sampling tasks)

- Jan./Feb.-Jun. 2016: ST1-ST5

- Apr. 2016: 3rd NPM meeting (FT student sampling)

- Sept. – Nov. 2016: FT data collection

- Dec. 2016-Jan 2017: ST6

Sampling Task Timeline

- Main Survey sampling activities

- Dec. 2016-June 2017: ST7-ST11

- May 2017: 5th NPM meeting (MS sampling activities and student sampling)

- Sept. – Nov. 2017: MS data collection

- Dec. 2017-Jan. 2018: ST12

Main Survey first and then Field Trial

- All things in the Field Trial are driven by plans for the Main Survey

- Decide on Main Survey sampling design and parameters first and then design an appropriate Field Trial

Thinking about the MS: Why Sample?

- Usually not feasible to assess all PISA students

- Need to assess about 4500 of them, chosen in such a way so that they represent the entire population

How to Sample for the MS

- Usually not feasible to have a list of all PISA students to select from, nor practical.

- PISA has a stratified two-stage sampling design: schools are sampled first, and then PISA students are sampled within sampled schools.

PISA Target Population

- First step in sampling is the need to define the target population of PISA students

- Generally referred to as "15-year-olds"
- More specifically, but importantly:
  - students of ages between 15 years and three completed months and 16 years and two completed months at the beginning of testing
  - historically, this meant that for the many countries that tested in April 2000 for the first cycle of PISA, the population was students born in 1984
  - attending ANY educational institution in the country
  - restricted to grades 7 and above

PISA Target Population

- Defined by student birth dates

- Student birth date definition needs to be adjusted for the chosen time of testing.

- If assessments are conducted in October 2017, the preferred definition for the population birthdates is:
  - born between July 1, 2001 and June 30, 2002.

Sample Design

- Stratified two-stage sample design

- Schools sampled probability proportional to size (PPS)

- Size = best estimate of school enrollment of PISA students

- Students within schools are sampled with equal probability, with the same number of students sampled per school in a stratum (TCS – Target Cluster Size)

Sample Design

- Why stratification?

  - To reduce sampling variance $\Rightarrow$ reduced sampling variance means increased precision in estimates produced from the sample

  - To ensure adequate representation of all major different school types and geographic regions, so that estimates can be made for these groups

Sample Design

■ Why the two-stage sampling design?

- Feasibility

- Cost reduction

Sample Design

- Why are schools sampled PPS and students sampled equal probability?

  - To produce approximately equal student weights; the student sample resembles the population

    - Undesirable to have any sampled student with greater weight than another sampled student

    - Unequal student weights increase sampling variance

Sample Design

- Sample size:
  - At least 150 sampled schools

  - At least 5250 sampled students

  - At least 4500 assessed students

What does the National Centre need to do?

- Decide on the testing period

- Determine the PISA population birth date definition

- Determine which school types could have ANY PISA students at the time of assessments

- Compile a list of all such schools $\Rightarrow$ school sampling frame

What does the National Centre need to do?

- Determine the best variables for the stratification of schools on the school frame, which possibly help to explain differences in student performance.

- Determine the best way to estimate, for each school on the school frame, the number of enrolled PISA students at the time of school sampling.

What does the National Centre need to do?

- Communicate with the OECD and contractors any questions that they have about population definitions, creation of the school frame, stratification, and the best estimate to use for enrolled PISA students in each school, after reviewing the relevant documentation.

Exclusions from the target population

- PISA aims to minimize exclusions

- Should not assume that acceptable exclusions from other national surveys will be acceptable for PISA

- Exclusions can be both at the school-level and at the student-level

Exclusions from the target population

- Within-school student-level exclusions:

  - Try not to exceed the maximum allowable of 2.5% of the student target population

    - Functionally disabled students who cannot take the test

    - Students with cognitive, behavioral or emotional disability who cannot take the test

    - Students with insufficient language experience

    - Students for whom there are no test materials available in the language of instruction

School Sampling

- Information required from the National Centre for FT and MS school sampling will be mainly collected on several sampling forms (ST0-ST11)

- The use of standard forms ensures that there is no misunderstanding of intentions, or oversight of requirements

- For both the FT and the MS the contractor staff will select the schools in close collaboration with the National Centre

School Response Rate Requirements

- PISA requires a minimum of 85%(weighted) of originally sampled schools to participate

- PISA requires at least 150 participating schools, including replacements

- If 85% of originally sampled (eligible and non-excluded) schools do not participate, an acceptable school response rate may still be achieved through the use of replacements (identified for each sampled school)

School Response Rate

- If the school response rate of original sampled schools is below 65%, the results of the study will be in question because of the substantial risk of school non-response bias

Student Sampling

- Why?

