

PISA for Development

Expected results from the meeting



*4th meeting of the International Advisory Group
Sokha Angkor Resort – Siem Reap, Cambodia
17-19 May 2017*

Expected results from the meeting

- Election of two Co-Chairs (representing participating countries and development partners)
- Review and discussion of annual report on project activities, including the contribution of the OECD and the project to the SDG Education 2030 agenda
- Review and discussion of detailed progress report for Strands A, B and C, including progress with the field trials, analysis of the data from the field trials, sampling design and detailed plans for 2017/2018, particularly the main study
- Review and discussion of the outcomes from the field trial of the cognitive tests (Strand A)
- Review and discussion of the outcomes from the field trial of the questionnaires (Strand B)

Expected results from the meeting (cont'd)

- Review and discussion of progress with capacity-building activities and outcomes
- Review and discussion of the PISA for Development Assessment and Analytical Framework
- Review and discussion of the Analysis and Reporting Plan, including the structure and content for the national reports and the table shells to be used
- Peer-to-peer learning with Brazil, Kosovo and GIZ regarding communicating with key stakeholders; analysing national data and preparing a national report; and disseminating results

Expected results from the meeting (cont'd)

- Review and discussion of progress with the PISA for Development engagement and communication strategy – participating countries reports on stakeholder engagement and communication activities
- Presentation by OECD of the PISA 2015 results (optional session during lunch break)
- Agreement of the ToR for an independent review of the project to be conducted in 2018-19
- Discussion of regional assessment experiences with NEQMAP and SEA-PLM
- Draft summary record and next steps agreed

Day One sessions

1. Introductions, Purpose of Meeting, Election of Co-Chairs, Project Progress Report and the contribution of the OECD and the project to education in the post-2015 agenda
2. Presentation by the Contractors of detailed progress report for Strands A, B and C, including the field trials (Strands A, B and C), instruments and technical standards (Strand C), sampling design and detailed plans for 2017-18
3. Presentation by ETS of the outcomes from the field trial of the cognitive tests (Strand A)
4. Presentation by TLB of the outcomes from the field trial of the questionnaires (Strand B)

Day Two sessions

5. Participating countries report on capacity-building activities and outcomes
6. Presentation by OECD of the PISA for Development Assessment and Analytical Framework
7. Presentation by OECD and TLB of analysis and reporting plan, including the structure and content for the country reports

Day Three sessions

8. Peer-to-peer learning with Brazil and Kosovo regarding communicating with key stakeholders, analysis of national PISA data, preparing a national report, and disseminating results
9. PISA for Development engagement and communication strategy – participating countries to report on stakeholder engagement and communication activities

Optional session: Lunch break with presentation by OECD of PISA 2015 results

Day Three sessions (cont'd)

10. Agreement of the ToR for an independent review of the project to be conducted in 2019
11. Presentation by UNESCO Bangkok on Network on Education Quality Monitoring in Asia and the Pacific (NEQMAP), and presentation by UNICEF East Asia and Pacific Regional Office (EAPRO) and the Southeast Asian Ministers of Education Organization (SEAMEO) on Southeast Asia Primary Learning Metrics (SEA-PLM)
12. Next steps and meeting conclusion

Key documents *sent prior to meeting

1. Agenda*
2. Participants list
3. Annual progress report & summary record of last meeting*
4. Strand A: Progress Report and Field Trials
5. Strand B: Progress Report and Field Trials
6. Strand C: Progress Report, Field Trials and Instruments and Technical Standards
7. Outcomes of the field trial of the cognitive tests*
8. Outcomes of the field trial of the questionnaires*
9. PISA-Development Assessment and Analytical Framework*
10. Analysis and Reporting Plan + table shells*
11. Terms of Reference for an independent review of the project*

Other documents in the room

- PISA-D Engagement and Communications Strategy
- Examples of national reports from PISA 2015:
 - Education in China: A Snapshot
 - Brazil in PISA 2015: Executive Summary
 - PISA 2015: Les défis du système éducatif français et les bonnes pratiques internationales
- PISA-D commissioned reports:
 - The Experience of Middle-Income Countries Participating in PISA 2000-15
 - A Review of International Large-Scale Assessments in Education
 - Making Education Count for Development: Data collection and availability in six PISA for Development countries
- PISA Partnership Options brochure

IAG Agenda

- The International Advisory Group is invited to adopt the agenda for its fourth meeting

PISA-D Annual Report on project activities



*4th meeting of the International Advisory Group
Sohka Angkor Resort – Siem Reap, Cambodia
17-19 May 2017*

Summary of achievements since last IAG

- The project plan for 2016 has been implemented in accordance with the 15 next steps agreed upon at the 3rd IAG meeting
- Plans for Strands A and B are on track: major technical milestones have been achieved – field trial has taken place and preparation for the main study is well under way
- Plans for Strand C have been slightly adjusted due to delays with the delivery of the tablets to the National Centres, but major technical milestones have been achieved – tests, questionnaires, manual, standards, sampling plan, adaptation and translation guidelines are all in place, and the field trial is under way
- Capacity building and peer learning activities are continuing

Reasons for success

- The commitment and performance of participating countries and their national teams has been outstanding.
- The international contractors have done a great job and delivered on their terms of reference in exactly the right way.
- Development partners have delivered their support (international and in-country) and technical guidance on time and keep us moving forward.
- Technical partners, especially UNESCO and UNICEF, have provided guidance and co-operation and facilitated the project's contribution to wider processes; i.e., Education SDG agenda.
- PISA participating countries (Brazil, Canada, Chile, Colombia, Kosovo, Korea, Mexico, Peru, Spain, Uruguay, United States) have been generous and valuable contributors to peer-to-peer learning

Annual report format

From 2015 onwards, the OECD submits a progress report to the IAG including the following:

- actual outputs compared to planned outputs
- summary of the use of funds compared to budget,
- explanation of major deviations from plans,
- assessment of problems and risks,
- assessment of the need for adjustments to activity plans and/or inputs and outputs, including actions for risk mitigation,
- assessment of achievements in relation to project purpose

15 next steps agreed upon at 3rd IAG

1. OECD to sign contract with international contractor (ETS) for Strand C as soon as possible
2. OECD to complete the report on system level data with UIS following receipt of feedback from the countries and to make this available on its website
3. OECD to sign participation agreement with Panama
4. OECD to complete the Capacity Needs Analysis and the design of Capacity Building Plans for Honduras and Panama
5. Honduras and Panama to finalise Project Implementation Plans
6. Participating countries to finalise outstanding agreements with development partners regarding contributions and support (e.g., international costs, in-country costs and activities)
7. OECD to further develop and confirm peer-to-peer learning partnerships and implement the planned capacity development activities with the countries

15 next steps agreed upon at 3rd IAG

8. OECD to finalise with development partners outstanding agreements for support to the project – general contributions and country-specific contributions
9. Third meeting of Technical Advisory Group on 6 June 2016
10. Project implementation: Strands A and B – Field Trial
 - 3rd international/NPM meeting (4-8 April 2016); 4th international/NPM meeting (4-8 July 2016)
11. Project implementation: Strand C – Technical Development and Field Trial
 - 4-a international/NPM meeting (1-4 November 2016)
12. IAG to send Secretariat comments on analysis and reporting plan by 30 April 2016; Secretariat to discuss second draft of plan with TAG on 6 June 2016; third draft for IAG review in Autumn 2016; final draft for IAG meeting in May 2017
13. OECD Secretariat with International Contractors to complete translation of key materials, such as manuals and frameworks, into French and Spanish
14. OECD to implement with participating countries the engagement and communication strategy for the project and country strategies
15. Fourth meeting of IAG from 17-19 May 2017 in Cambodia.

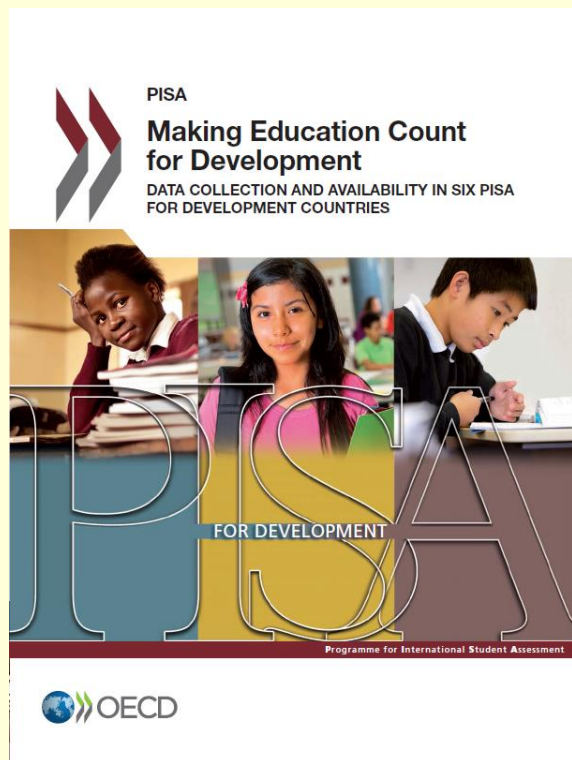
i) OECD to sign contract with international contractor (ETS) for Strand C as soon as possible

OECD signed Strand C contract with international contractor ETS on 4th May 2016

- Preparations for Strand C have been underway since late 2015, and work formally began with the signing of the contract
- 1st monthly PISA-D Strand C newsletter issued by ETS on 29 April 2016

ii) OECD to complete the report on system level data with UIS following receipt of feedback from the countries and to make this available on its website

- Published on 26 Sept 2016, available online
- Reviews the availability and quality of system-level data and metadata on education in 6 PISA-D countries: CAM, ECU, GUA, PAR, SEN, ZAM
- Identifies country-specific issues and provides options to address the challenges, i.e. improving data quality, completeness and international comparability
- HON joined PISA-D too late to be included in the UIS report, so OECD is collecting their system-level data; PAN data is collected through PISA 2018



iii) OECD to sign participation agreement with Panama

- PISA-D has been launched in Panama
- Agreement signed on 22nd April 2016



iv) OECD to complete CNAs and CBPs for Honduras and Panama

Honduras: CNA published Sept 2016; CBP Nov 2016

Panama: CNA published Apr 2017; CBP August 2017

- The CNAs reveal solid foundational capacity for implementing the project in both countries, particularly technical capacity and knowhow
- There are some gaps, which are being addressed through the CBPs, but most core capacities required for successful implementation are already in place

v) Honduras & Panama to finalise PIPs

The PIPs for **Honduras** and **Panama** were finalised shortly after the countries' CBPs were put in place

- Describes actions to be carried out by the specific entities and agents that are named and commissioned for implementation by the authorities of the country
- Also includes policy priorities and expectations, communication strategy, risk management strategy, evaluation and monitoring strategy, and a fully costed project budget

vi) Countries to finalise outstanding agreements with development partners

- All 9 countries have finalised or are in the process of finalising their respective budgets for financing their participation in PISA-D
 - Countries' contributions to the international costs of participation (budget managed by the OECD)
 - In-country costs of participation (budget managed by the countries themselves and, in some cases, their development partners), which is described in the PIPs

vii) OECD to further develop and confirm peer-to-peer learning partnerships and implement the planned capacity development activities



vii) OECD to further develop and confirm peer-to-peer learning partnerships and implement the planned capacity development activities (cont'd)



Pedro Sanoja, Diana Muñoz and Luis Saldivia at an ETS-facilitated PISA workshop last month in Ecuador.



viii) OECD to finalise with development partners outstanding agreements for support to the project

- All outstanding agreements for support to the project – general contributions and country-specific contributions – have been finalised
- Fund-raising is an on-going process
- Since the 2015-16 progress report, new development partners include:



ix) Third meeting of Technical Advisory Group

3rd meeting of Technical Advisory Group (TAG) took place on 6th June 2016 in Princeton, United States

- Test design, timeline and field trial
- PISA-D technical standards: Strands A, B and C
- Data analysis and reporting plan
- Factors affecting data quality
- Comparing PISA-D results to PISA

4th TAG meeting (virtual) took place on 9th May 2017:
results of field trial and design of the main data collection

x) Project implementation: Strands A and B – Field Trial

3rd international/NPM meeting, Asuncion (4-8 April 2016)

- Sampling plan for the school assessment
- Arrangements for data management and survey operations



4th international/NPM meeting, Livingstone (4-8 July 2016)

- Coding reading, mathematics, science and occupations
- Survey operations for the field trial
- Data management training
- ICT training for the out-of-school assessment



x) Project implementation: Strands A and B – Field Trial (continued)

Field Trial of school-based instruments completed successfully from Aug-Dec 2016 in 7 countries



- ECU, GUA, HON & PAR delivered data sets to the contractors from Jan-Mar 2017: data has been cleaned, verified & analysed
- CAM, SEN & ZAM delivered data sets to the contractors in Apr 2017
- All efforts have been made by the countries, the contractors & the OECD to incorporate analysis of the Field Trial data from all countries into the design of the instruments for the main study

xi) Project implementation:

Strand C – Technical Development and Field Trial

4-a international/NPM meeting, Madrid (1-4 November 2016)

- Interviewer training, using interactive role plays with the tablets and interview software
- Survey operations for the field trial, including sampling updates and the tablet-based delivery system
- Data management training



Countries also finalised their out-of-school instruments with the contractors in readiness to load these on the tablets, once delivered

- Delivery occurred with slight delays
- Customs clearance took much longer than expected (2-3 months)
- Strand C is therefore running behind the original schedule: countries will go to the field 3 months later than planned (but project will still end Dec 2019)

xi) Project implementation: Strand C – Technical Development and Field Trial



- 1st Field Trial of out-of-school population instruments launched in PAR in April 2017
- Next up GUA, followed by HON, PAN, SEN & ZAM
- All countries will be out of the field by the end of August 2017

xii) IAG to send Secretariat comments on analysis and reporting plan by April 2016; Secretariat to discuss 2nd draft of plan with TAG in June 2016; 3rd draft for IAG review in Autumn 2016; final draft for IAG meeting in May 2017

- IAG feedback was incorporated by the Secretariat in a 2nd draft
- TAG discussed 2nd draft and provided feedback
- Secretariat incorporated TAG feedback along with feedback from consultations with the contractors from Apr-Nov 2016
- Secretariat created 136 proposed table shells for reporting results based on the draft outline of the national reports included as an annex to the Analysis and Reporting Plan
- 3rd draft + table shells were issued to IAG in Dec 2016
- IAG feedback plus analysis of Field Trial data informed the final draft of the Plan + revised table shells discussed at IAG May 2017

xiii) OECD Secretariat with International Contractors to complete translation of key materials, such as manuals and frameworks, into French and Spanish

Materials translated into French and Spanish include:

- PISA-D NPM Manual and PISA-D Technical Standards together with a selection of other key documents, including the frameworks
- New PISA-D brochure issued in Nov 2016
- The entire OECD PISA-D website in Mar 2017

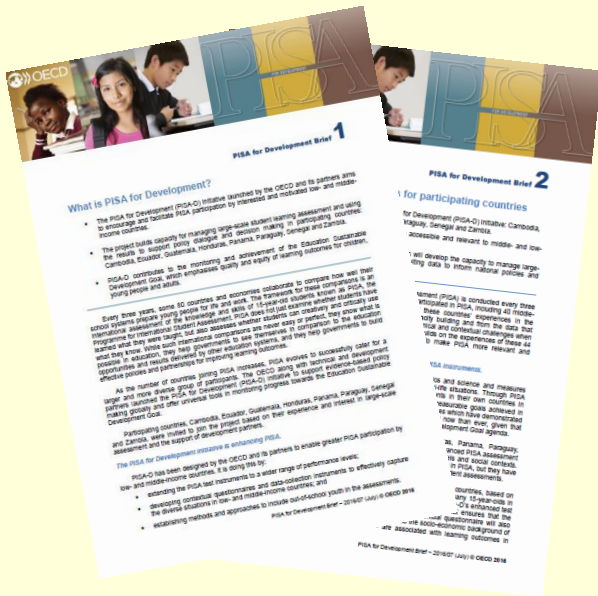
OECD is seeking funds to translate the PISA-D Assessment and Analytical Framework, a key tool for communicating with stakeholders

xiv) OECD to implement with participating countries the engagement and communication strategy for the project and country strategies

May 2015: Final PISA-D Engagement and Communication Strategy endorsed by IAG

- All NCs have linked a communications focal point with the OECD

July 2016:



xiv) OECD to implement with participating countries the engagement and communication strategy for the project and country strategies

Aug 2016: OECD reviewed countries' online PISA-D communication & provided tailored feedback with suggestions to strengthen tools

Sep-Dec 2016: OECD provided communications support to countries leading up to Field Trial launches, e.g. suggesting Tweets/Facebook posts to announce the launch, responding to country inquiries regarding their websites, brochures, etc.



xiv) OECD to implement with participating countries the engagement and communication strategy for the project and country strategies

Jan 2017: OECD published the PISA Partnership Options brochure, which explains that the OECD now offers PISA-D-style support to countries participating in PISA



- PISA-D instruments and capacity-building activities have been incorporated within the Terms of Reference for the PISA 2021 Call for Tender
- OECD's outreach efforts to low- and middle-income countries beyond those participating in PISA-D have been reinforced by the inclusion of PISA in the Sustainable Development Goals (SDGs) global indicators framework that will be used to monitor progress towards the Education SDG by 2030

xv) 4th meeting of IAG 17-19 May 2017 in Cambodia



PISA for Development use of funds

Expenditure Analysis (EUR)	Expenditures			Commitments			Total 6 years
	2014 Y1	2015 Y2	2016 Y3	2017 Y4	2018 Y5	2019 Y6	
A1. International Advisory Group	162 406	62 473	54 887	51 000	45 000		375 766
A2. Instrument development, field trials, local assessment implementation and related services	56 638	810 460	2 123 286	2 299 238	951 750	713 269	6 949 371
A3. Technical oversight, co-ordination, analysis and reporting	344 454	391 106	480 769	440 000	440 000	215 000	2 311 329
A4. Engagement for peer-to-peer learning and contribution to UN-led post-2015 process	17 070	91 236	82 144	40 000	40 000	40 000	310 450
Administrative cost recovery	147 715	88 236	93 775	95 000	80 000	75 000	579 726
Total	728 283	1 443 511	2 834 861	2 925 238	1 551 480	1 043 269	10 526 642

Challenges, risks and assumptions

1. The institutional, technical and administrative capacity and conditions that are needed to successfully implement PISA at country and sub-national levels – **risks low**
2. The technical challenges are many and various and relate to the central question of how far PISA for Development can go in enhancing and adapting PISA to be more relevant to developing countries while still ensuring that the results of the assessment contribute to the establishment of an international benchmark in the context of PISA – **risks low**

Challenges, risks and assumptions (cont'd)

3. PISA is changing rapidly and PISA-D needs to ensure that it provides an effective bridge for participating countries into main PISA – **risks low**
4. In all countries, PISA-D is already helping to strengthen institutions devoted to the production of reliable data and evidence for informing policy design. As the project moves forward, it is essential to maintain the teams, sustain the support for the project, institutionalise a culture of evaluation and draw support for PISA in the country across the entire political spectrum – **risks are manageable**

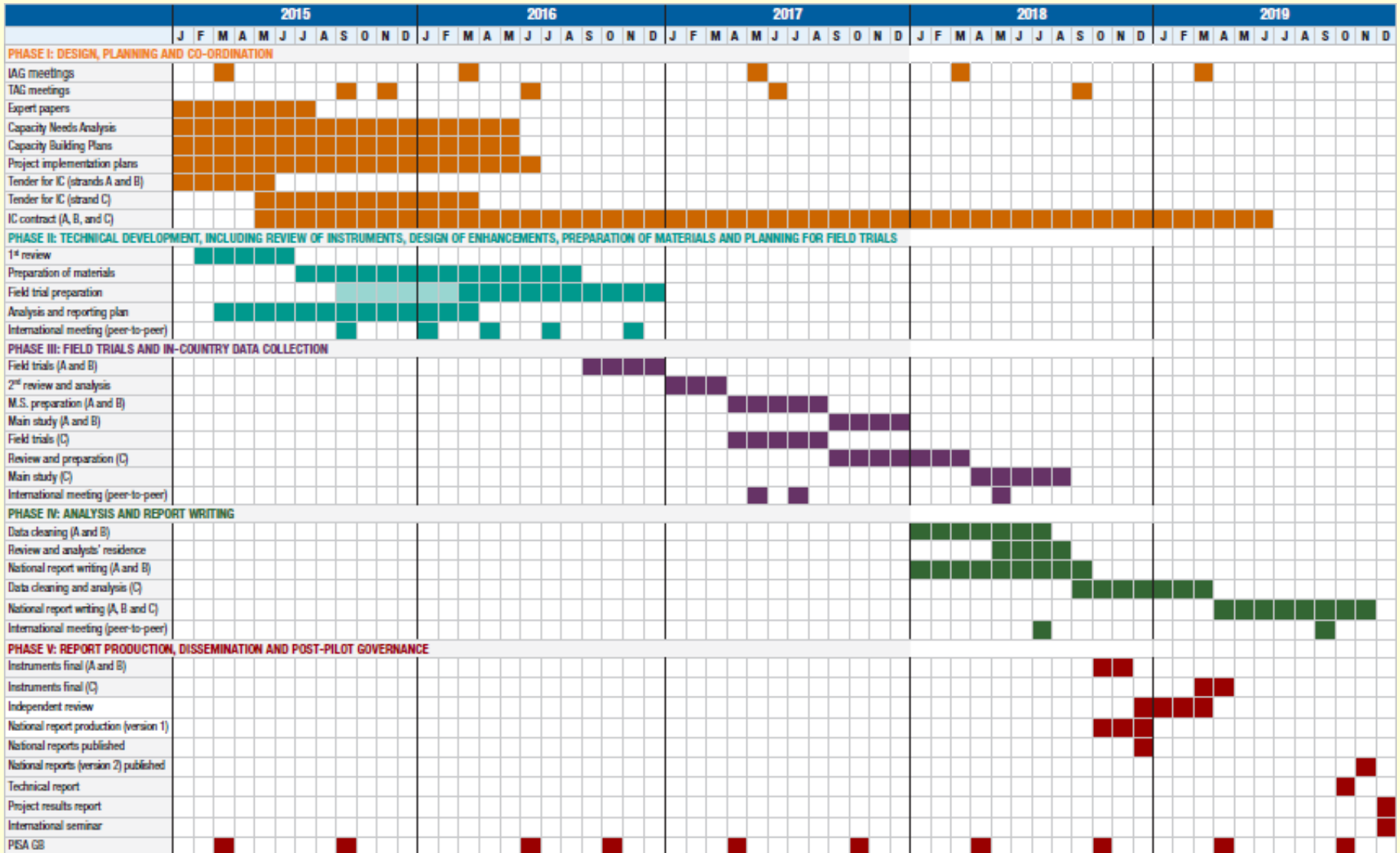
OECD and Education SDG

OECD a member of the
Education SDG Steering
Committee and architecture

PISA and PIAAC included
In SDG global monitoring
Framework from 2017
onwards



Adjustments: Project Implementation Schedule



Annual Report on Progress

The International Advisory Group is invited to:

- Note the progress that has been made since March 2016
- Discuss any issues arising from the progress report

PISA

for Development



PISA for Development Progress Report

PISA for Development IAG Meeting
17 May 2017
Siem Reap, Cambodia

Ann Kennedy, ETS



cApStAn

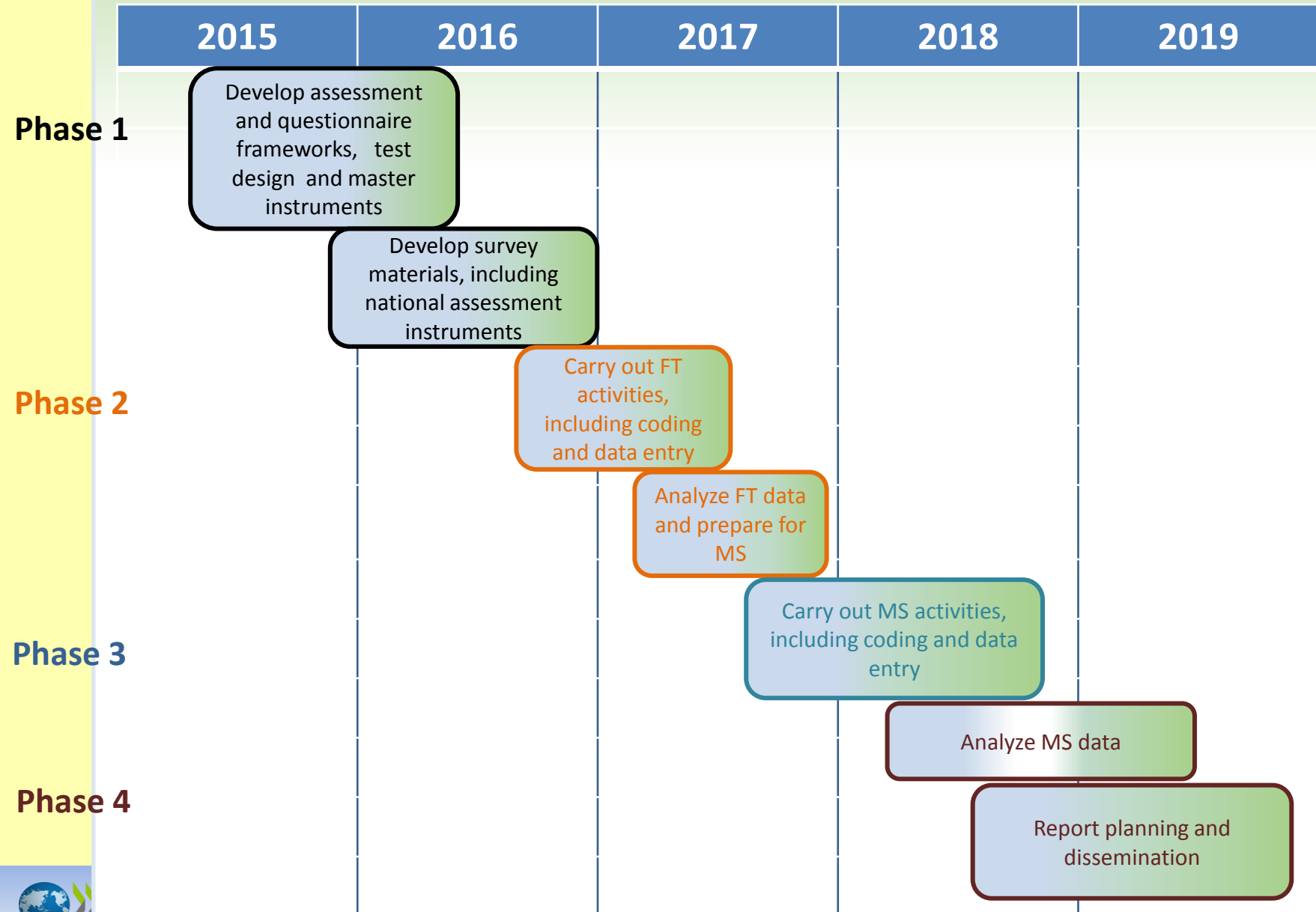


PEARSON

Overview of Progress Update

- Key project phases
- Timelines
- Key activities and milestones (March 2016-May 2017)
- Key activities and milestones through the next IAG meeting in 2018
- Technical Standards

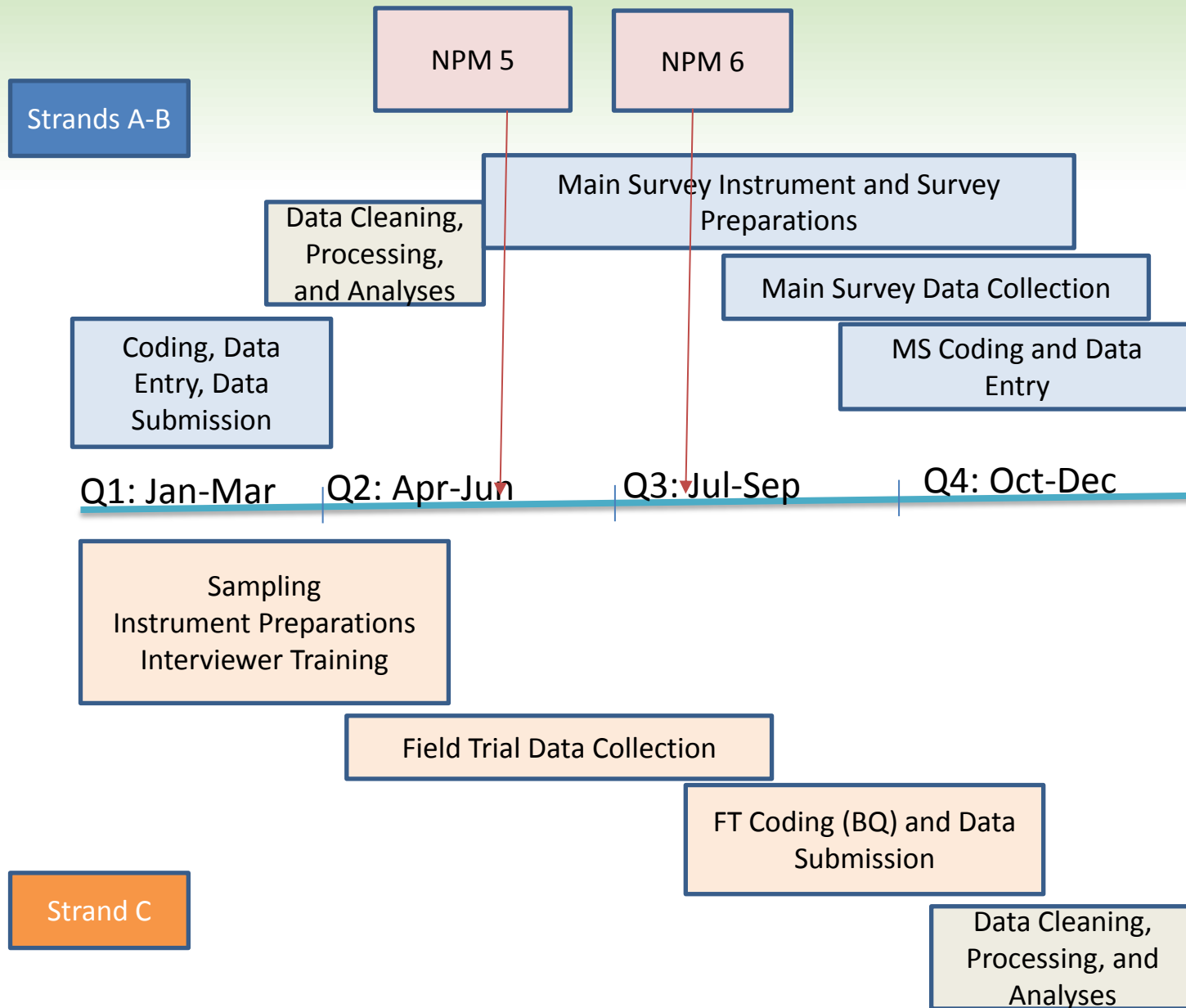
Timeline Summary



NPM Meetings

	Description	Date
4	A+B: FT survey operations, coding (BQ, cognitive) and data management training C: ITC training	4-8 Jul 2016
4a	Strand C only: Interviewer and data management training	1-4 Nov 2016
5	A+B: analysis and interpretation of FT results; preparation for MS, including sampling and proposed MS instruments C: Update on FT data collection	22-25 May 2017
6	A+B: MS survey operations, student sampling, test administrator training, coding and data management C: FT update, FT data management, and MS sampling workshop	24-28 Jul 2017
7	A+B: Data processing procedures and results, scaling methodology, and preparations for data analysis, including training C: MS Quality control and interviewer training	7-11 May 2018
8	A+B: Preparations for reporting and dissemination of country reports, including training C: Update on status of MS data collection	23-27 Jul 2018
8a	Strand C only: Reporting and dissemination of country reports	16-19 Sep 2019

Timeline: 2017



Key Activities and Milestones

March 2016 – May 2017

Strands A-B

Key Activities 2016-2017: Phase 1

- Completed FT sampling tasks
- Finalised translation/adaptations and assembled cognitive test booklets and questionnaires
- Worked with schools
- Recruited and trained test administrators
- Planned staffing for coding and data entry
- Updated data codebooks

Strands A-B

Key Activities 2016-2017: Phase 2

- During FT Data Collection
 - Dispatched test and school-level materials
 - Monitored data collection
 - Arranged follow-up sessions
 - Organised materials from schools
- Following FT Data Collection
 - Completed coding and data entry
 - Submitted sampling checks
 - Submitted database and documentation
 - Archived FT materials

Strand B (The Learning Bar Inc.)

Key Activities 2016-2017: Phases 1 & 2

- **Completed Student, Teacher and School Questionnaires**
 - QAS Files
 - Coordinated translation with cApStAn
 - Developed rotation schedule
 - Worked with countries in adaptation process
 - Final Optical Check (FOC) of each county's questionnaires
 - Provided syntax for data management of all variables
 - Provided syntax for harmonization of variables

Strand B

Key Activities 2016-2017: Phases 1 & 2

- **Developed Analysis and Reporting Plan**
 - Design of national report
 - Draft of table shells

- **Developed Student and Teacher Tracking Forms in collaboration with WESTAT**

- **Prepared and delivered ISCO training session**
 - Assessed typical coding errors in PISA 2015 for other middle-income countries
 - Prepared document on theoretical foundations for ISCO
 - Developed training files with examples
 - Delivered workshop in Zambia
 - Provided guidelines on coding by contractors, in collaboration with the OECD

Strand C

Key Activities 2016-2017: Phase 1

- Completed FT sampling tasks
- Finalised reviews of cognitive test units (adapted from Strand A) and translation/adaptation of new questionnaire content
- Configured tablets
- Reviewed and tested the computer delivery system
- Recruited and trained interviewers

Strand C (The Learning Bar Inc.)

Key Activities 2016-2017: Phase 1

- **Developed Student Interview Schedule, Parent Questionnaire, and Household Observation Schedule**

Finalized instruments in collaboration with countries and the QEG

Coordinated translation with cApStAn

Worked with countries in adaptation process

Developed interview filtering scenarios

Finalized adaptations in collaboration with countries and cApStAn

Contributed to 'dress rehearsal' for interviewer training

Strand C

Key Activities 2016-2017: Phase 2

Begin/continue data collection

Key Activities and Milestones

May 2017 – May 2018

Strands A-B

Key Activities 2017-2018: Phase 2

- Review Field Trial data and data analysis reports
- Review operations procedures
- Begin Main Survey school sampling work
- Prepare Main Survey cognitive instruments and questionnaires
- Prepare Main Survey school materials
- Identify Main Survey testing period
- Plan staffing and resources for coding and data entry

Strands A-B

Key Activities 2017-2018: Phase 3

- During MS Data Collection
 - Dispatch test and school-level materials
 - Monitor data collection (including PQMs)
 - Arrange follow-up sessions
 - Organise materials from schools
- Following MS Data Collection
 - Complete coding and data entry
 - Submit sampling checks
 - Submit database and documentation
 - Archive MS materials

Strands A-B

Key Activities 2017-2018: Phase 4

- Review Main Survey data and analysis reports
- Prepare draft table shells

Strand C

Key Activities 2017-2018: Phase 2

- During FT Data Collection
 - Monitored data collection through case management reports
 - Manage completed case files from interviewer tablets
- Following FT Data Collection
 - Upload data files to the DME database
 - Submit sampling checks and final reports
 - Submit database and documentation

Strand C

Key Activities 2017-2018: Phase 2

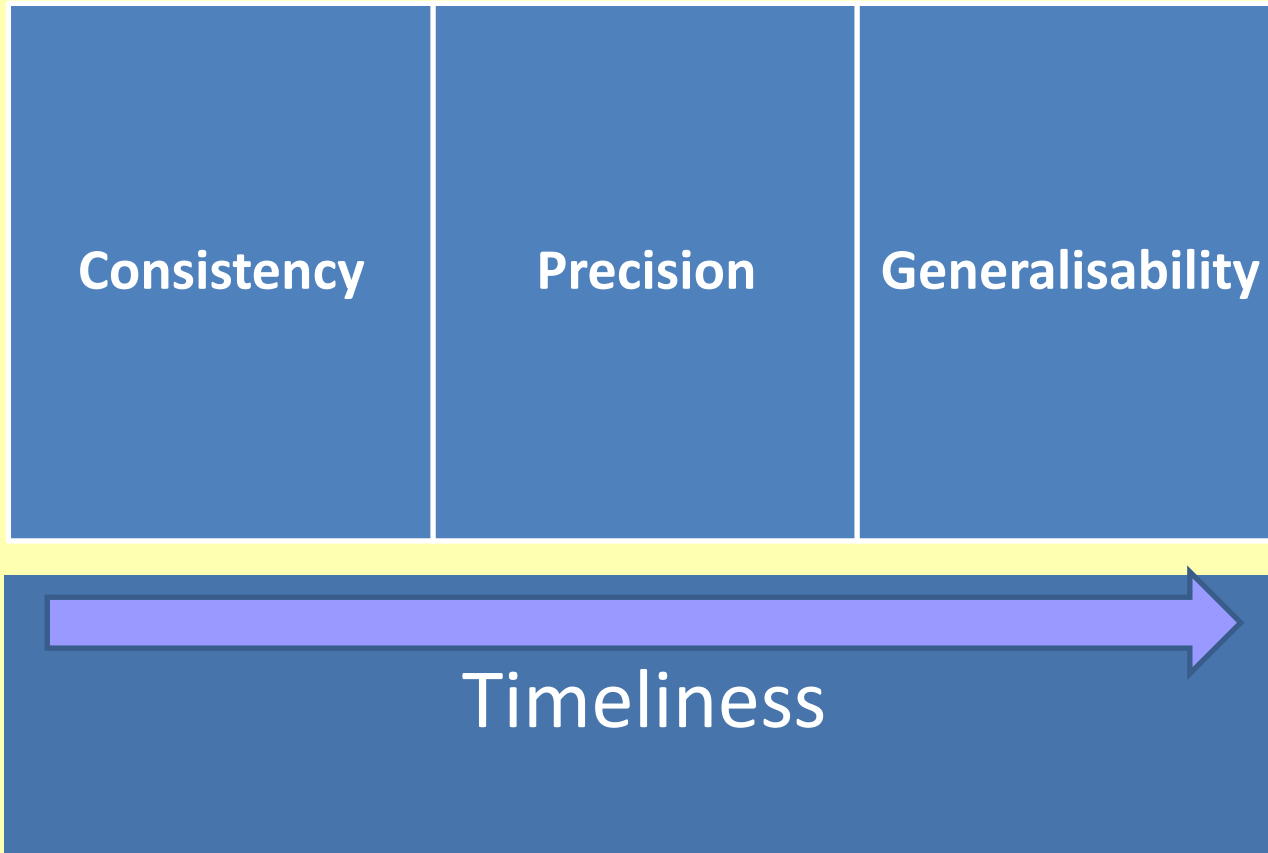
- Review Field Trial data and data analysis reports
- Review operations procedures
- Begin Main Survey school sampling work
- Prepare Main Survey cognitive instruments and questionnaires
- Prepare Main Survey interviewer training materials
- Identify Main Survey testing period
- Plan staffing and resources for coding and data management

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



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 *
- 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

PISA

for Development



Field Trial Analysis Cognitive Instruments

PISA for Development
International Advisory Group
17-19 May 2017

Claudia Tamassia, ETS



PEARSON

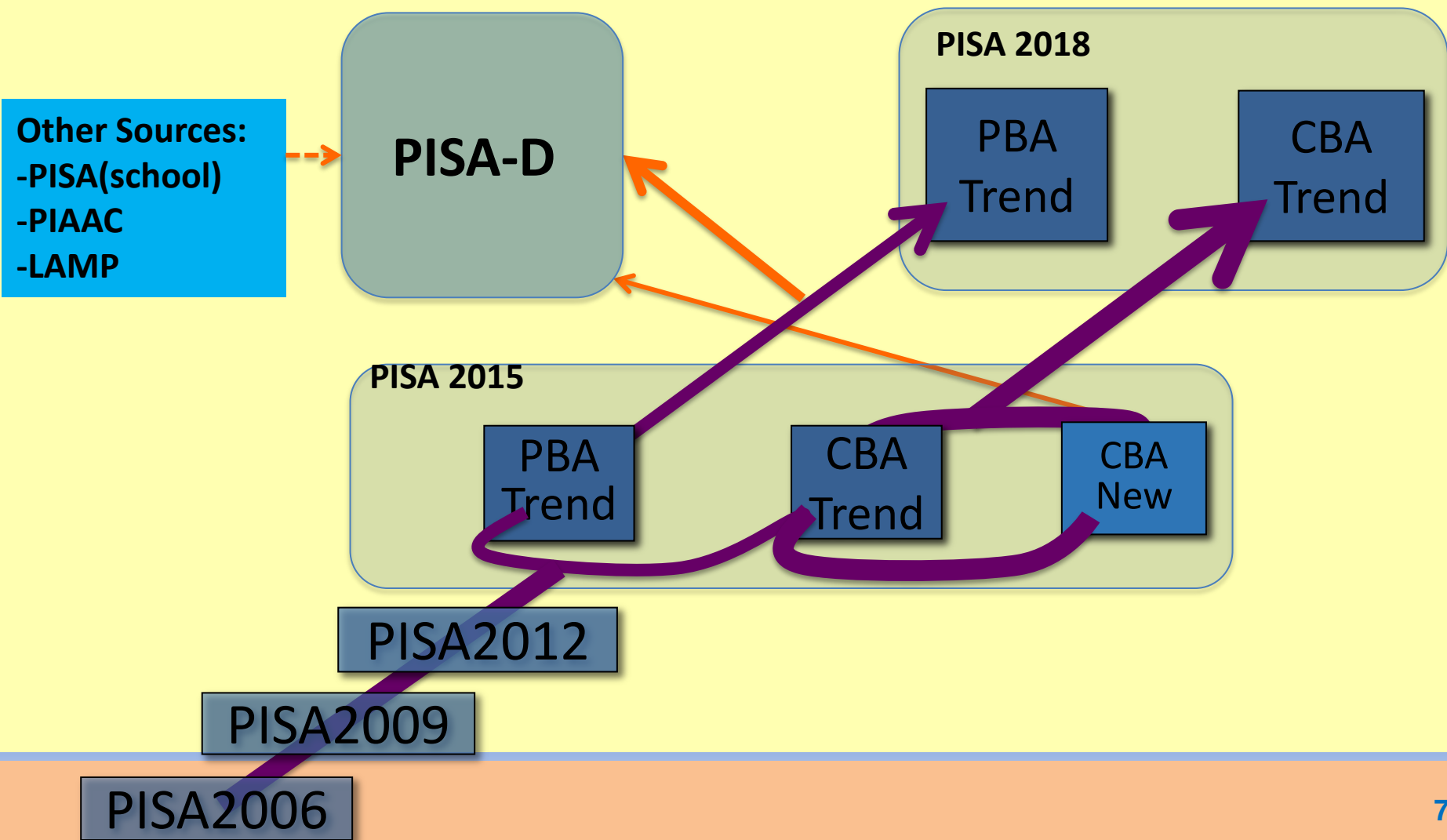
Overview

- Assessment Design
- Data Processing Overview
- Data Quantity and Quality
- IRT Scaling
- Proficiencies
- Concluding Remarks

Assessment Design

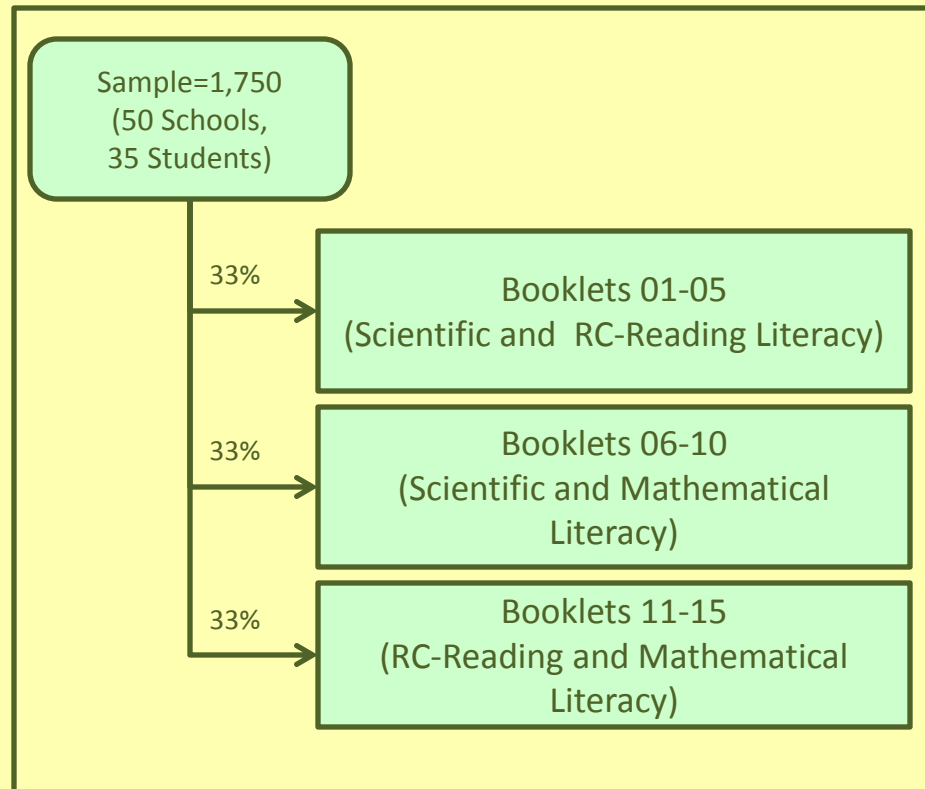
Linking Across Cycles and Modes

PISA 2015/18 and PISA-D



Field Trial Assessment Design

Strand A

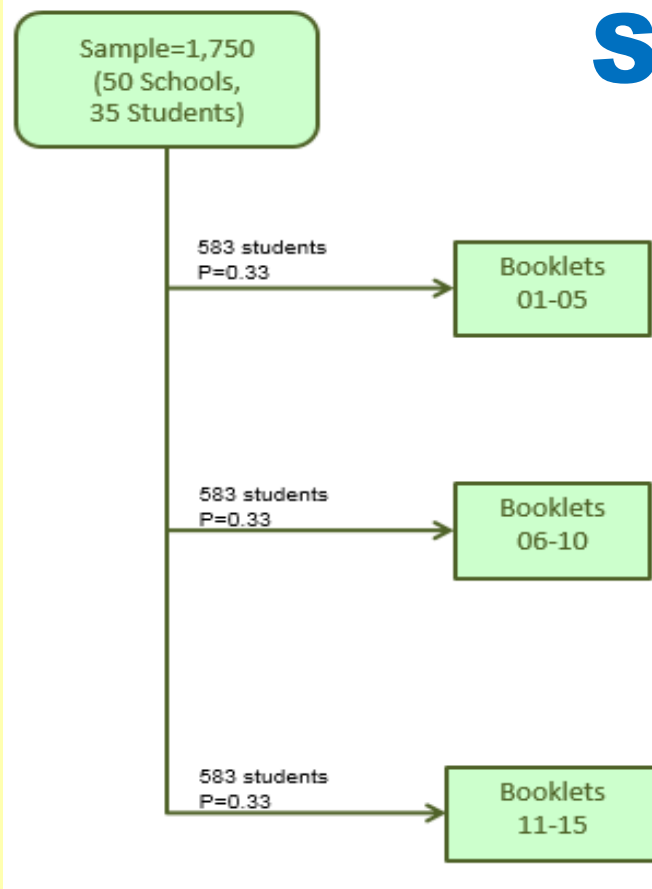


Reading Components (RC):

- 5 5-min clusters (one per Reading cluster pair) examining sentence processing and passage comprehension.
- Each of these clusters include 15 sentences + 1 passage (8-10 items).

Field Trial Assessment Design

Strand A



Booklet	Cluster 1	Cluster 2	Cluster 3	Cluster 4
1	RC-Block1	R1	R2	S1
2		S2	S3	RC-Block5
3	RC-Block3	R3	R4	S3
4		S4	S5	RC-Block2
5	RC-Block5	R5	R1	S5
6		S1	S2	M1
7		M2	M3	S2
8		S3	S4	M3
9		M4	M5	S4
10		S5	S1	M5
11		M1	M2	RC-Block4
12	RC-Block2	R2	R3	M2
13		M3	M4	RC-Block1
14	RC-Block4	R4	R5	M4
15		M5	M1	RC-Block3

Data Processing Overview

Data Processing Overview

- After data submission, national data encountered three phases of data processing:
 1. Pre-processing Phase
 - Review of submitted data file and documentation to ensure data was complete and that all comments were included in documentation
 - Review of consistency checks reports
 2. Processing and Cleaning Phase
 - Data extraction from the raw data
 - Generation of reports for countries to use during verification and/or revision of data for accuracy and consistency
 - Analysis of the OERS Reports to check consistency in cognitive coding procedures.
 3. Psychometrics Analysis
 - Data examination to check whether the design was implemented as specified and to check data quantity and quality
 - Classical and Item Response Theory analysis

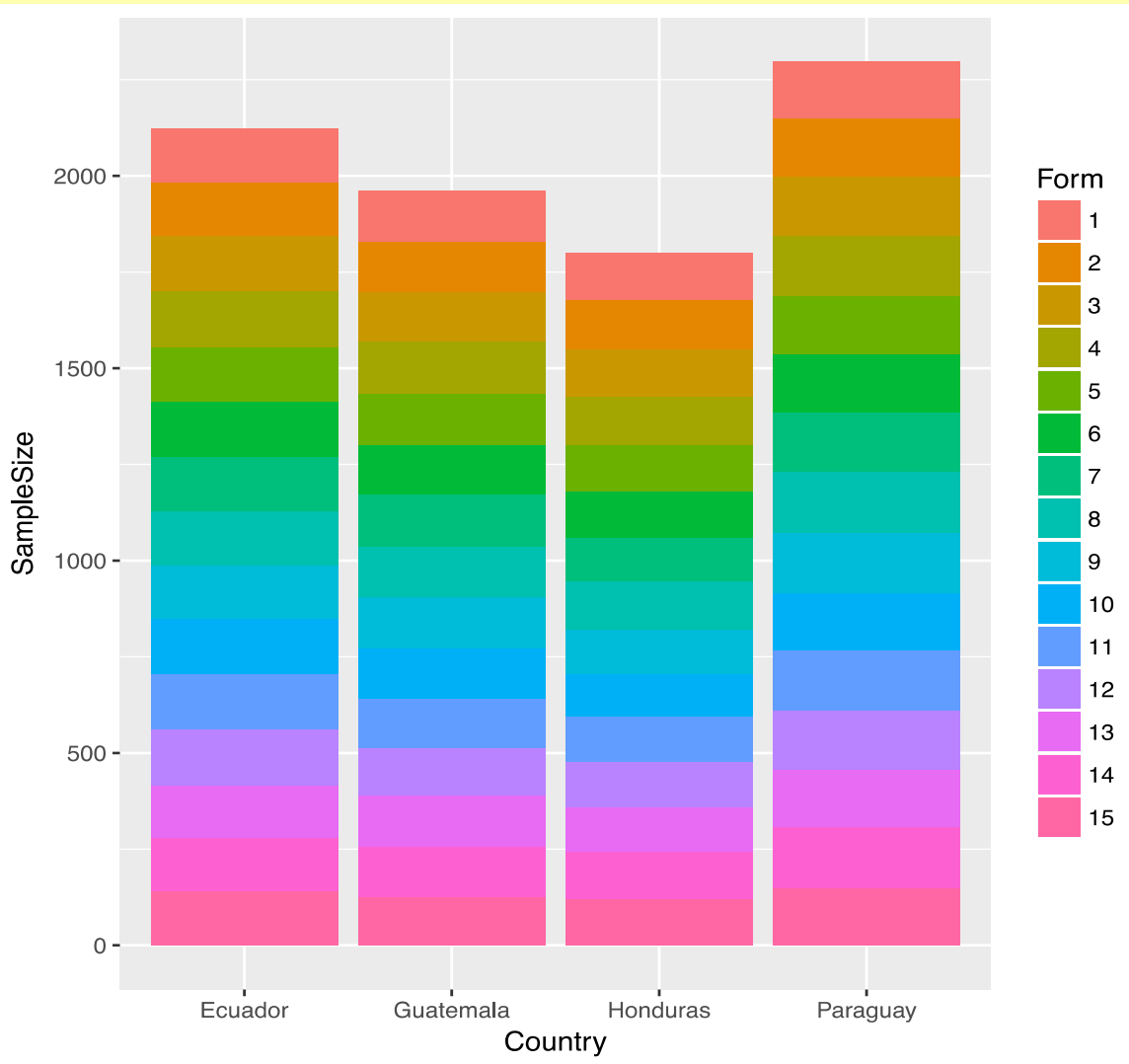
Data Quantity and Quality

Field Trial Data Overview

	PRY	ECU	GTM	HND
FT Yield	2,297	2,123	1,963	1,800

Domain	Number of PISA-D Items
Math	76
Reading	180 (Reading 80 + Reading Components 100)
Science	76

Distribution of Students Across Booklets



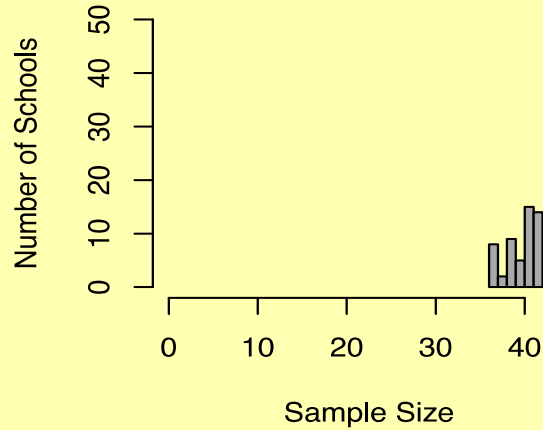
- Evenly distributed with 33-34% of students receiving booklets 1-5 (Reading and Science), 32-33% of students receiving booklets 6-10 (Science and Math), and 32-33% of students receiving booklets 11-15 (Math and Reading) in each country

Distribution of Students per School

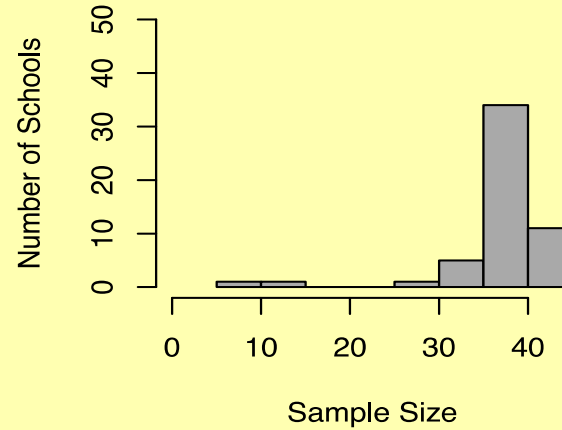
Country	# of Schools	Average # of assessed students per school	# of Schools with less than 10 assessed students
Ecuador	53	40.1	0
Guatemala	53	37.0	1
Honduras	53	34.0	2
Paraguay	64	35.9	3

Distribution of Students per School

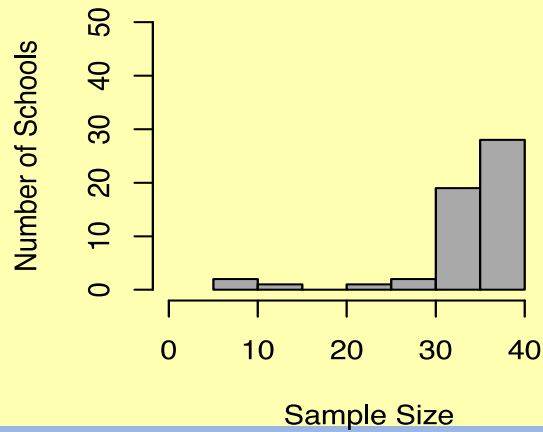
Ecuador



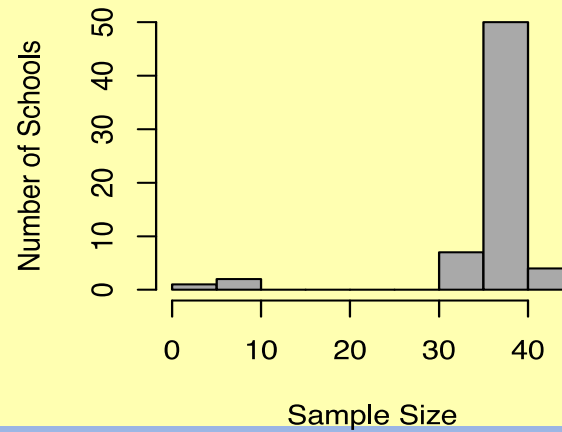
Guatemala



Honduras



Paraguay



Source of Items – Math

Source	# of PISA-D Items	% of PISA-D Items
PISA 2015 Trend	45	59.20
PIAAC	13	17.10
PISA for Schools	18	23.70
Total	76	100.00

→ All PISA-D items were linked to the PISA 2015 Trend scale.

Source of Items – Reading

Source	# of PISA-D Items	% of PISA-D Items
PISA 2015 Trend	54	67.50
PISA for Schools	10	12.50
E&S Online	6	7.50
LAMP	5	6.25
PIAAC	5	6.25
Total	80	100.00
Reading Components	100	100.00

→ All PISA-D items were linked to the PISA 2015 Trend scale.

Source of Items – Science

Source	# of PISA-D Items	% of PISA-D Items
PISA 2015 Trend	48	63.20
PISA 2015 New	12	15.80
PISA for Schools	16	21.00
Total	76	100.00

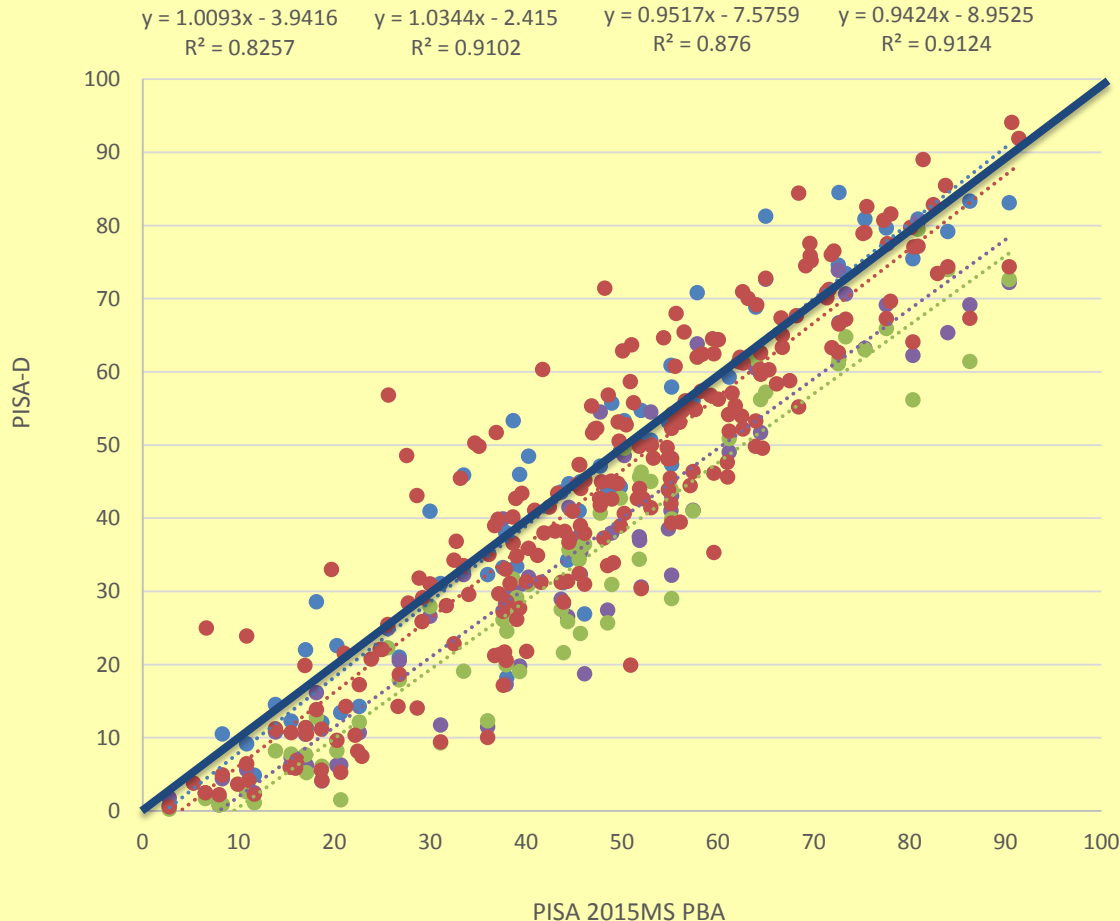
- All PISA-D items were linked to the PISA 2015 Trend scale.
- New PISA 2015 Science CBA items were transformed into PBA items for PISA-D; then the PISA 2015 CBA item parameters were used in the IRT scaling.