  - Ensures that the desired sample size of students is obtained

  - Ensures that students are sampled according to standardized procedures

  - Procedure provides information required for weighting the student samples

What the National Centre needs to do?

- Make sure that schools send complete lists of all students with the specified birth dates, including those

  - students from grades 7-12
  - students from different classes
  - students in all school tracks/programs
  - students in all shifts/attendance sessions
  - students of all languages of instruction.

What the National Centre needs to do?

- Check back with the school when the number of students from the school frame differs greatly from the actual number in the school, especially when the actual number in the school is LESS than the number noted on the school frame

- Send the verified student lists to the contractors who will select the student samples and return them

Student Response Rates

- PISA requires a national student response rate, including students in any replacement schools, of 80%(weighted).

Field Trial 2016

- Currently planned to occur between September and November

- The target population birth date definition should not overlap the birth date definition to be used in the Main Survey – in most cases should be exactly twelve months older

- PISA Standards stipulate that a field trial should occur in any language of instruction used for at least 5% of the target population

Field Trial 2016

- The number of sampled schools will likely be at least 25 plus a few others in reserve to replace some of those that do not participate

Sampling in PISA

- Questions?

- Thank you

Survey Operations – Project Phases

- Project Design

- Field Trial Preparation

- Field Trial Data Collection

- Field Trial Analysis and Reporting

- Main Survey Preparation

- Main Survey Data Collection

- Main Survey Analysis and Reporting

- Next Cycle Preparation

Project Timeline

2015	2016	2017	2018

Develop assessment frameworks and test design

Develop survey materials, including assessment instruments

Carry out FT activities, including coding and data entry

Analyze FT data and prepare for MS

Carry out MS activities, including coding and data entry

Analyze MS data

Report planning and dissemination
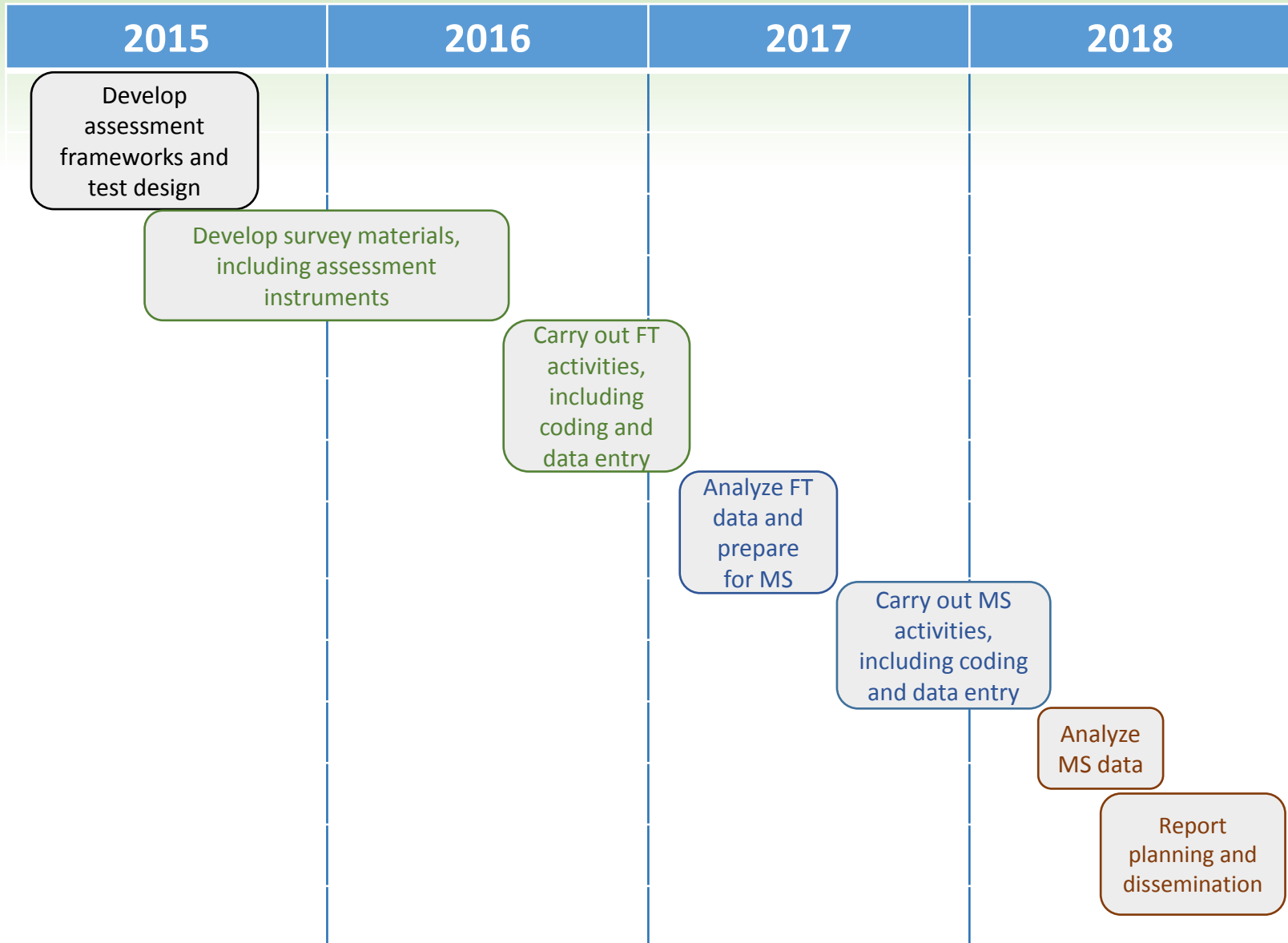
Project Design

Project Design

Project Design

Consultants (National & International)

- Psychometricians/Statisticians

- Subject Area Experts

- Test and Questionnaire Item Developers

- Translators and Translation Verifiers

- Coding Experts and Data Analysts

- Survey Operations

Project Design

Goals

- Produce Data of Highest Quality

- Cost-effective

Questions and Discussion

Field Trial

Field Trial Preparation

- Communicating

- Preparing Survey Operations Materials

- Informing/Recruiting Sampled Schools

- Hiring Field Staff

- Training Field Staff - Materials and Webinars

- Developing Quality Control Procedures

- Planning for receipt and processing of data

Field Trial

Field Trial Preparation

Communications

- Email

- Telephone/Skype

- Internet conferences (webinars)

- International Meetings

- National Centre Visits

Questions and Discussion

Field Trial

Field Trial Preparation

Survey Operations Materials

- Guides for National Project Managers

- Resource Materials

- School-level Materials

- Guides for Adapting/translating School-level Materials

- Guides for Hiring Data Collection Staff

- Printing

Field Trial

Field Trial Preparation

Survey Operations Materials

- Manuals for School Co-ordinators and School Associates

- Manuals for Test Administrators

- Test Script

- Student Tracking Form

- Session Attendance Form

- Session Report Form

Survey Operations Materials

School Materials Adaptation Process

Westat gives international English version to countries	→	Countries adapt to fit country-specific procedures (English)	→	Westat reviews and approves countries' adaptations (English)	→	Countries translate into own language(s)
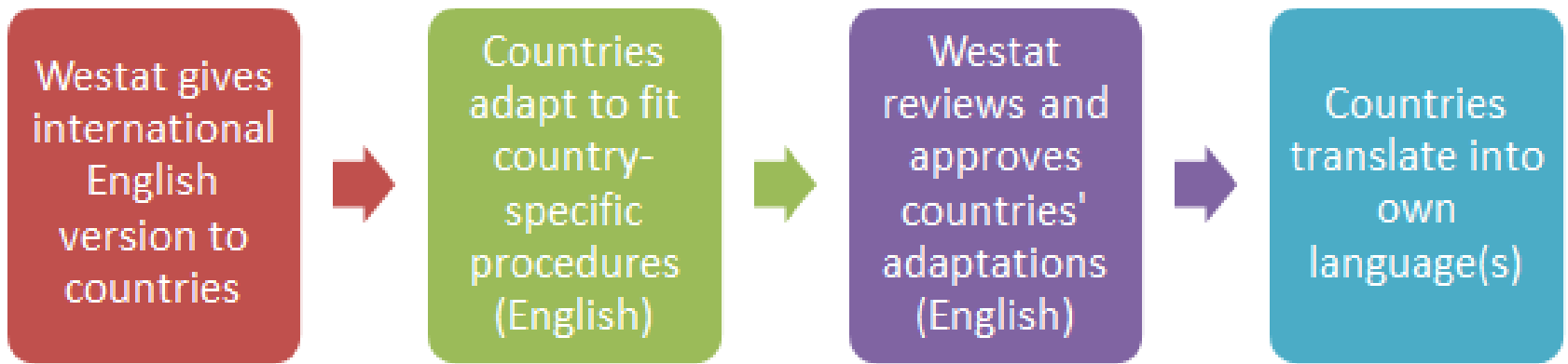
Questions and Discussion

Field Trial

Field Trial Preparation

Informing and Recruiting Schools

- **Introductory letter** (emailed or mailed)

- **Brochures and Informational Materials** (emailed, mailed or provided in-person)

- **National Centre Web Site**

- **Toll-free telephone, email, and fax numbers**

Field Trial

Field Trial Preparation

Informing/Recruiting Schools

- Countries must implement effective school refusal conversion strategies to meet PISA response rate goals

- **Incentives** (monetary, non-monetary)

Questions and Discussion

Field Trial

Field Trial Preparation

Field Staff Hiring

- Test Administrator – prepares for and administers the test and student questionnaire

- School Co-ordinator - Appointed by National Centre or the school. Co-ordinates activities between the school, Test Administrator, and National Centre

- School Associate – serves as both the Test Administrator and the School Co-ordinator

Field Trial

Field Trial Preparation

Field Staff Hiring

- recruiting,
- hiring,
- training

Questions and Discussion

Field Trial

Field Trial Preparation

Training Materials and Webinars

- Guides for National Project Managers

- Resource Materials

- Training Materials

- Webinars

- NPM mentors

Field Trial

Field Trial Preparation

Training Materials and Webinars

- Focus is on in-person training

- Countries should use experienced training staff to conduct training sessions

- Trainings should be standardized preferably with a script that all trainers follow

- Trainings should occur at least 2 weeks before the beginning of data collection

Field Trial

Field Trial Preparation

Training Materials and Webinars

Major objectives/purpose

- Standardise approach and content of staff training across countries as much as possible

- Provide staff with thorough training on PISA instruments and procedures

- Instil in staff importance of following study protocol, maintaining data confidentiality and data security

- Review the PISA Technical Standards and the importance of Quality Control

Questions and Discussion

Field Trial

Field Trial Data Collection Phase

- Communications

- Receiving and Processing Data

- Quality Control

- Begin preparing for the Main Survey

Survey Operations
Quality Control (QC)

National Quality Control

- Countries must develop and implement procedures to monitor in-field performance of test administrators to ensure that field staff are administering instruments following PISA policies

- Countries must document QC procedures and outcomes of monitoring

Survey Operations Quality Control (QC)

QC activities for survey operations are part of PISA Technical Standards

Major objectives/purpose of Quality Control

- Ensure that all PISA countries and organisations adhere to the PISA Technical Standards

- Allow for reliable and valid results for countries

- Allow for reliable and valid comparisons between countries

Survey Operations Materials

Field Trial Analysis and Reporting

Questions and Discussion

Transition to Main Survey

Main Survey

Main Survey Stages

- Communicating

- Preparing Survey Operations Materials

- Informing/Recruiting Schools

- Hiring Field Staff

- Training Field Staff - Materials and Webinars

- Developing Quality Control Procedures

- Planning for receipt and processing of data

Overview

- Coding

- Data Management

- Data Analysis

- Data Products

- Support for Report Preparations

Coding

The Importance of Coding

- Accurate and reliable coding are key components of quality control necessary for ensuring valid assessment results.