Item Analysis

- Classical Test Theory was used to analyze items at both the national and international level
- More specifically, item difficulties (proportion correct, $P+$), frequencies of scores, cluster score and point-biserial (correlation between item score and cluster score) were examined separately for each country
- These allow for the identification of outliers and cases with obvious coding rules deviations

P+ comparison for MATH linking items

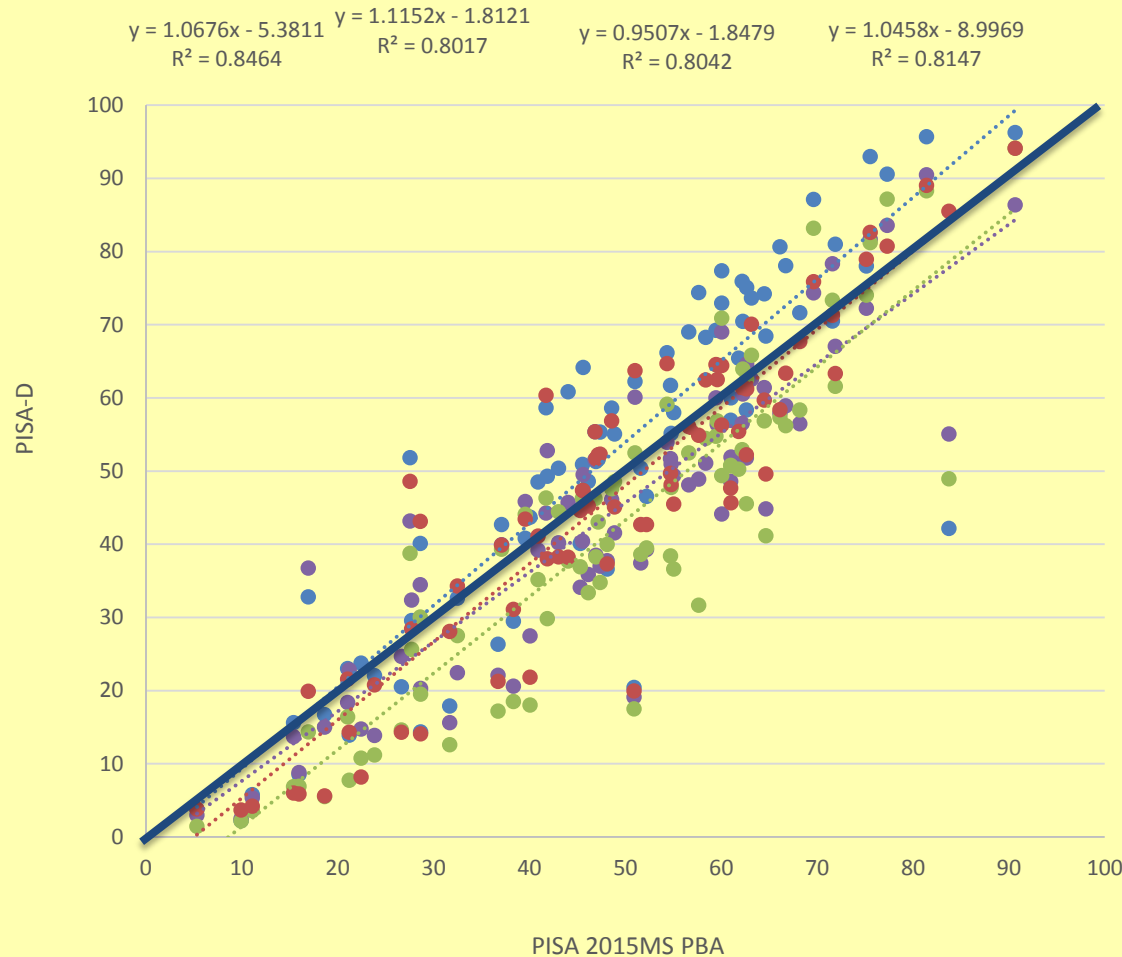
Math Items



- Items for PISA-D behaved similarly to PISA 2015 MS (i.e., parallel regression lines to the 45-degree line)
- Wide range of item difficulties (we have a good distribution of easy, medium, difficult items in PISA-D)
- No systematic differences were observed across countries (i.e., assessment worked in a similar way across countries)

P+ comparison for READING linking items

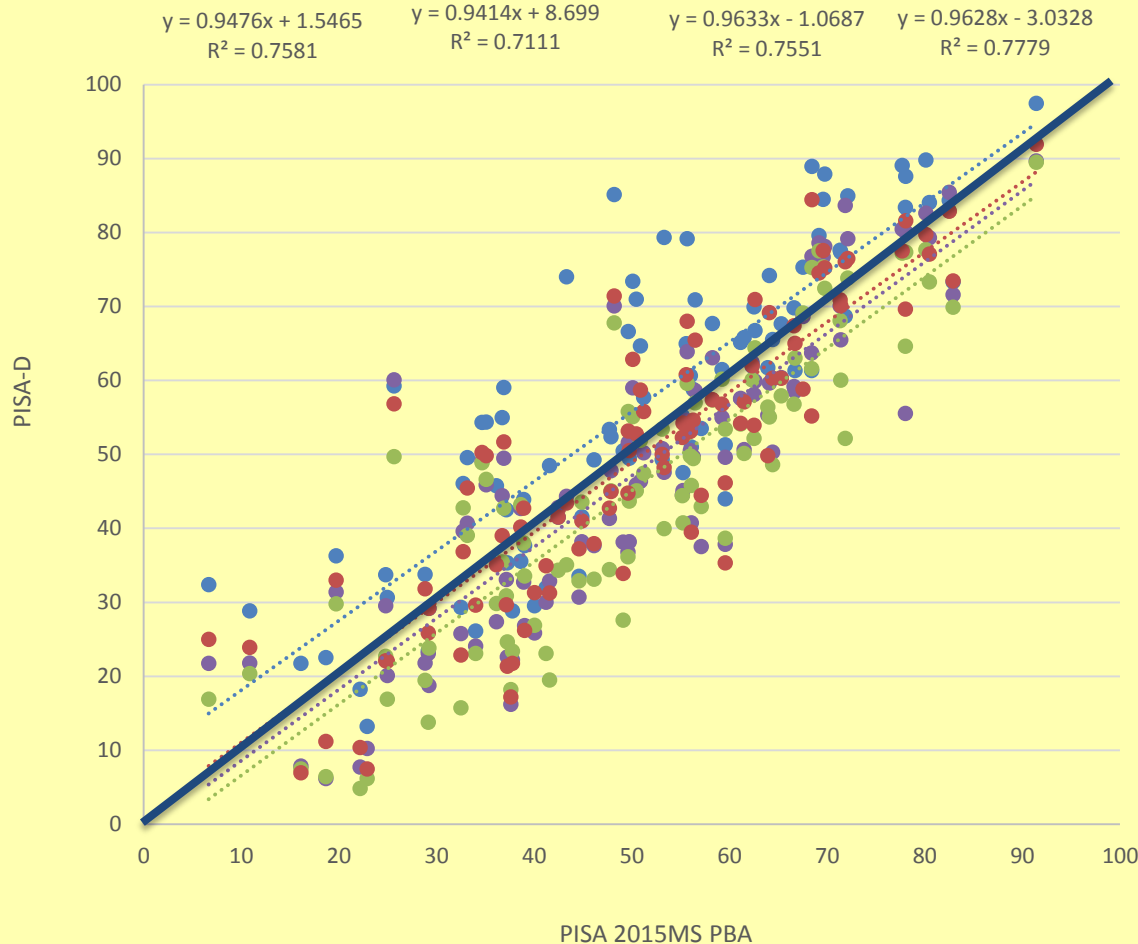
Reading Items



- Items for PISA-D behaved similarly to PISA 2015 MS (i.e., parallel regression lines to the 45-degree line)
- Wide range of item difficulties (we have a good distribution of easy, medium, difficult items in PISA-D)
- No systematic differences were observed across countries (i.e., assessment worked in a similar way across countries)

P+ comparison for SCIENCE linking items

Science Items



→ Items for PISA-D behaved in a similarly to PISA 2015 MS (i.e., parallel regression lines to the 45-degree line)

→ Wide range of item difficulties (we have a good distribution of easy, medium, difficult items in PISA-D)

→ No systematic differences were observed across countries (i.e., assessment worked in a similar way across countries)

Item Correlation

- Item-by-item correlations were examined to monitor local dependencies among items, which is an assumption for IRT.
- Local dependency means: if one item could not be solved, the other (dependent) item cannot be solved either.
- Within-cluster correlations are compared against the across-cluster correlations.
- High correlations indicate items are too similar to each other (e.g., repetitive items).
- Low to medium correlations are more desirable because they contribute to the test information.

Item Correlation – Math

	mean	min	max
M1	0.147	-0.010	0.298
M2	0.149	-0.002	0.404
M3	0.143	-0.010	0.365
M4	0.209	0.019	0.445
M5	0.151	-0.029	0.429
Math	0.157	-0.071	0.445

Item Correlation – Reading

	mean	min	max
R1	0.123	-0.021	0.382
R2	0.122	-0.037	0.523
R3	0.124	-0.023	0.451
R4	0.162	-0.014	0.359
R5	0.161	0.011	0.333
Reading	0.132	-0.080	0.523

Item Correlation – RC

	mean	min	max
RC1	0.244	0.037	0.724
RC2	0.195	-0.026	0.593
RC3	0.158	-0.024	0.537
RC4	0.128	-0.046	0.499
RC5	0.154	-0.044	0.443
RC	0.178	-0.046	0.724
Reading	0.132	-0.080	0.523
Reading+RC	0.101	-0.129	0.724

Item Correlation – Science

	mean	min	max
S1	0.097	-0.025	0.330
S2	0.144	0.002	0.364
S3	0.123	-0.025	0.298
S4	0.137	0.047	0.278
S5	0.099	0.002	0.269
Science	0.117	-0.050	0.364

Monitoring Reliability

- Comparability within and across countries are important design criteria in PISA
 - **Cross-country inter-rater reliability** evaluates consistency of coders across countries to ensure they are applying the same criteria when coding the items
 - *Anchor booklets*: Sets of responses in English to be coded by the bilingual coders
 - **Within-country inter-rater reliability** evaluates the equivalence of coding within a country and identify coding inconsistencies
 - *Students booklets*: Multiple coding students responses by the entire coding team
- Detailed coding designs are provided to ensure the balanced distribution of responses across coders and the appropriate link within and across coders to support the monitoring and evaluation of reliability
- Results from both reliability estimates were found to be acceptable across these countries

IRT Scaling

Using IRT to Create Reporting Scales

- Large scale international assessments such as PISA and PIAAC have been reporting results using described proficiency scales that summarize performance for each participating country as well as subgroups within each country.
- Item parameters, based on response data from all countries, are estimated using a particular IRT model.
- Item Response Theory (IRT) models estimate item parameters (item difficulty α and item discrimination β) and person parameters (students' ability). The scaling approach used in PISA is based on a hybrid Rasch/2PLM for dichotomous data and the GPCM for polytomous data.
- Comparability is achieved through common (international) item parameters across countries, languages and assessment cycles and modes (Trend and New, PBA and CBA).

IRT Scaling in PISA-D

- Comparability of the cognitive domains (Math, Reading, Science) in PISA-D was established (1) to PISA 2015 MS and (2) across 4 countries (PRY, GTM, ECU, HND)
- Linking to PISA 2015 MS was established through a concurrent calibration
 - PISA-D item parameters are fixed to parameters obtained from the PISA 2015 MS where applicable
 - Imposing equal item constraints across countries
- This approach helps to reduce estimation errors assuming that item characteristics for most items are stable

Evaluating item parameters

- In case of deviations (item misfit), item constraints have been released to allow the estimation of unique item parameters – this was expected for a relatively small number of cases across items and countries
- Item fit was evaluated using the Mean Deviation (MD) and the Root Mean Squared Deviation (RMSD)
- MD and RMSD were calculated for all items in each country

IRT Scaling Results Math

	# of items	% of items
Fixed to PISA 2015 MS	38 (Trend)	50.0
Estimated (common across countries)	12 (PIAAC) 18 (PISA-S; 1 no PRY) 5 (Trend)	46.1
Estimated (unique to specific countries)	2 (1 ECU from PIAAC, 1 GTM from Trend)	2.6
Deleted	1 (very low slope, 0.10)	1.3

- Strong linkage established to PISA 2015 MS and across countries
- 99.3% of common item parameters (when calculated based on the item-by-country interactions)

IRT Scaling Results Reading

	# of items	% of items
Fixed to PISA 2015 MS	50 (Trend)	62.50
Estimated (common across countries)	25 (1 Trend + 24 Other)	31.25
Estimated (unique to specific countries)	3 (2 ECU, 1 GTM from Trend)	3.75
Deleted	2 (low item-total corr.)	2.50

- Strong linkage established to PISA 2015 MS and across countries
- 99.0% of common item parameters (when calculated based on the item-by-country interactions)

IRT Scaling Results Reading Components

	# of items	% of items
Fixed to PISA 2015 MS	--	--
Estimated (common across countries)	99	99.0
Estimated (unique to specific countries)	1 (HND)	1.0
Deleted	0	0.0

- Strong linkage established across countries
- 99.75% of common item parameters (when calculated based on the item-by-country interactions)

IRT Scaling Results Science

	# of items	% of items
Fixed to PISA 2015 MS	44 (37 Trend + 7 New)	57.9
Estimated (common across countries)	14 (10 Trend +4 New) 16 (PISA for Schools)	39.5
Estimated (unique to specific countries)	1 (HND, New)	1.3
Deleted	1 (low item-total corr.)	1.3

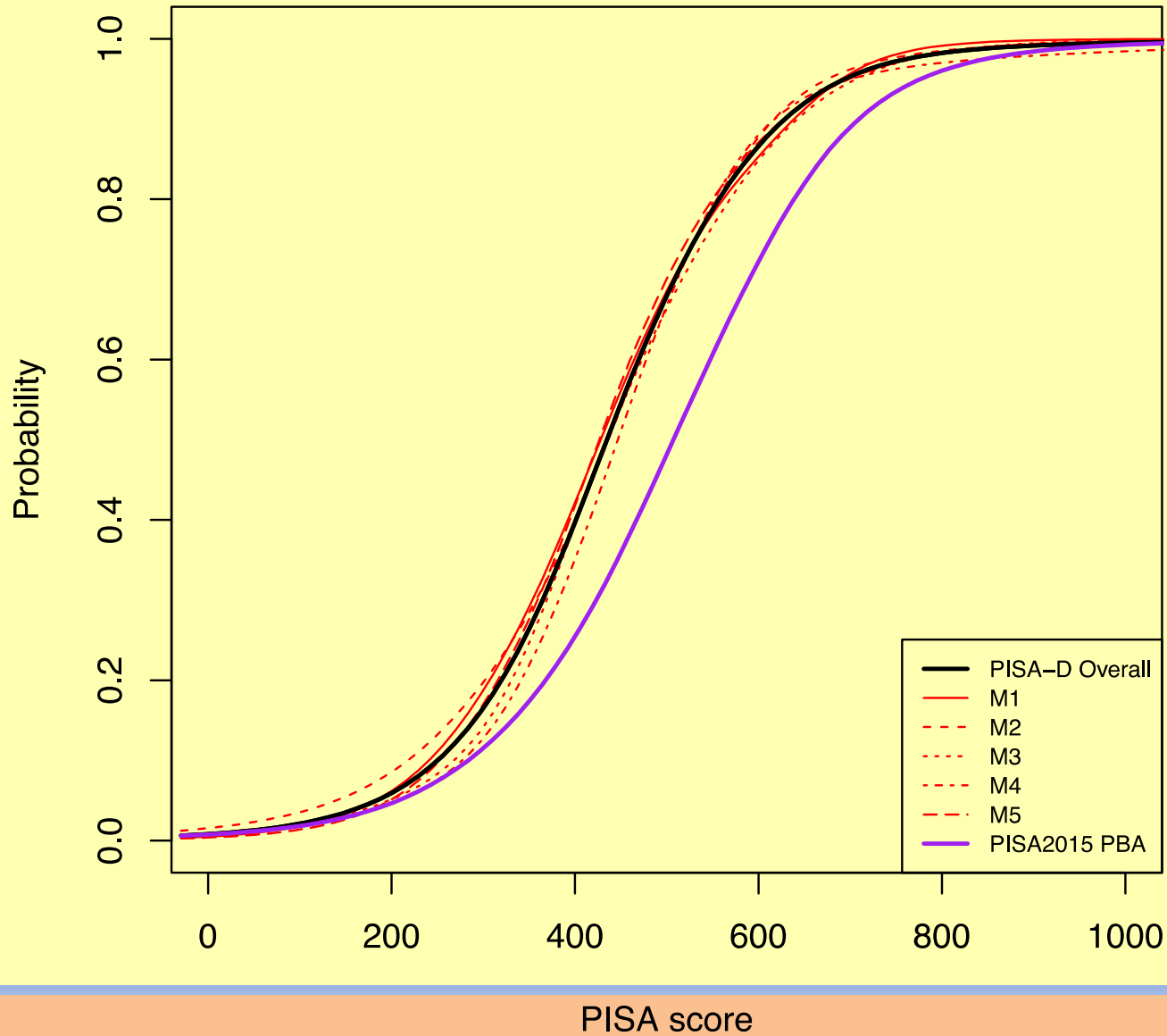
- Strong linkage established to PISA 2015 MS and across countries
- 99.7% of common item parameters (when calculated based on the item-by-country interactions)

Test Characteristic Curve

- Using item parameters (discrimination and difficulty), response probabilities are calculated for each item
- Test characteristic curve is the summation of item characteristic curves at the test level or cluster level
- We can see how items and clusters behave on average on the PISA 2015 reporting scale
- We expect clusters behave similarly (i.e., useful to compare the performance of clusters on the PISA 2015 reporting scale)