- The scales on which we build the PISA statistical framework are only as good as the codes/scores that comprise them.
  - Quality in coding cannot be recovered after the fact.

Coding Rationale

- <u>Coding</u> is required to classify responses into prescribed categories and to determine whether respondents have answered the questions correctly. All questions will need to be coded: some automatically, some by trained coders because these involve human judgment.

- <u>Recoding</u> is required as quality assurance to determine whether the coding rubrics are being applied consistently. Being able to compare results between countries requires the consistent application of coding rules both within and across countries.

Goals of the Coding Process

- To code each response at least once
- To recode some of the responses
  - Determine whether coders are applying the coding rubrics consistently
- To ensure that coders within countries are applying consistent coding rules
  - Established by recoding a set of responses by two independent coders (with careful monitoring)
- To ensure that coders across countries are applying consistent coding rules
  - Established by recoding the same set of common responses by coders from all countries

Coding Requirements

- Assemble a coding team:
  - Three separate coding teams for cognitive assessment (one for each domain—reading, mathematics, and science)
    - Relatively high skills (i.e., more than a secondary qualification)
    - Good understanding of mid-secondary level studies in the relevant domain
    - Available for the duration of the coding period
  - Each domain-specific team will have at least 3 coders, from which one is the Lead Coder
- Additional coders for background questionnaires
- Data entry support staff

Coding Timeline

- FT coding procedures and item-by-item coding guidelines (Lead Coders required) (Fourth NPM Meeting, July 2016)

- FT National coding training activities (near the end of the FT data collection)

- FT coding and manual data entry (December 1 2016—January 31, 2017)

- MS coding procedures and focused item-by-item coding guidelines (Lead Coders required) (Sixth NPM Meeting, July 2017)

- MS National coding training activities (near the end of the MS data collection)

- MS coding and manual data entry (December 1 2017—January 31, 2018)

Data Management, Analysis, and Products

Data Management

- Contractor responsibilities
- National tasks and roles
- Resources and materials
- Timeline

Data Management

- Contractor (ETS) responsibilities
  - Provide software, codebooks, manuals, training, and support to countries for national database building
  - Manages, processes, and cleans data on the international level
  - Prepares analysis and dissemination databases

Data Management

- National tasks and roles
  - National Project Manager (NPM)
    - Supervises all data related work
    - Responsible for ensuring that all tasks relating to the production of a quality national database are carried out on schedule and in accordance with specified international standards

  - National Data Manager (NDM)
    - Responsible for the day-to-day data management tasks
    - Possesses technical skills with respect to databases and the conversion between different file formats
    - Experienced in planning, conducting, and supervising data management in large-scale surveys
    - Main contact person to ETS data analysis team

Data Management

- Key resources and materials
  - Data Management Manual
    - Single most important resource for information regarding the upcoming data tasks
  - Data Management Expert (DME) software
    - The most important application for the within-country work
  - International master codebook template
    - Includes all variable definitions as used in the international English source version of any instrument and required file

Data Management Timeline

- FT data management training (NDM required) (Fourth NPM Meeting, July 2016)

- FT data processing, manual data entry, and data submission (eight weeks following the end of data collection)

- FT international data cleaning and processing

FT Data Analysis

- Uses of Field Trial data by national centres in order to review results and prepare Main Study instruments

  - Examine data yield (quantity)

  - Examine data quality and comparability

  - Use of classical item analysis and IRT model-based output to identify item-specific inconsistencies and/or serious errors (i.e., translation or printing problems)

FT Data Products

- National Database
  - Background and assessment data
  - SAS and SPSS formatted files
- Summary Data Tables (cognitive and BQ)

Data Management Timeline

- FT data management training (NDM required) (Fourth NPM Meeting, July 2016)

- FT data processing, manual data entry, and data submission (eight weeks following the end of data collection)

- FT international data cleaning and processing

- MS data management training (NDM required) (Sixth NPM Meeting, July 2017)

- MS data processing, manual data entry, and data submission (eight weeks following the end of data collection)