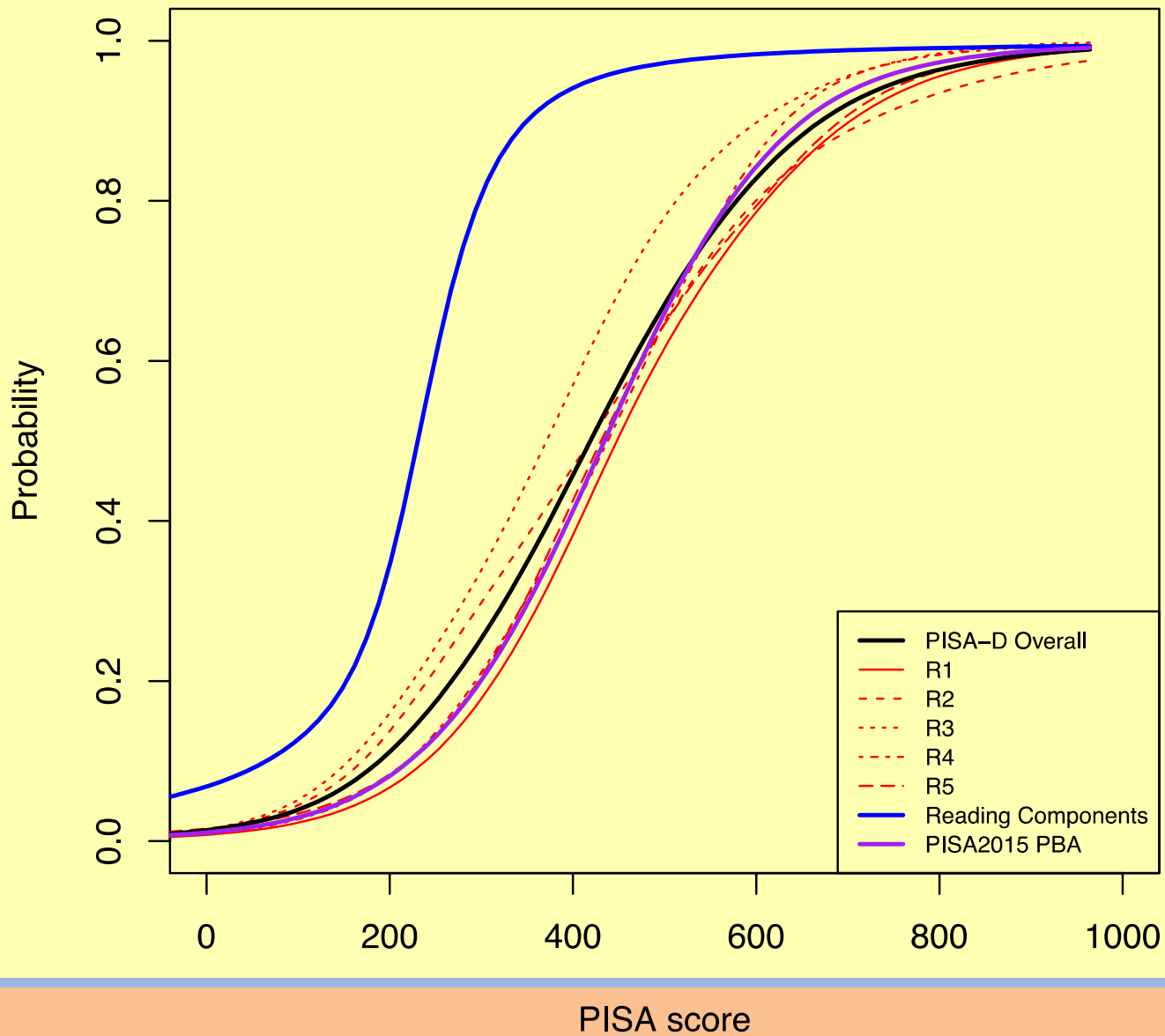
TCC FT Clusters

Math



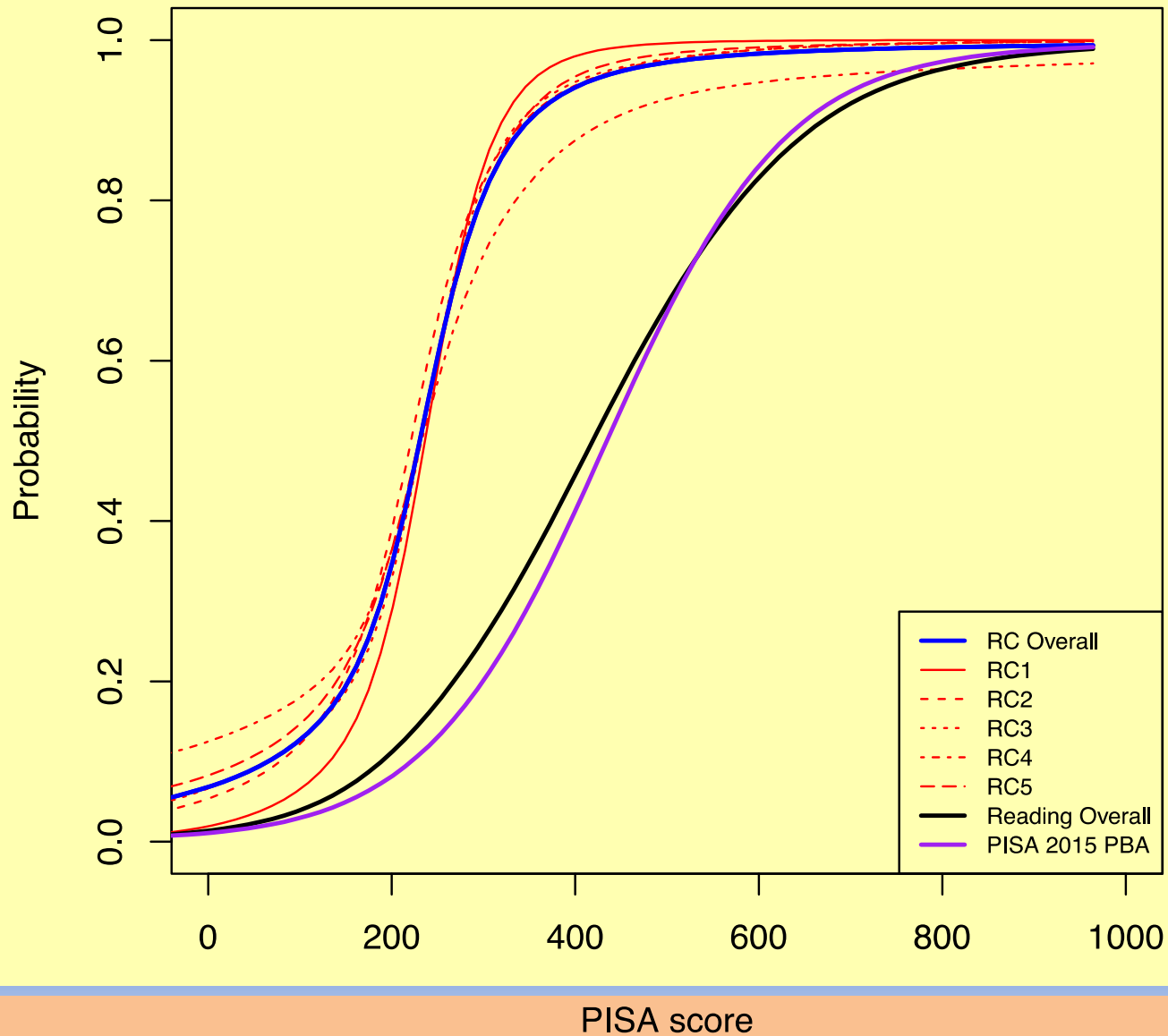
TCC FT Clusters

Reading



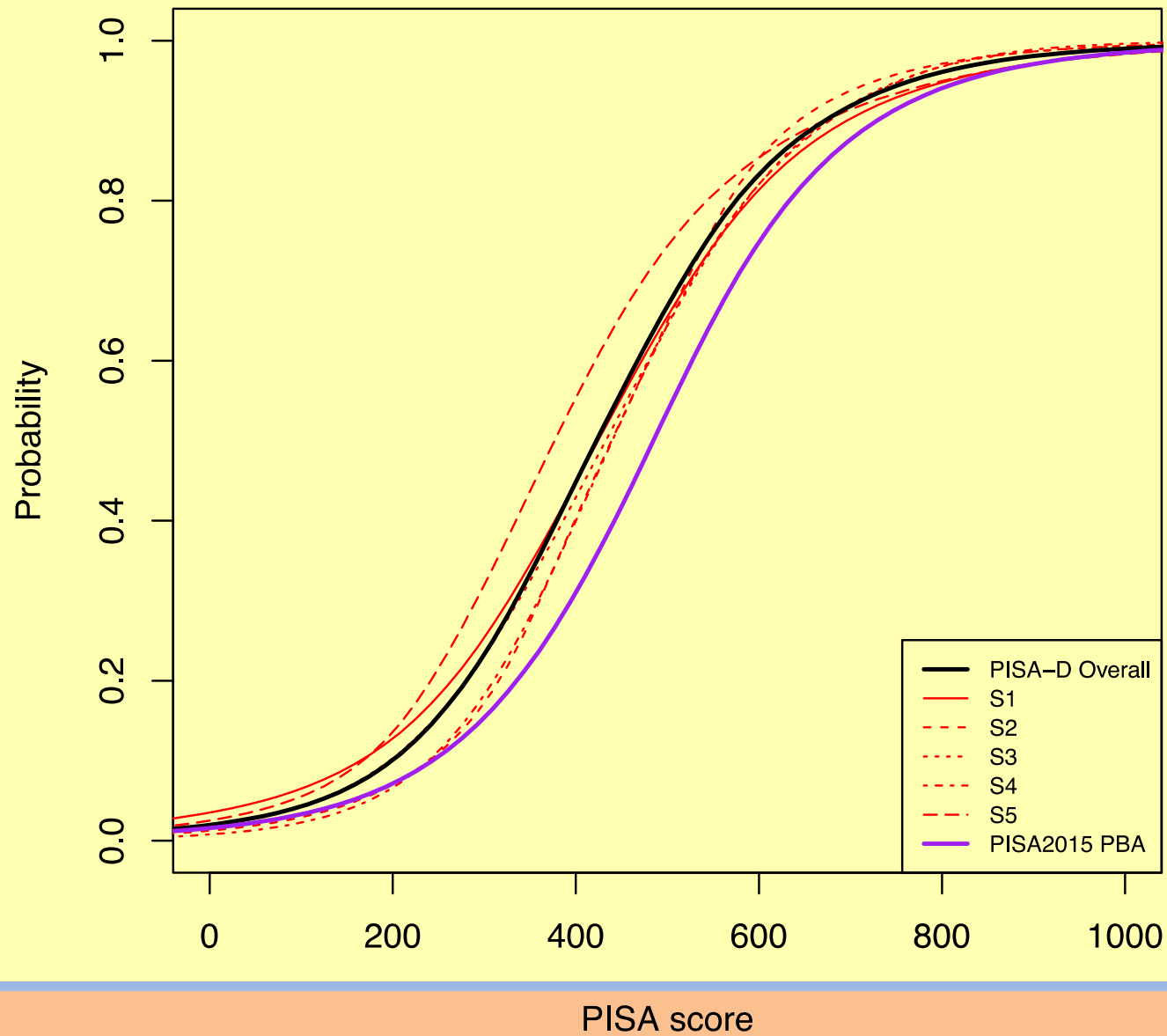
TCC FT Clusters

Reading Components



TCC FT Clusters

Science

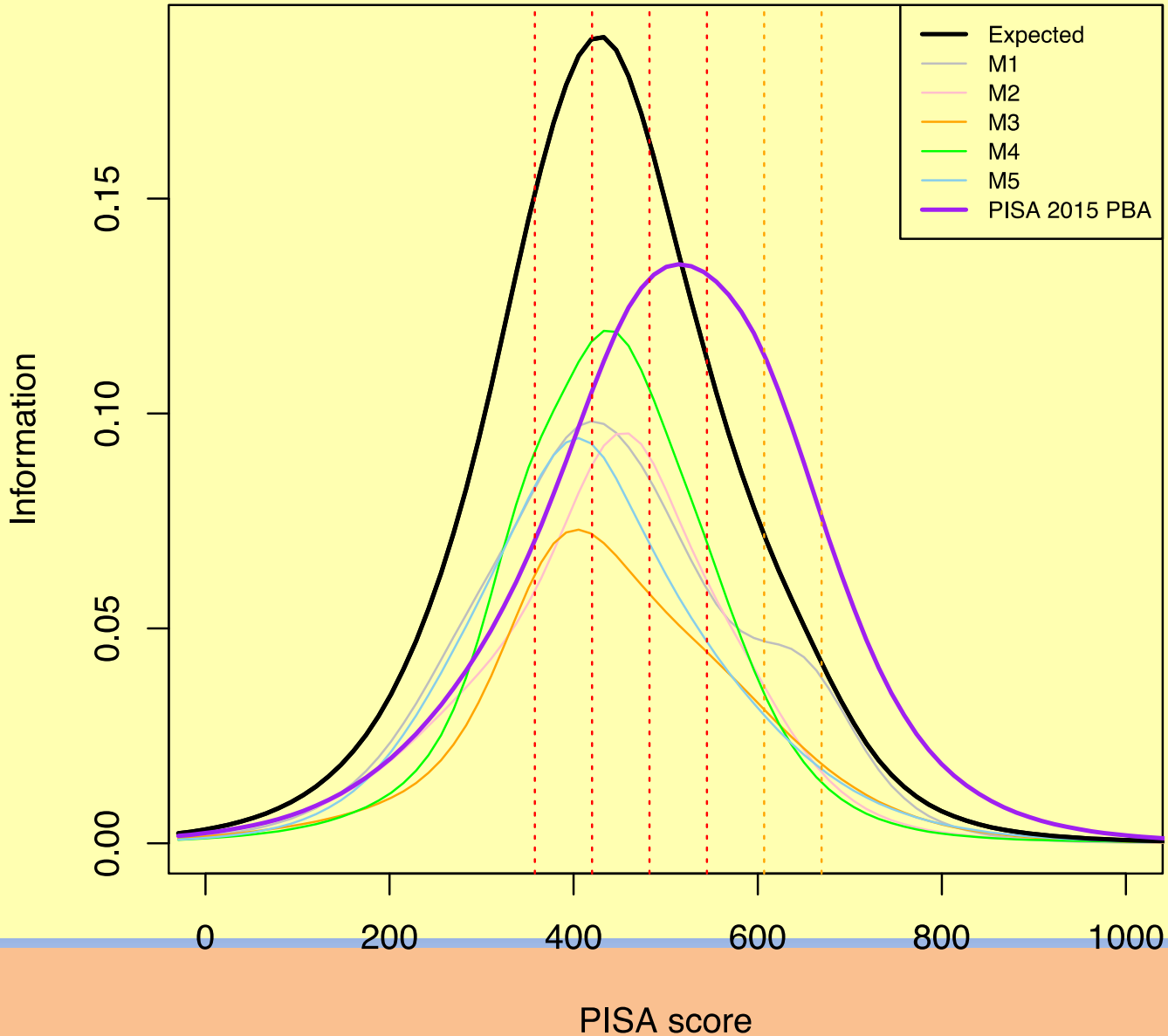


Test Information Function

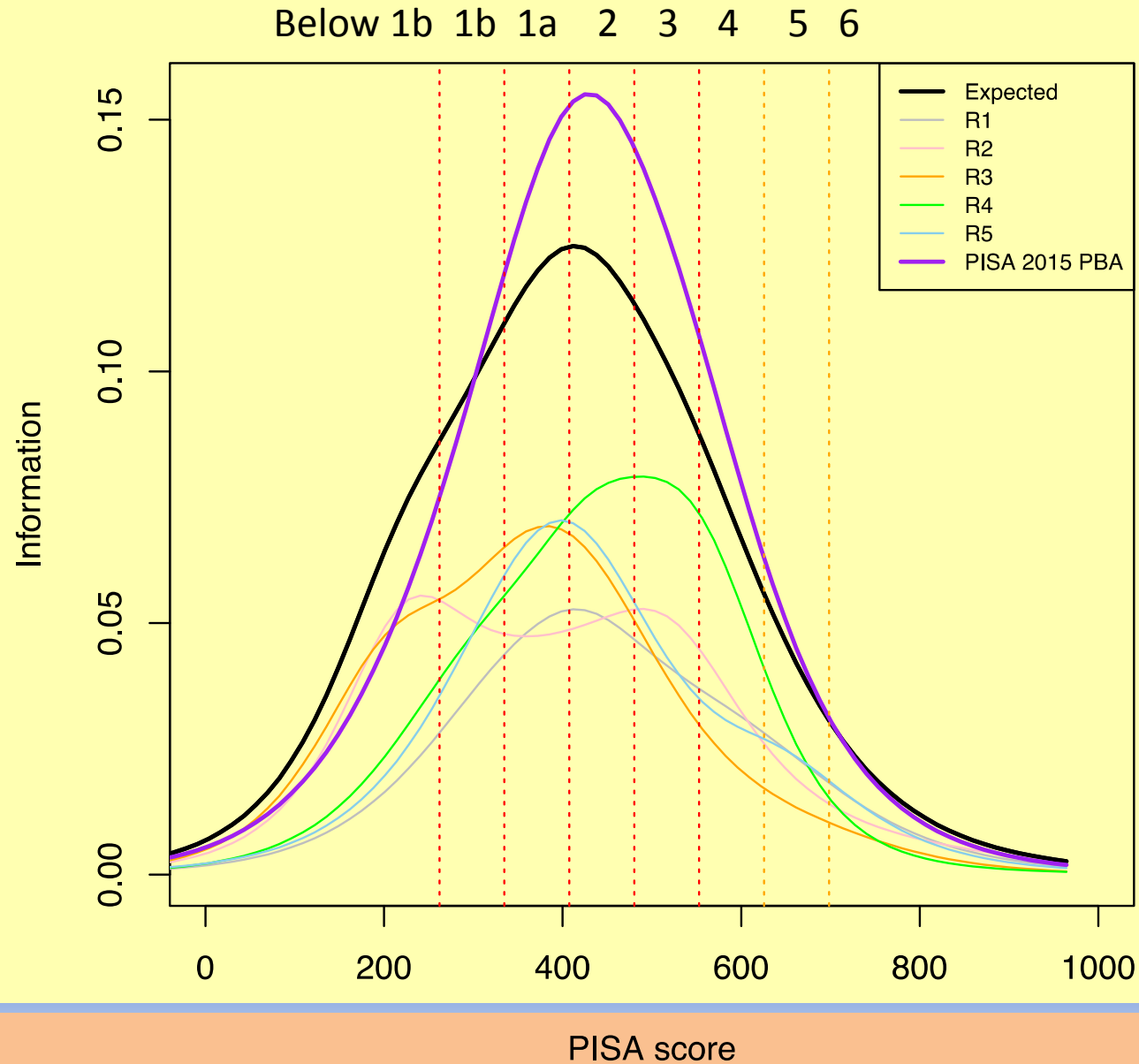
- Using item parameter estimates (discrimination and difficulty), the extent to which information is provided from the test is calculated on the PISA 2015 scale
- Size of the information corresponds to the measurement accuracy
- The highest point of the curve is most accurately measured
- Useful to understand the target population on the PISA 2015 scale

FT Clusters - Math

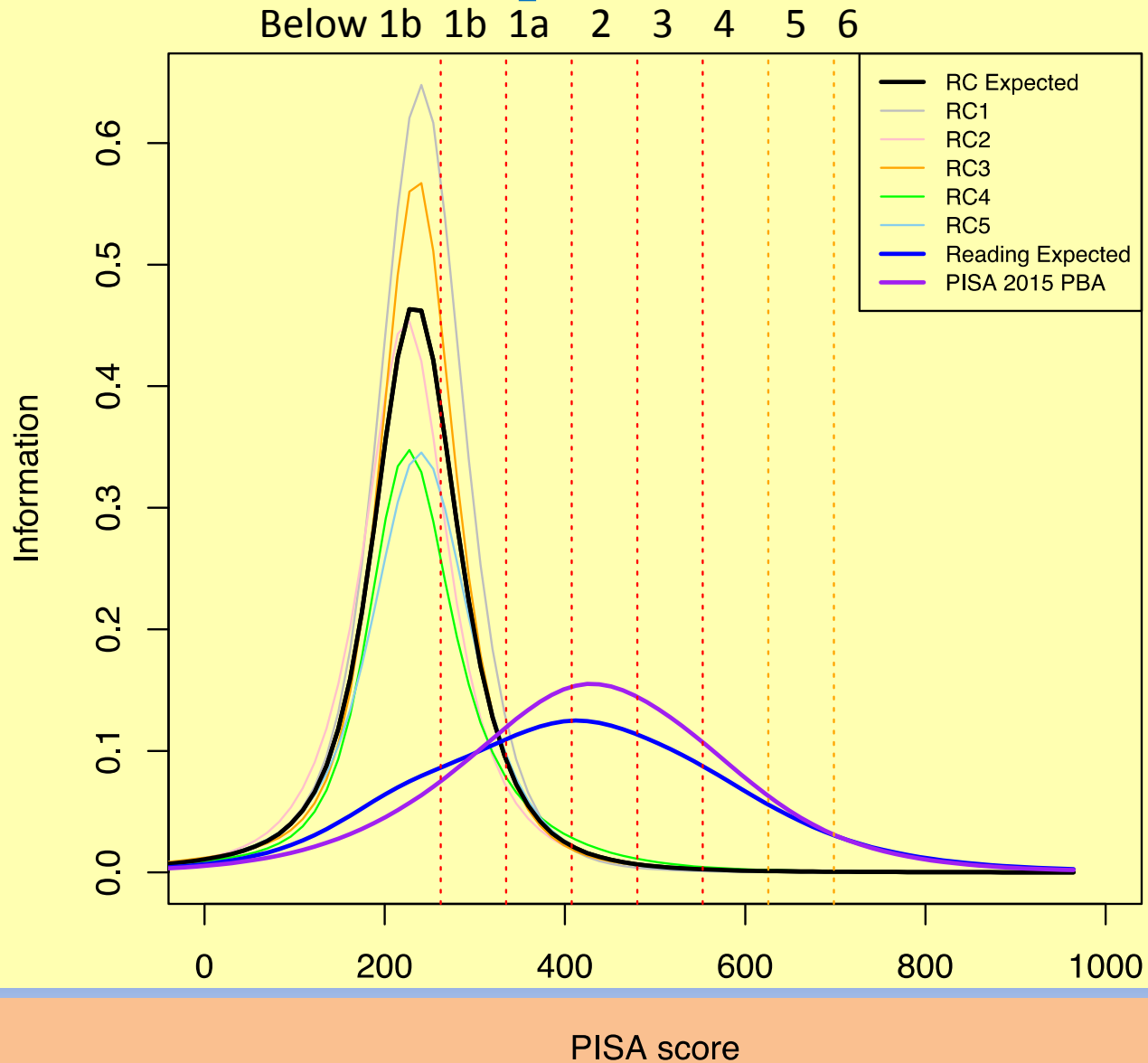
Below 1 1 2 3 4 5 6



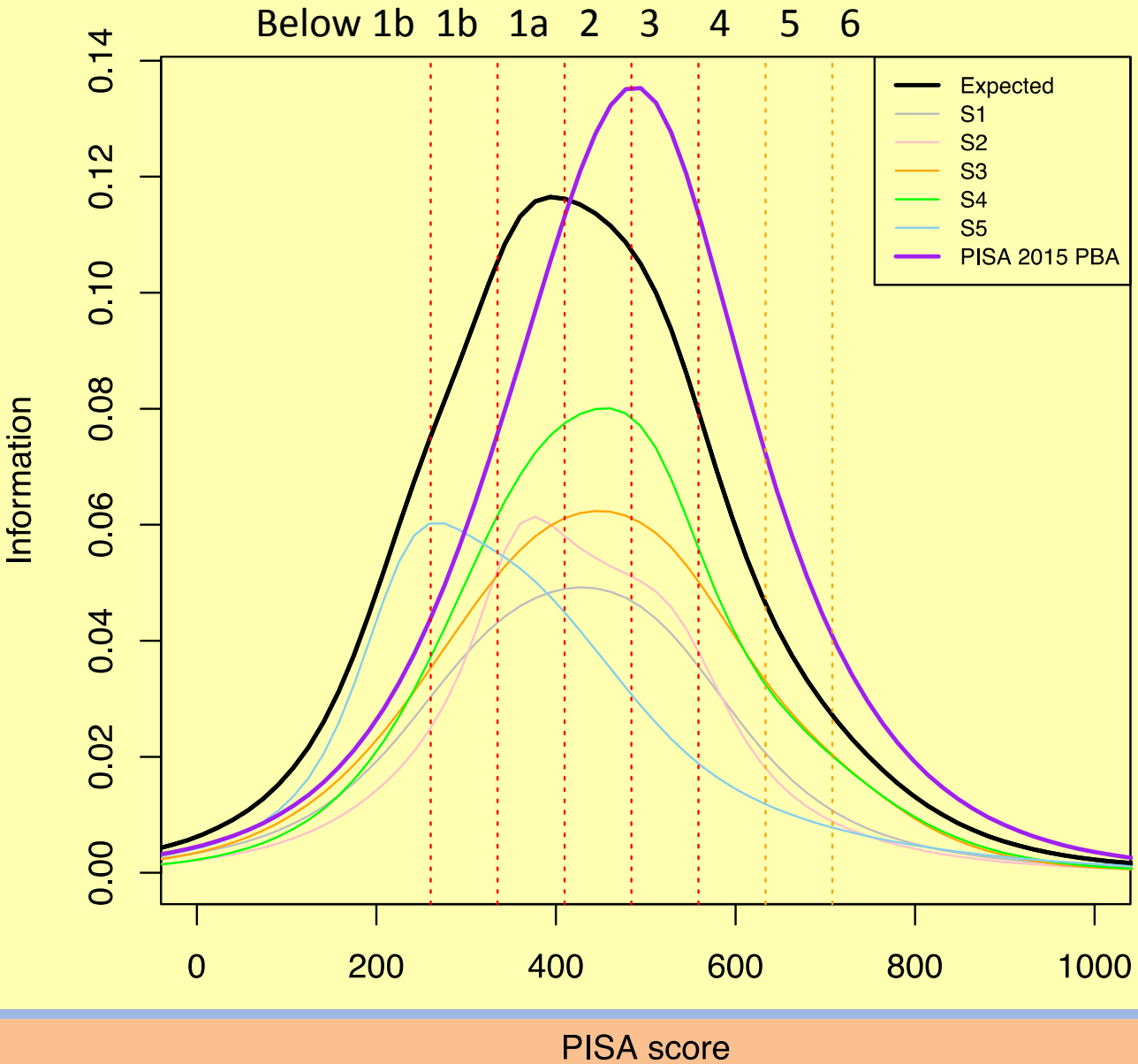
TIF FT Clusters – Reading



TIF FT Clusters – Reading Components



TIF FT Clusters - Science



Summary of IRT scaling results

The preliminary results of the IRT Scaling demonstrated:

- Strong linkage was established to PISA 2015 MS → nearly all linking items received common item parameters and very few item-by-country interactions
- High comparability between 4 countries was achieved through strongly linked design and extended IRT models
- Test information curves demonstrate that the FT assessment targeted low performing students
- The overall assessment worked well

Proficiency Scale

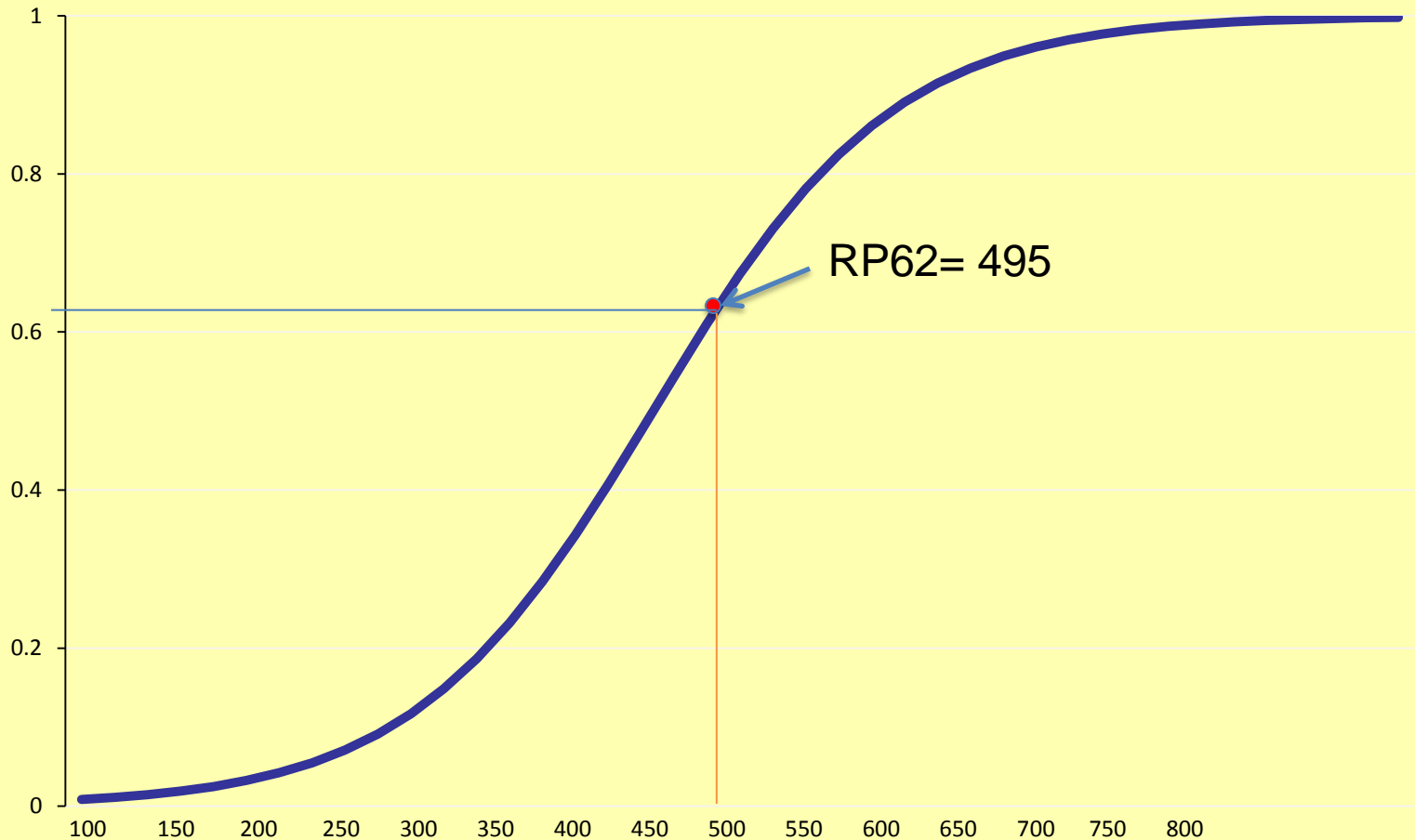
Developing Described Proficiency Scales

- As the tasks used in PISA were drawn from a “population” of tasks representing the constructs as defined in the frameworks, the goal is to draw appropriate inferences about how proficient respondents are with the population of tasks that represent the construct.
- That is, we are interested in drawing inferences about how well they performed on items used in the assessment as well as items having similar characteristics that also represent the construct but were not included in this particular assessment.
- Policy makers and survey organizations have moved to the development of “described proficiency scales”. One characteristic of these scales is the creation of item maps that place items along the scale based on a selected response probability (RP value).

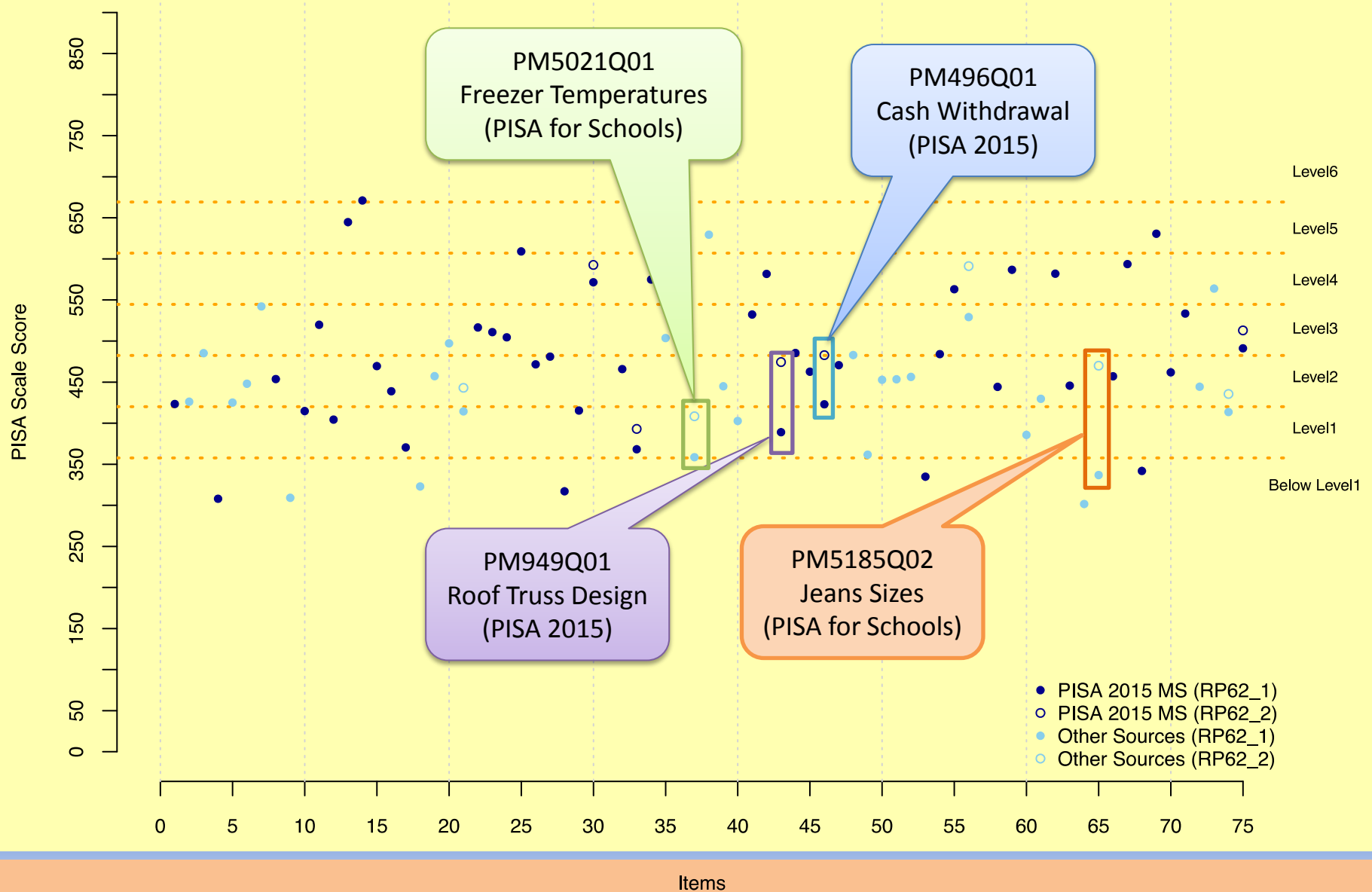
Defining the Proficiency Levels

- An RP value is selected to predict the probability of making correct responses to the group of items that share characteristics.
- An item map that places items along a scale based on a selected response probability (RP) is used to represent the range of difficulty of tasks
- Test items do not discriminate perfectly and each person has a chance (however small) of responding correctly to any given item
- In PISA an RP value of 0.62 is used: **RP62**
- It is important to keep in mind that the selection of a response probability comes after estimation of the item parameters and ability

RP62 and Item Characteristics Curve (ICC)



FT Math Items: RP62



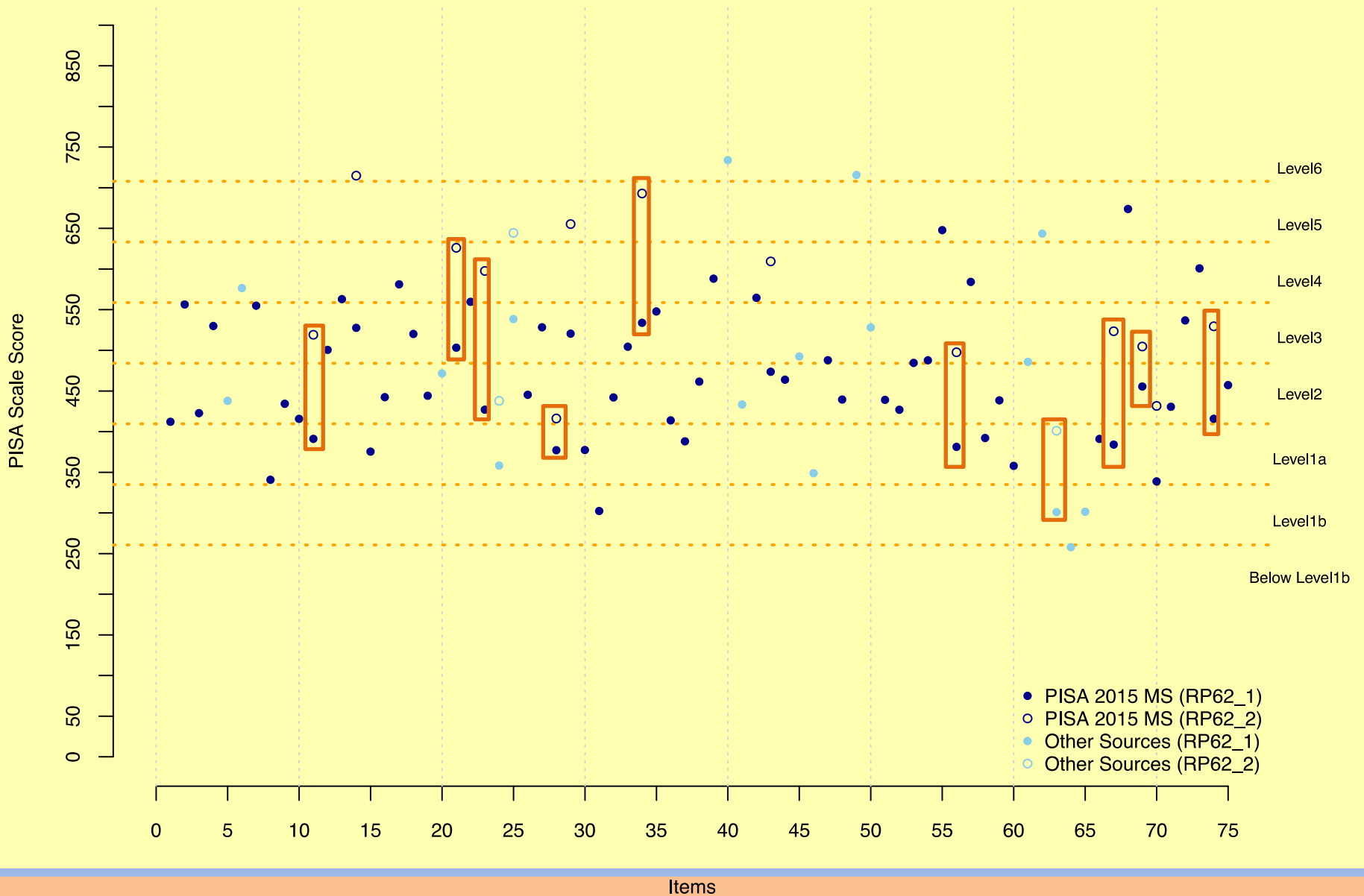
RP62 Summary – Math

Proficiency Level	Score points on the PISA scale	PISA 2015 Trend (%)	PIAAC (%)	PISA for Schools (%)	PISA-D FT (%)	PISA 2015 MS PBA (%)
Below Level 1	Less than 357.77	4 (9.1)	3 (23.1)	1 (5.6)	8 (10.7)	7 (5.7)
Level 1	Higher than 357.77 and less than or equal to 420.07	5 (11.4)	2 (15.4)	2 (11.1)	9 (12.0)	9 (7.4)
Level 2	Higher than 420.07 and less than or equal to 482.38	14 (31.8)	7 (53.8)	6 (33.3)	27 (36.0)	21 (17.2)
Level 3	Higher than 482.38 and less than or equal to 544.68	10 (22.7)	1 (7.7)	3 (16.7)	14 (18.7)	27 (22.1)
Level 4	Higher than 544.68 and less than or equal to 606.99	7 (15.9)	0 (0.0)	3 (16.7)	10 (13.3)	29 (23.8)
Level 5	Higher than 606.99 and less than or equal to 669.30	3 (6.8)	0 (0.0)	1 (5.6)	4 (5.3)	13 (10.7)
Level 6	Higher than 669.30	1 (2.3)	0 (0.0)	2 (11.1)	3 (4.0)	16 (13.1)
Total		44	13 (100)	18 (100)	75 (100)	122 (100)

RP62 Summary – Reading

Proficiency Level	Score points on the PISA scale	PISA 2015 Trend (%)	PISA for Schools (%)	PIAAC, LAMP, ESO (%)	PISA-D FT (%)	PISA 2015 MS PBA (%)
Below Level 1b	Less than 262.04	1 (1.9)	0 (0.0)	5 (35.7)	6 (7.7)	2 (1.5)
Level 1b	262.04 to less than or equal to 334.75	2 (3.7)	0 (0.0)	1 (7.1)	3 (3.8)	7 (5.3)
Level 1a	Higher than 334.75 and less than or equal to 407.47	15 (27.8)	1 (10.0)	2 (14.3)	18 (23.1)	31 (23.3)
Level 2	Higher than 407.47 and less than or equal to 480.18	14 (25.9)	3 (30.0)	4 (28.6)	21 (26.9)	33 (24.8)
Level 3	Higher than 480.18 and less than or equal to 552.89	8 (14.8)	1 (10.0)	0 (0.0)	9 (11.5)	28 (21.1)
Level 4	Higher than 552.89 and less than or equal to 625.61	10 (18.5)	2 (20.0)	0 (0.0)	12 (15.4)	21 (15.8)
Level 5	Higher than 625.61 and less than or equal to 698.32	2 (3.7)	0 (0.0)	2 (14.3)	4 (5.1)	6 (4.5)
Level 6	Higher than 698.32	2 (3.7)	3 (30.0)	0 (0.0)	5 (6.4)	5 (3.8)
Total		54 (100)	10 (100)	14 (100)	78 (100)	133 (100)

FT Science Items: RP62



RP62 Summary – Science

Proficiency Level	Score points on the PISA scale	PISA 2015 Trend (%)	PISA 2015 New (%)	PISA for Schools (%)	PISA-D FT (%)	PISA 2015 MS PBA (%)
Below Level 1b	Less than 260.54	0 (0.0)	0 (0.0)	1 (6.25)	1 (1.3)	0 (0.0)
Level 1b	260.54 to less than or equal to 334.94	0 (0.0)	1 (8.3)	1 (6.25)	2 (2.7)	1 (0.9)
Level 1a	Higher than 334.94 and less than or equal to 409.54	3 (6.4)	4 (33.3)	2 (12.5)	9 (12.0)	5 (4.6)
Level 2	Higher than 409.54 and less than or equal to 484.14	14 (29.8)	5 (41.7)	4 (25.0)	23 (30.7)	23 (21.3)
Level 3	Higher than 484.14 and less than or equal to 558.73	16 (34.0)	1 (8.3)	3 (18.75)	20 (26.7)	39 (36.1)
Level 4	Higher than 558.73 and less than or equal to 633.33	10 (21.3)	0 (0.0)	1 (6.25)	11 (14.7)	27 (25.0)
Level 5	Higher than 633.33 and less than or equal to 707.93	3 (6.4)	1 (8.3)	2 (12.5)	6 (8.0)	8 (7.4)
Level 6	Higher than 707.93	1 (2.1)	0 (0.0)	2 (12.5)	3 (4.0)	5 (4.6)
Total		47 (100)	12 (100)	16 (100)	75 (100)	108 (100)

Concluding Remarks

Concluding Remarks

- Good quality data
- Strong linkage with PISA 2015 through common item parameters for many items
- Very few item-by-country interactions
- Items that showed item-by-country interactions were identified and recommended that they should not be used in MS
- Limitation:
 - Decisions on MS instruments were based only on available data from Spanish speaking countries only.
 - It will be important to check whether results will hold with remaining countries

Questions and Discussion

PISA

for Development



Main Survey Integrated Design

PISA for Development
International Advisory Group
17-19 May 2017

Ann Kennedy, ETS



PEARSON

Integrated Design

Students:

- Cognitive assessment covering the three domains (120 minutes per student)
- A Student Questionnaire session (35 minutes)

School:

- School Questionnaire completed by the school administrator (45 minutes)
- Teacher Questionnaire completed by any teacher who teaches at least 25% of the time (20 minutes)

Cognitive Assessment Goals and Constraints

- Four 30-minute clusters per domain
- Maintain as many PISA trend items as possible (approx. half of the items per domain)
- Maintain distribution with greater number of items at the lower end of the proficiency scale (approx. 60% at or below PISA proficiency scale Level 2)
- Maintain adequate construct coverage across framework aspects
- Consideration of whether the unit also is included in Strand C for adequate linking for Strand C

Student Sample Size and Cognitive Booklet Design

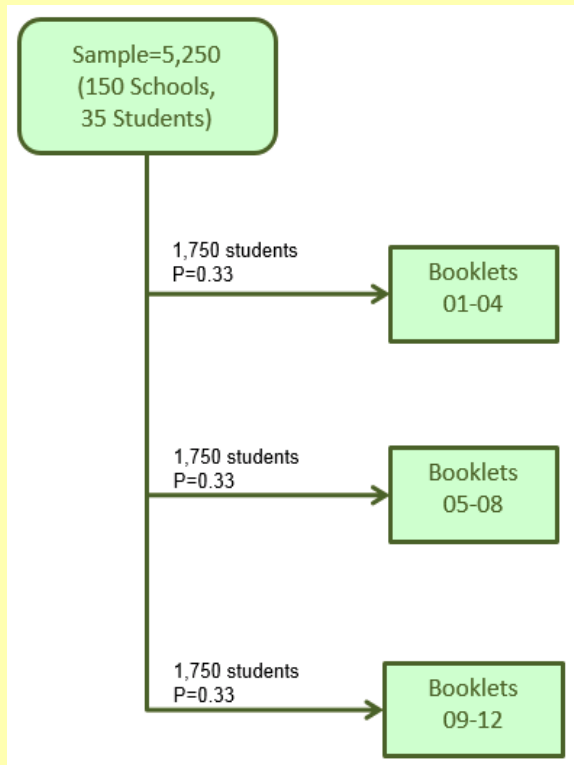
Student sample

- Minimum of 35 students from 150 school
- Representative sample of 15-year-old students
- Total sample size of 5,250 students

Booklet Design

- Each student randomly assigned to one of 12 test booklets following an equal probability design
- Each student takes four 30-minute clusters (two clusters from two of the three domains)
- At least half of the items in each cluster will provide trend information from PISA 2015

Cognitive Assessment Design



Booklet		Cluster 1	Cluster 2		Cluster 3	Cluster 4
1	RC-Block1	R1	R2		S1	S2
2		S2	S3	RC-Block2	R2	R3
3	RC-Block3	R3	R4		S3	S4
4		S4	S1	RC-Block4	R4	R1
5		S1	S2		M1	M2
6		M2	M3		S2	S3
7		S3	S4		M3	M4
8		M4	M1		S4	S1
9		M1	M2	RC-Block1	R1	R2
10	RC-Block2	R2	R3		M2	M3
11		M3	M4	RC-Block3	R3	R4
12	RC-Block4	R4	R1		M4	M1

Selection of Main Survey Items from the Field Trial Item Pool

- Data analysts produced CTT and IRT statistics
- Test developers and psychometricians reviewed items based on:
 - Data quality
 - Psychometric quality of items
 - Coder reliability
 - Content
 - Construct representation
 - Content appropriateness
 - Factors that affect difficulty
- ETS assembled a draft set of MS clusters

Assembling MS Clusters

- Criteria for assigning units to clusters included:
 - maintained FT cluster groupings and order as much as possible
 - relatively equivalent overall difficulty across clusters
 - range of item types by
 - Difficulty
 - Item type (i.e., simple multiple choice, complex multiple choice and constructed response)
 - Classification (i.e., processes, situations)
 - Source (i.e., combination of PISA trend / other sources)

Selection of Main Survey Items from the Field Trial Item Pool, cont.