- MS international data processing and editing

MS Data Analysis

- Uses of Main Study data by national centres in order to review results and prepare international database

  - Examine data yield (quantity)

  - Examine data quality and comparability

  - Scaling

  - Preparation of international database and products

Final Data Products

- International Database
  - Background and assessment data
  - SAS and SPSS formatted files
- Data Analysis Tools
  - PISA-D Data Explorer (Web based)
  - Micro-Data Analyzer for SPSS
  - SAS Macros
- Summary Data Tables (cognitive and BQ)
- Compendia
- Editable International Report Tables
- Technical Documentation

Data Tools

- **Analysis Possibilities**
  - Percentages and means
  - Percentiles
  - Correlations
  - Linear and logistic regression
  - Benchmarks
- **Output Options**
  - Text files
  - SAS or SPSS datasets
  - MS Word or Excel
  - HTML

Support for Report Preparation

- Training
  - Survey methods and procedures
  - Data analysis tools

- Summary output
  - International report tables
  - Country-specific report tables

Questions and Discussion

PISA FOR DEVELOPMENT

International Advisory Group Meeting – Next Steps
Closing Session: Day Three

13 March 2015

OECD Secretariat

OECD
BETTER POLICIES FOR BETTER LIVES

- OECD Secretariat with International Contractors to facilitate translation of key materials, such as manuals, into French and Spanish

- Countries to nominate at least two subject experts from at least two of the domains by 31 March 2015

Contextual framework

The Learning Bar to facilitate:

- Countries to identify policy questions (additional to existing PISA) that countries need answered – gaps – by June 2015
- Identify questions (or measures) to include in the PISA-D contextual questionnaires for each country by July 2015
- TLB presentations to QEG August 2015
- Finalise contextual materials for November 2015

OECD and countries to work together to strengthen NPM networks among participating countries in PISA and PISA-D including NPM mentoring, continue to explore opportunities for further peer-to-peer learning in IAG and NPM meetings

Project Implementation Schedule

	2014	2015	2016	2017	2018
	J F M A M J J A S O N D	J F M A M J J A S O N D	J F M A M J J A S O N D	J F M A M J J A S O N D	J F M A M J J A S O N D

PHASE I: DESIGN, PLANNING AND CO-ORDINATION
- IAG meetings
- TAG meetings
- Expert papers
- Workshops
- CNA
- CBP design
- Project implementation plans
- Tender for IC (A, B, C)
- IC contract

PHASE II: TECHNICAL DEVELOPMENT, INCLUDING REVIEW OF INSTRUMENTS, DESIGN OF ENHANCEMENTS, PREPARATION OF MATERIALS AND PLANNING FOR FIELD TRIALS
- 1st review*
- Preparation of materials*
- OOSC preparation
- Field trial preparation
- Country report focus
- International meeting (peer-to-peer)

PHASE III: FIELD TRIALS AND IN-COUNTRY DATA COLLECTION
- Field trials*
- 2nd review and analysis*
- M.S. preparation*
- Main study*
- OOSC field trials
- OOSC review and preparation
- OOSC survey
- International meeting (peer-to-peer)

PHASE IV: ANALYSIS AND REPORTING
- Data cleaning and analysis
- Review
- Country report writing
- International meeting (peer-to-peer)

PHASE V: POST-PILOT GOVERNANCE
- Instruments final
- OOSC final
- Country reports final
- Independent review
- Results report
- PISA GB
- International seminar

*contextual questionaires and cognitvie instruments for students in school only

International Advisory Group

- IAG 2015 March   Paris, France   Europe

- IAG 2016 March   Asunción, Paraguay LAC

- IAG 2017 March   Zambia?   Asia?

- IAG 2018 March   Saly, Senegal   Africa

Capacity Building Meetings

- NPM#1: 28 September 2015, Quito, Ecuador

- NPM#2: January 2016, Washington D.C, US

- NPM#3: April 2016, Asunción, Paraguay

- NPM#4: July 2016, Zambia

- NPM#5: May 2017, Saly, Senegal or Princeton/Washington D.C., US

- NPM#6: July 2017, Antigua, Guatemala

- NPM#7: March 2018, Saly, Senegal

- NPM#8: July 2018, Antigua, Guatemala