**Subject Matter Expert Group Meeting
(18-19 April 2017)**

- Data quantity and quality
- Overview of item analyses
- Group discussions based on item statistics and review comments from test developers and psychometricians
- Recommendations for minor revision to draft clusters (i.e., unit sequencing, reclassification)

Number and Distribution of Selected MS Reading Items

	Level 1 and Below	Level 2	Level 3	Level 4	Levels 5-6	Total
PISA 2015 Trend	18	14	8	10	3	53 (79%)
Other Sources	8	3	1	1	1	14 (21%)
Total	26	17	9	11	4	67

Approximately 65% of the total number of items

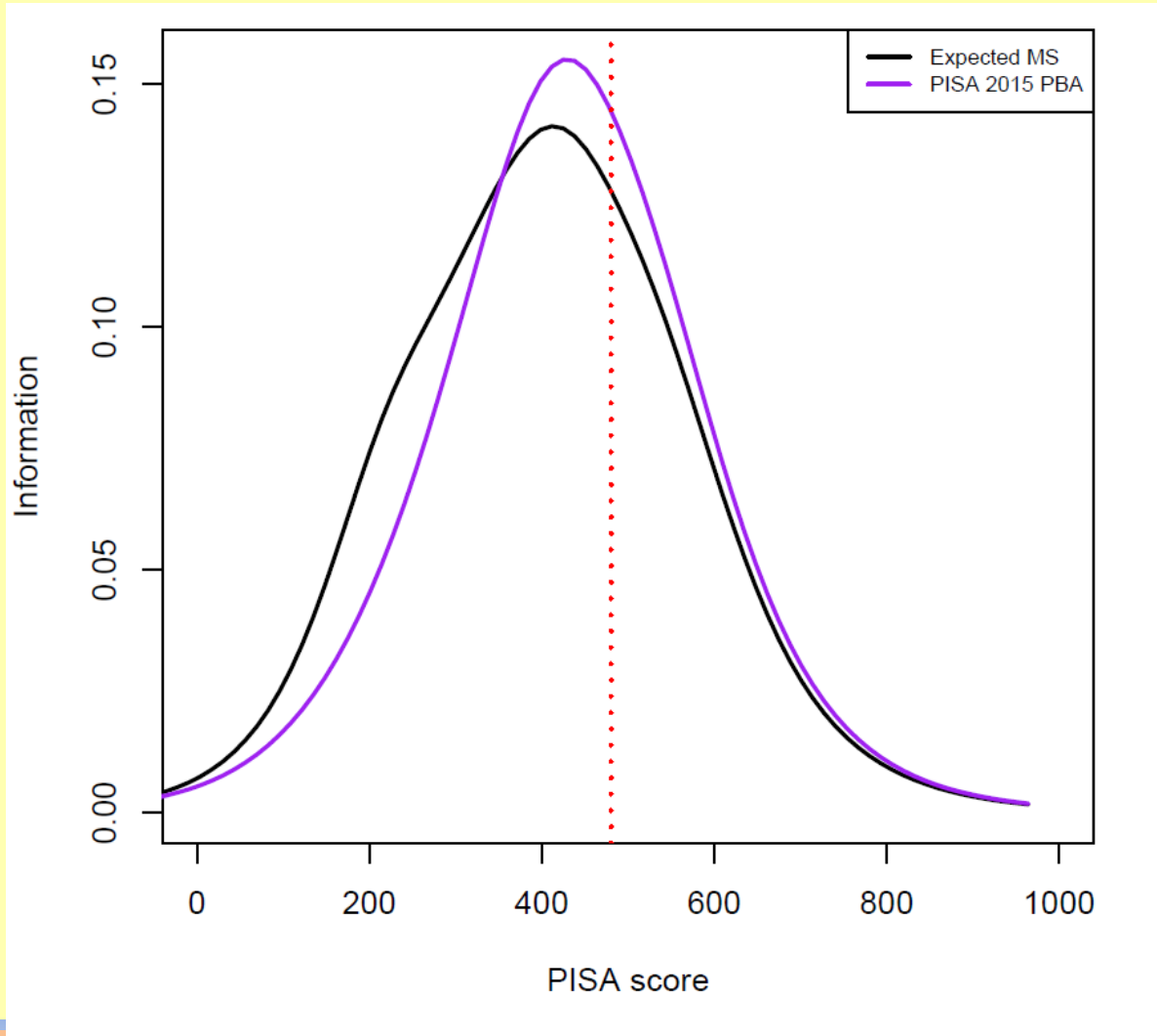
Reading Items by Process

Process	PISA-D Proposed MS		PISA 2015 Trend	Framework Goal
	Strand A			
	Number	%	%	
Access and retrieve	22	33%	29%	25-30%
Integrate and interpret	31	46%	49%	45-55%
Reflect and evaluate	14	21%	23%	15-25%
Total	67	100%	100%	100%

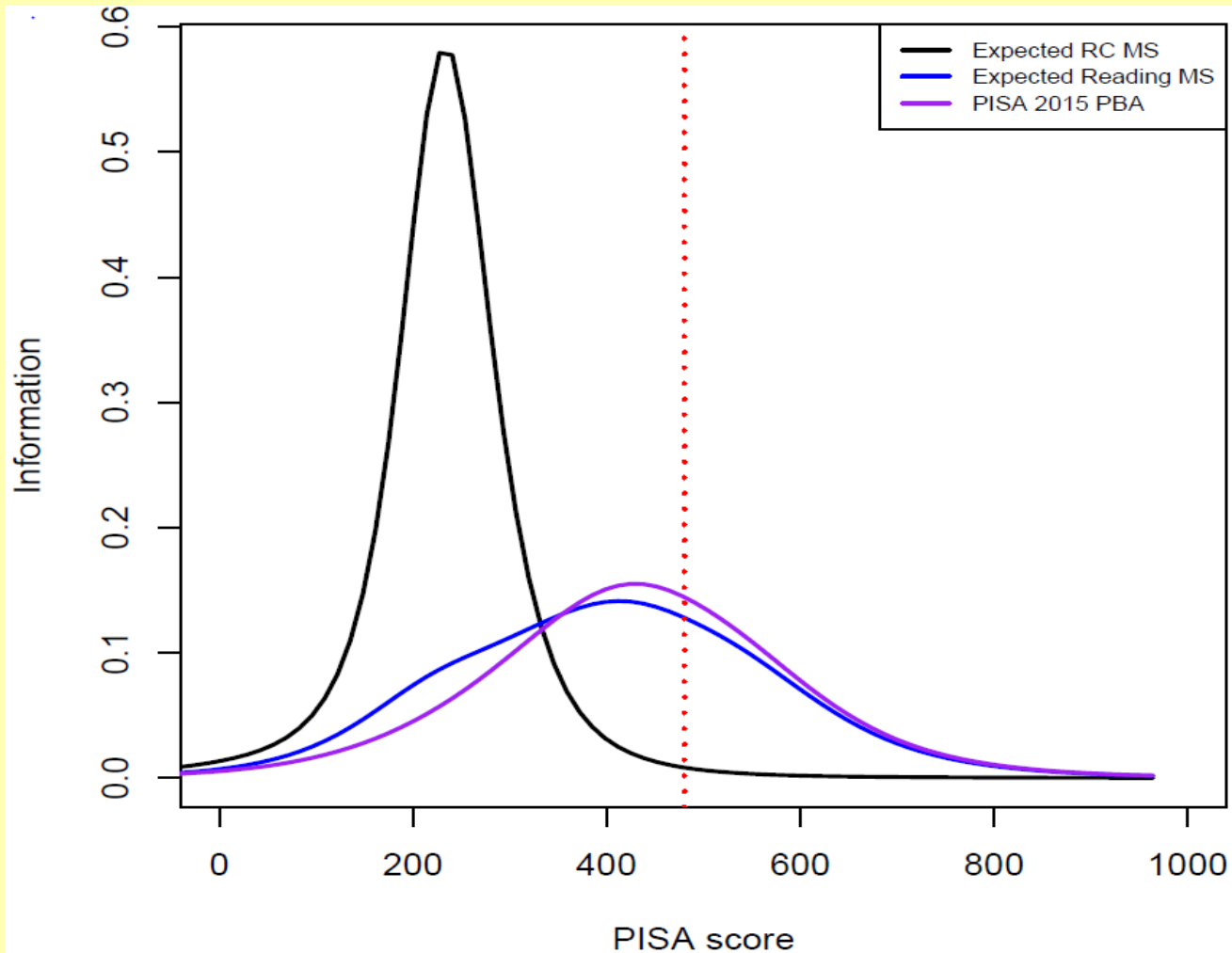
Reading Items by Situation

Situation	PISA-D Proposed MS Strand A		PISA 2015 Trend %	Framework Goal
	Number	%		
Personal	22	33%	27%	25-45%
Educational	21	31%	28%	25-45%
Occupational	4	6%	23%	15-25%
Public	20	30%	22%	5-15%
Total	67	100%	100%	100%

Test Information Curve: Reading



Test Information Curve: Reading Components



Number and Distribution of Selected MS Mathematical Items

	Level 1 and Below	Level 2	Level 3	Level 4	Levels 5-6	Total
PISA 2015 Trend	8	14	8	7	2	39 (62%)
Other Sources	8	12	4			24 (38%)
Total	16	26	12	7	2	63

Approximately 65% of the total number of items

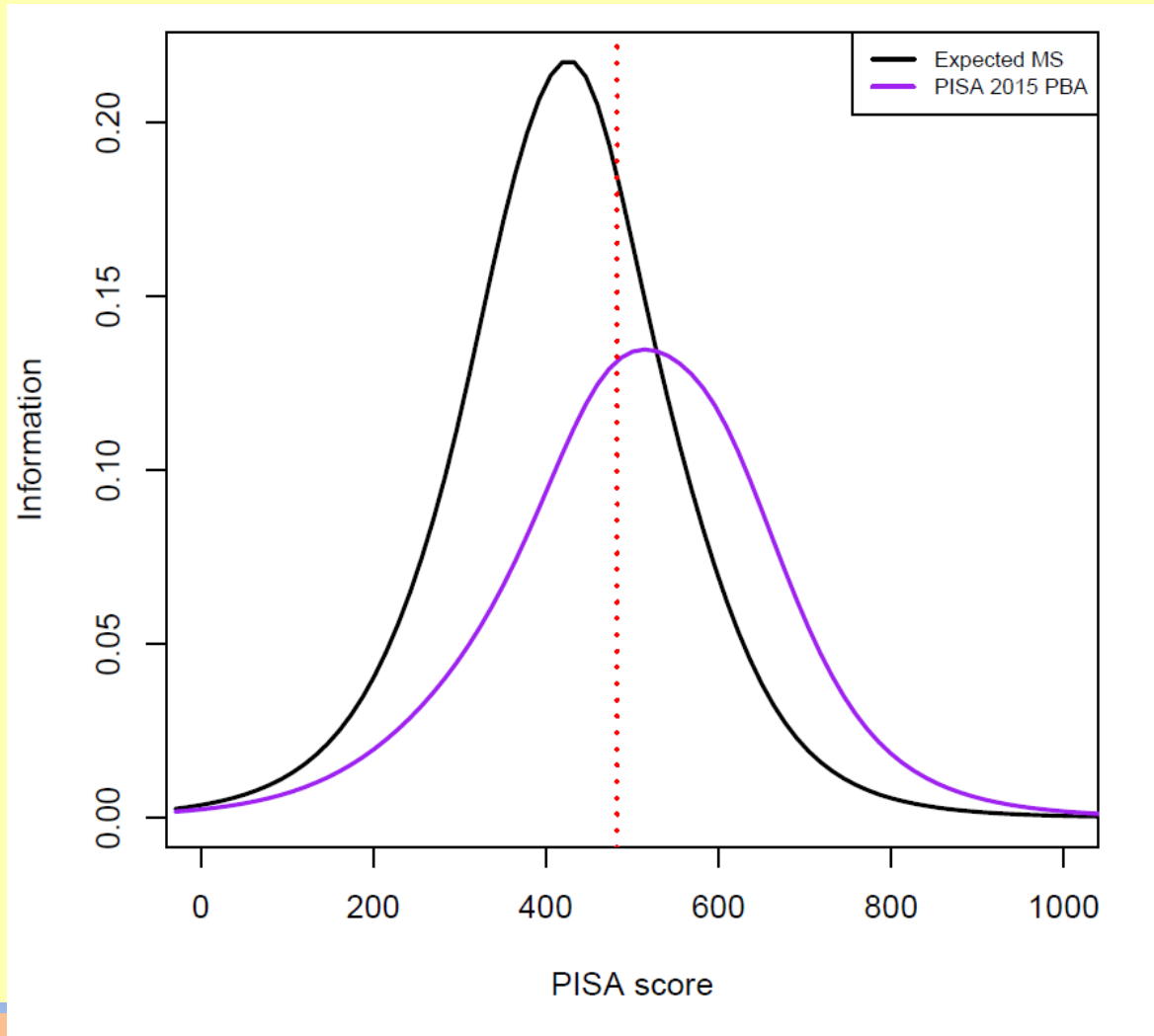
Mathematics Items by Process

Process	PISA-D Proposed MS		PISA 2015 Trend	Framework Goal
	Strand A			
	Number	%	%	
Formulating situations mathematically	13	21%	29%	Approx. 25%
Employing mathematical concepts, facts, procedures	28	44%	43%	Approx. 50%
Interpreting, applying, and evaluating mathematical outcomes	22	35%	28%	Approx. 25%
Total	63	100%	100%	100%

Mathematics Items by Content

Content	PISA-D Proposed MS Strand A		PISA 2015 Trend	Framework Goal
	Number	%	%	
Change and relationships	12	19%	27%	25%
Space and shape	9	14%	23%	25%
Quantity	27	43%	25%	25%
Uncertainty and data	15	24%	25%	25%
Total	63	100%	100%	100%

Test Information Curve: Mathematics



Number and Distribution of Selected MS Science Items

	Level 1	Level 2	Level 3	Level 4	Levels 5-6	Total
PISA 2015 Trend	3	13	16	10	3	45 (68%)
Other Sources	8	7	3	1	2	21 (32%)
Total	11	19	19	11	5	66

Approximately 45% of the total number of items

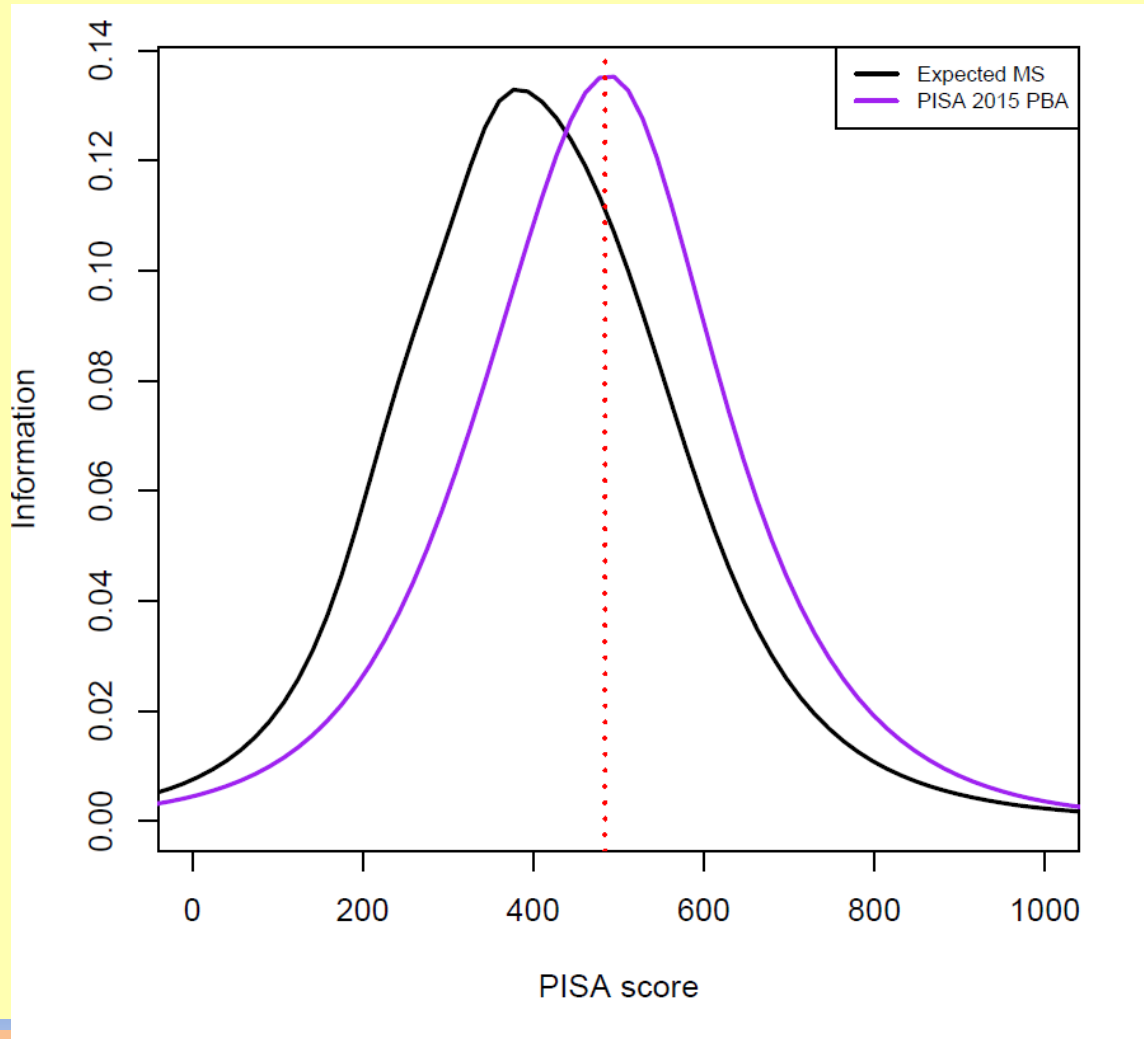
Science Items by Competency

Scientific Competency	PISA-D Proposed MS Strand A		PISA 2015 Trend	Framework Goals
	Number	%	%	
Explaining phenomena scientifically	35	53%	46%	40-50%
Evaluate and design scientific enquiry	13	20%	19%	20-30%
Interpreting data and evidence scientifically	18	27%	34%	30-40%
Total	66	100%	100%	100%

Science Items by Types of Scientific Knowledge

Knowledge Types	PISA-D Proposed MS Strand A		PISA 2015 Trend	Framework Goals
	Number	%	%	
Content	42	64%	62%	55-65%
Procedural	16	25%	27%	20-30%
Epistemic	8	12%	12%	10-20%
Total	66	100%	100%	100%

Test Information Curve: Science



Questions and Discussion

PISA-D Questionnaires

Rationale for selection of questionnaire content and summary of field trial analyses

4th Meeting of the International Advisory Group
17-19 May 2017

Presentation by The Learning Bar



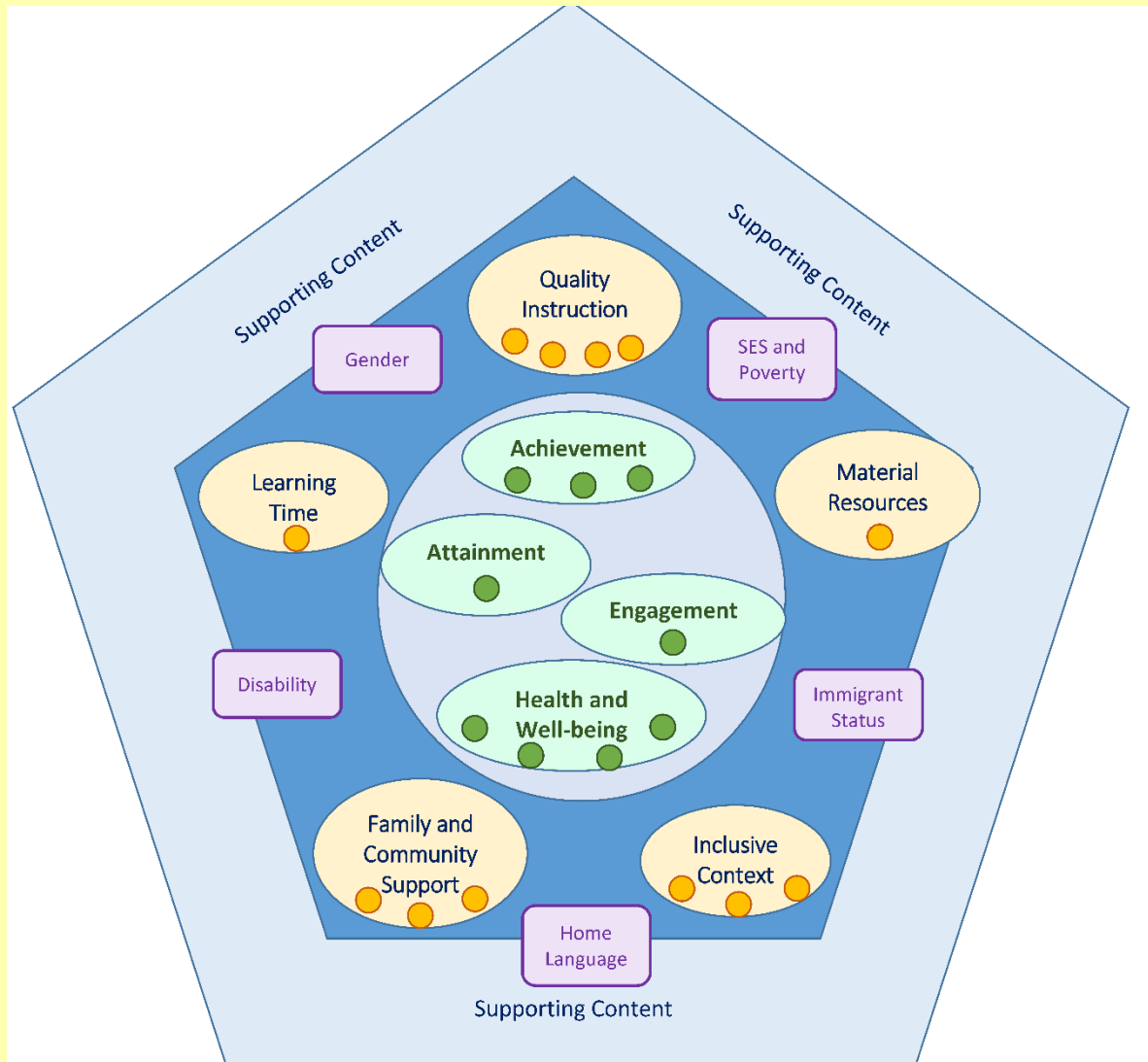
PISA-D Strand B Questionnaires

- Rationale for proposed selection of questionnaire content
 - Core indicators
 - Summary content
- Student Questionnaire prosperity outcomes
- Student Questionnaire foundations for success
- Student Questionnaire equality-equity variables
- Student Questionnaire supporting content
- Teacher Questionnaire
- School Questionnaire

Rationale for proposed selection of questionnaire content

- Consistent with the contextual framework, which is based on the Educational Prosperity model.
- Over-arching goals:
 - Provide reliable measures of the core measures
 - Include measures that can be used to link to 2015 main PISA
 - Provide the constituent components to measure SES and poverty
 - Include a number of measures of supporting content
- Considerable work was conducted to extend the measure of ESCS, to consider an alternate approach for measuring ISCO, and to develop a schema for assessing material possessions.

Core indicators and supporting content for PISA-D



PISA-D Main Study – Proposed Number of Questions for Strand B

	Student	Teacher	School
Prosperity Outcomes			
Educational Attainment	●●●●●●	●●	●●
Health and well-being	●●●●	●	
Student engagement	●		
Foundations for Success			
Inclusive environments	●●●●	●●	●●●●
Quality instruction	●●●	●	●
Learning time	●●●	●	●●
Material resources		●●●●●●	●●●●●
Family and community support	●	●	●●
Demographic factors to assess equity and equality			
Gender	●		
Language spoken at home	●●●●	●●●	
Disability	●		
Immigrant status	●		
Socioeconomic status and poverty	●●●●●●●● ●●●●●●●● ●●		●●
Context factors		●●●●●● ●●●●●●●●	●●●●●● ●●●●
Total	49	33	28

● Trend Questions linked to PISA 2015

● New questions

Student Questionnaire Prosperity Outcomes

9 Core Indicators			
Achievement	1	Reading	Percent Level 2 and higher
	2	Math	Percent Level 2 and higher
	3	Science	Percent Level 2 and higher
Attainment	4	On-track	Percent 'on-track'
Health and Well-Being	5	Life satisfaction	Percent above 5
	6	Anxiety	Prevalence no Anxiety (< 1.5)
	7	Depression	Prevalence not depressed (<1.5)
	8	General health	Percent above 5
Engagement	9	Values schooling outcomes	Percent engaged (SA and Agree)

Educational Attainment

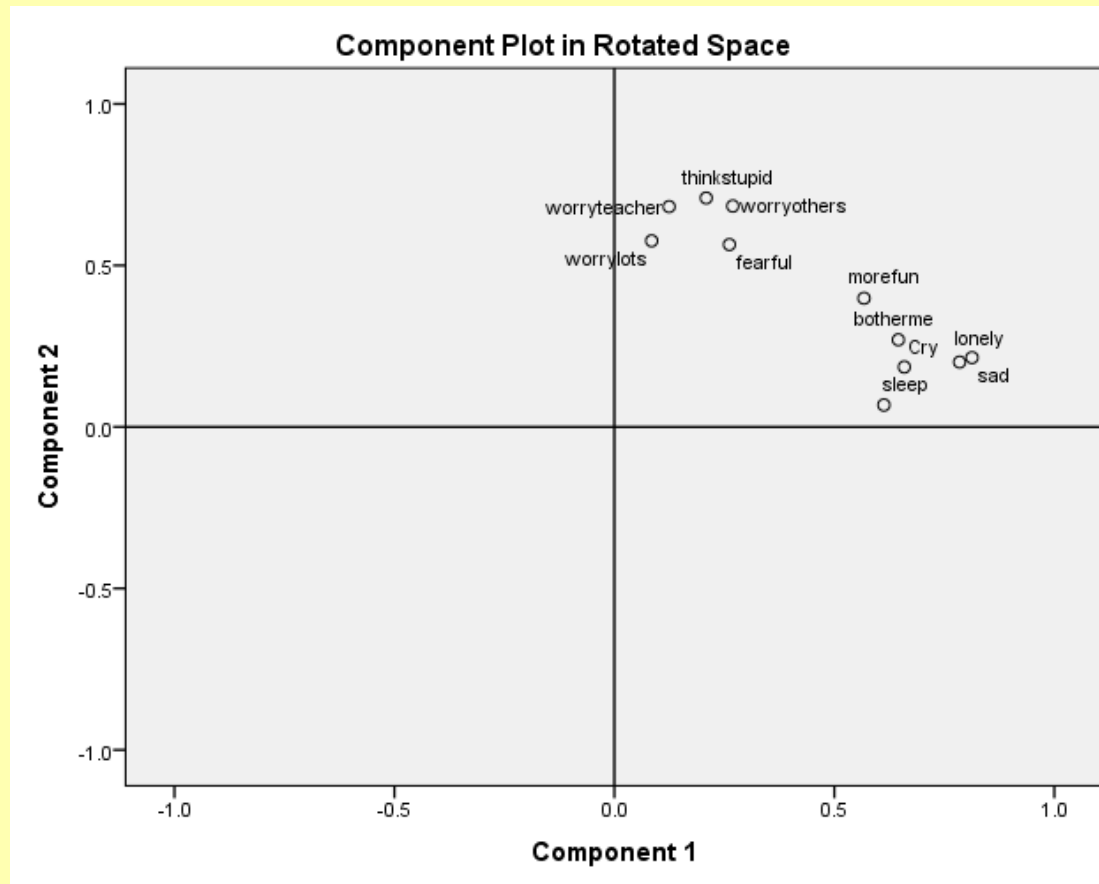
For Strand B, data on students' current grade level and their birthdate were used to construct an ordinal variable describing three levels of attainment:

- **1. On-track.** Students are in their expected grade, given their birthdate; that is, they started school on schedule and have not repeated a grade. In most cases, this would be grade 9 or 10.
- **2. One year below expected grade.** These students have usually repeated a grade or were out of school for a prolonged period. They would typically be in grades 8 or 9.
- **3. Two or three years below expected grade.** In most cases, these students have repeated two or three grades, but some may have started late or simply faded in and out of school for a year. They would typically be in grades 7 or 8.

The tracking form provides more reliable data for this purpose than the student responses for birthdate and grade.

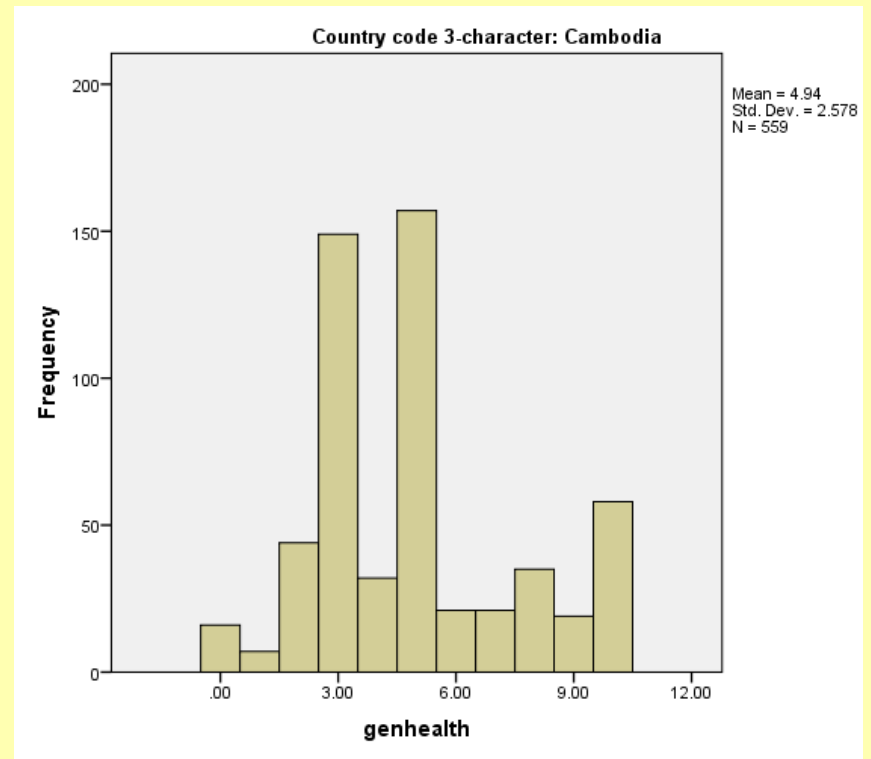
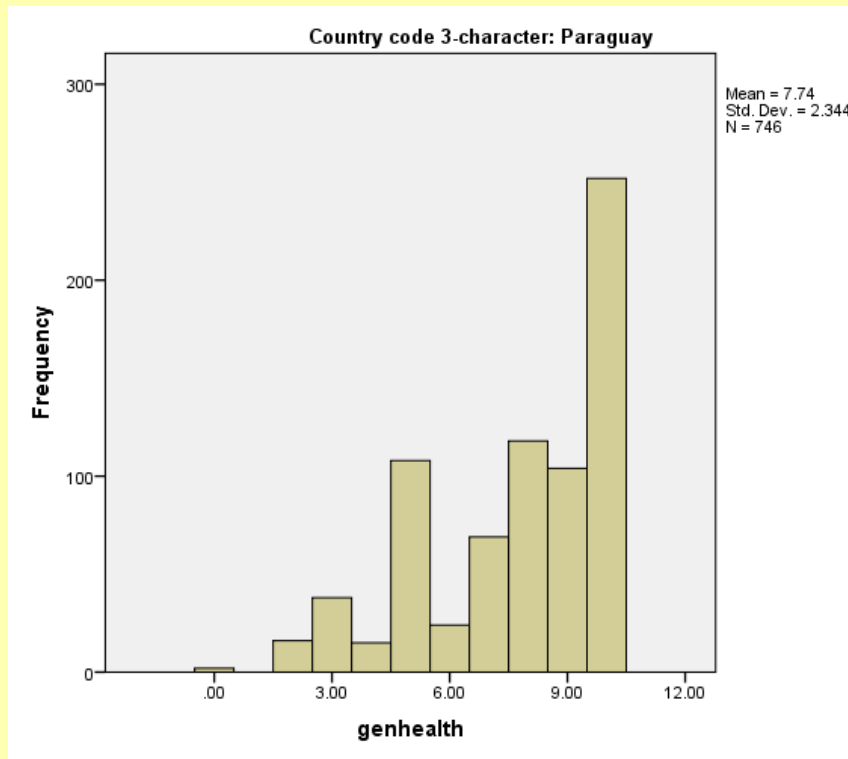
Anxiety and Depression

The items split into two factors in a way that is consistent with the literature.



General Health

Large variation among countries.



Material Resources

The aim of the analyses was to create continuous scale for material resources that could be used to establish four levels of resources. We used the following strategy:

- (1) aggregate each of the items of the student-level measures to the school level, and set a cut-point for each variable, such that there is a set of dichotomous variables.
- (2) aggregate each of items from the teacher-level measures to the school level, and set a cut-point for each variable, such that there is a set of dichotomous variables.
- (3) create a set of dichotomous variables from the school-level measures.
- (4) Conduct IRT analyses of the dichotomous variables.

The analysis revealed that there is a large range of difficulty across the items, and most items have strong discrimination.

At this stage, we recommend dropping the student level measure (ST082), and keeping the full sets of questions for the teacher measures (TQ31 to TQ35) and the school questions (SQ22 to SQ25). These results provide support for a 'proof of concept'. More detailed analyses will be conducted when data are available from all seven countries.

Student Questionnaire Equality Equity Factors

5 Core Indicators				
Gender	1	ST004	Students' Sex	Item (PISA 2015)
Language Spoken at Home	2	ST021	Home language	Item (PISA 2015)
Disability	3	ST016	Physical health in the school setting	Scale
Immigrant Status	4	ST026	Migration Background	Items (PISA 2015)
Poverty and SES	5		Multiple measures	Scale

Socioeconomic Status and Poverty

The aim of the analyses was to:

- (1) Develop and extended measure of home possessions that can be integrated into the main PISA ESCS.
- (2) Compare ESCS using the traditional main PISA approach (i.e., ISCO/HiSei measures, pared, extended homeposs) with an SES variable created using the new closed-format parental occupation questions.
- (3) Compare the measure of books in the home using the PISA 2015 measure (ST065) and the new measure (ST066).
- (4) Develop a measure of poverty that can be used across countries.

Socioeconomic Status and Poverty:

An extended measure of Home Possessions

Question	Label	Difficulty1	Difficulty2	Difficulty3	Discrimination
63	TVs	-1.90	-0.19	0.82	1.97
64	Table to have meals	-1.59			2.35
62	Dictionary	-1.38			2.28
64	Stove	-1.37			2.22
62	Books	-1.36			1.11
63	Smartphones	-1.09	-0.42	0.12	2.18
62	Quiet place	-1.04			1.14
63	Bathrooms	-0.96	0.68	1.58	2.14
64	Fridge	-0.89			3.00
62	Room on your own	-0.85			1.03
62	Desk to study	-0.73			1.29
62	Internet	-0.19			2.83
63	PCs	-0.18	0.84	1.57	2.62
64	Washer	-0.16			2.26
62	PC	-0.13			2.88
62	Ref book	-0.00			0.97
62	Art books	0.06			1.18
63	Cars	0.07	1.13	2.01	1.87
62	Poetry books	0.20			0.53
63	Music instruments	0.47	1.76	2.73	1.14
62	Art work	0.48			0.98
62	Software	0.73			1.69
62	Classic Literature	0.92			1.24

Socioeconomic Status and Poverty: An extended measure of Home Possessions

ST062 Which of the following are in your home?

(added64)

(Please tick one box in each row.)

		Yes	No
ST062Q01TA	A desk to study at	<input type="checkbox"/>	<input type="checkbox"/>
ST062Q02TA	A room of your own	<input type="checkbox"/>	<input type="checkbox"/>
ST062Q03TA	A quiet place to study	<input type="checkbox"/>	<input type="checkbox"/>
ST062Q04TA	A computer you can use for school work	<input type="checkbox"/>	<input type="checkbox"/>
ST062Q05TA	Educational software	<input type="checkbox"/>	<input type="checkbox"/>
ST062Q06TA	A link to the internet	<input type="checkbox"/>	<input type="checkbox"/>
ST062Q10TA	Books to help with your school work	<input type="checkbox"/>	<input type="checkbox"/>
ST062Q12TA	A dictionary	<input type="checkbox"/>	<input type="checkbox"/>
ST064Q01NA	A table to have meals	<input type="checkbox"/>	<input type="checkbox"/>
ST064Q03NA	A washer	<input type="checkbox"/>	<input type="checkbox"/>
ST064Q04NA	A refrigerator or freezer	<input type="checkbox"/>	<input type="checkbox"/>
ST064Q06NA	A stove or burner for cooking	<input type="checkbox"/>	<input type="checkbox"/>

Socioeconomic Status and Poverty:

Comparison of ESCS based on open- and closed-format questions for ISCO

We constructed an ESCS measure using the same approach as in the main PISA.

We scaled mother's and father's occupations for the closes items using two approaches – one based on levels of home possessions and another based on SEI values. These are called SES1 and SES2.

ESCS, SES1 and SES2 are highly correlated at the student and school levels:

Student Level			
	ESCS	SES1	SES2
ESCS	1	.951	.956
SES1	.951	1	.987
SES2	.956	.987	1

School Level			
	ESCS	SES1	SES2
ESCS	1	.953	.954
SES1	.953	1	.996
SES2	.954	.996	1

As this would be a major change to the traditional PISA approach we decided to keep both the open-format ISCO questions and the new closed-format questions for the Main Study

Socioeconomic Status and Poverty:

A new Measure of Poverty for Low- and Middle-Income Countries

We conducted the following analyses:

(1) Developed derived dichotomous or ordinal variables for the following measures:

Private WC

Lighting with electricity

Food security

Piped water

Flush toilet

Cook with gas or electricity

Bank account

Finished flooring

Uncrowded living space

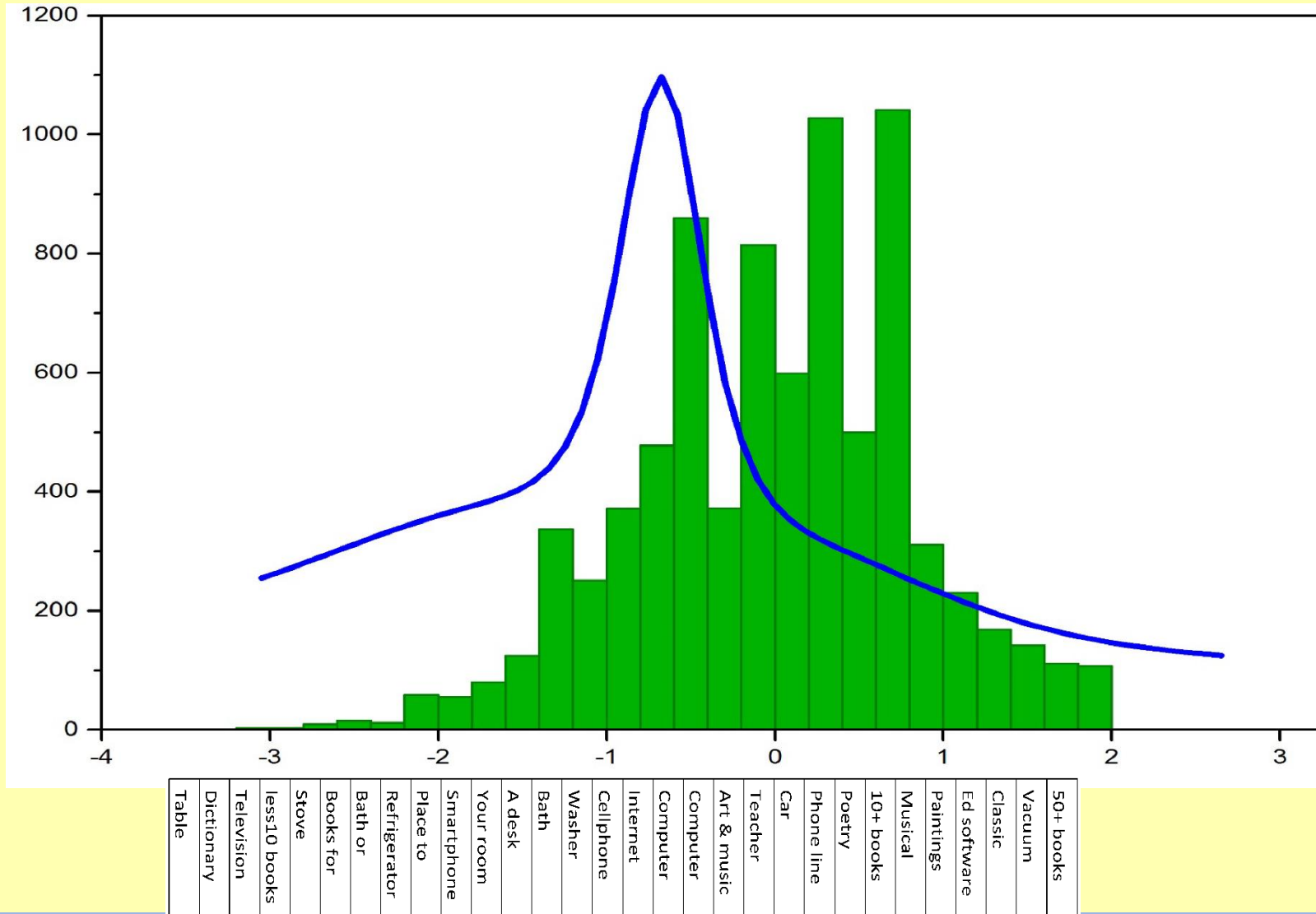
(2) Combined these variables with dichotomous variables used in the extended measure of SES.

(3) Conducted DIF tests for selected items.

(4) Examined the distribution of item theta values versus the population theta values.

Socioeconomic Status and Poverty:

A new Measure of Poverty for Low- and Middle-Income Countries



Supporting Content

A number of questions were included in the questionnaire that were intended to provide supporting content for the core measures of Prosperity Outcomes, Foundations for Success, and the Equality-Equity factors.

The majority of questions that were deleted provided reliable data but needed to be deleted to satisfy space limitations.

Summary

- (1) The Student Questionnaire is of comparable length to the 2009 PISA Questionnaire.

We are confident that we have reliable measures for the core indicators and a strong set of measures for supporting content.

The analyses yielded a reliable measure for the extended measure of home possessions and a new measure of poverty.

- (2) Three items were deleted from the Teacher Questionnaire, mainly due to considerations of length.
- (3) The School Questionnaire is unchanged.