IEA's LaNA and Rosetta Stone: Extending the Reach of TIMSS and PIRLS

Michael O. Martin, Matthias von Davier

& Ina V. S. Mullis

62nd IEA General Assembly

October 2021

IEA TIMSS & PIRLS International Study Center
Lynch School of Education
BOSTON COLLEGE

Extending TIMSS and PIRLS

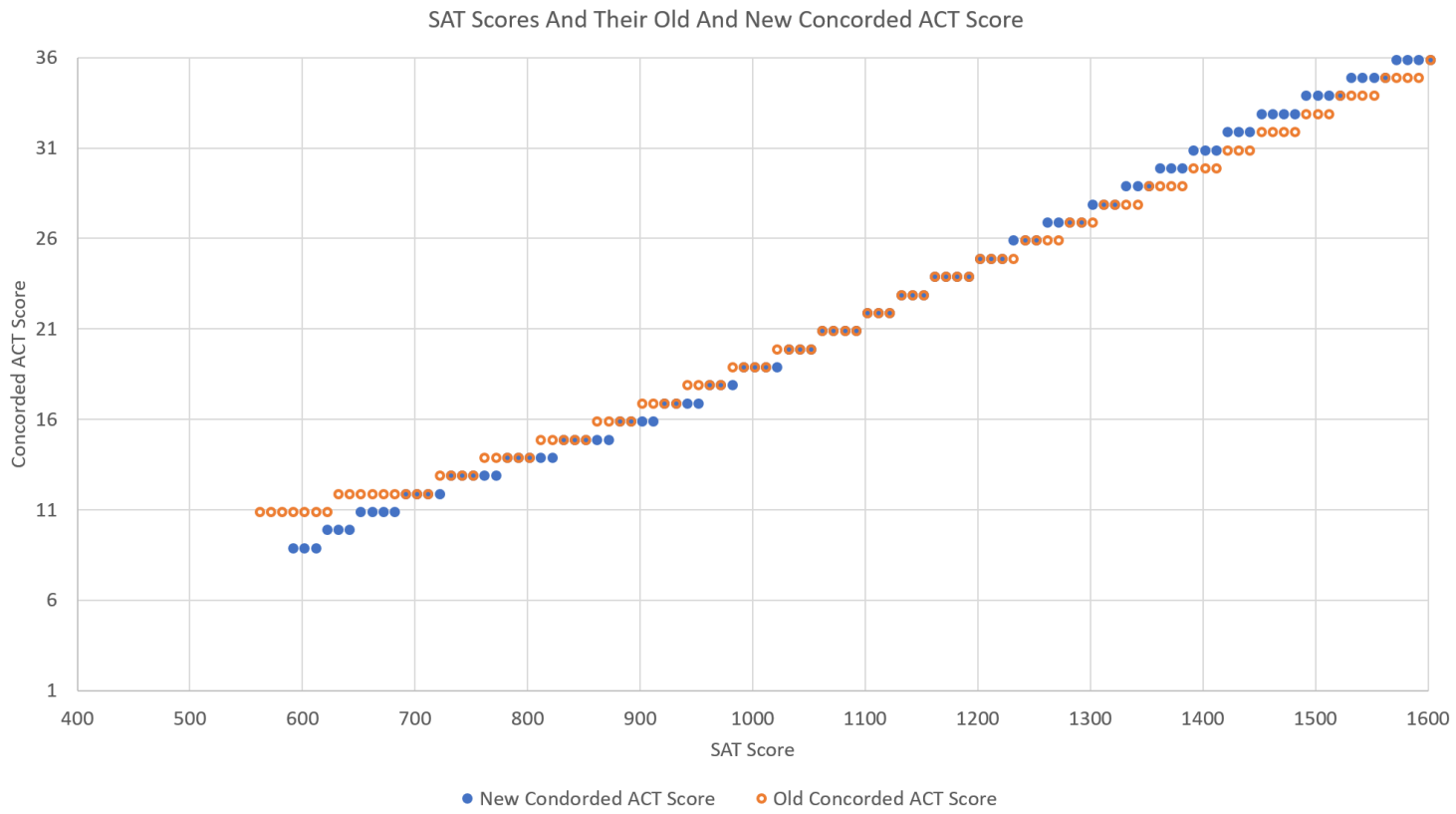
TIMSS and PIRLS achievement scales and benchmarks long established global indicators of student achievement

Two distinct efforts to build on TIMSS and PIRLS:

- LaNA: a new, less difficult Literacy and Numeracy assessment for countries where TIMSS and PIRLS are too difficult
- Rosetta Stone: linking TIMSS and PIRLS to regional assessments – ERCE, PASEC, SACMEQ, SEA-PLM, PILNA
- Address progress toward UN Sustainable Development Goal 4.6
  - Ensure inclusive and quality education for all and promote lifelong learning.
- Linking through concordance tables – projecting TIMSS and PIRLS scores from scores on other assessments

Example Concordance: ACT and SAT

- LaNA and Rosetta Stone concordance tables will be similar to the ACT – SAT example
- Will provide a range of likely TIMSS/PIRLS scores for each LaNA or regional assessment score
  - Account for measurement error



ACT Composite to SAT Total

This table shows a range of SAT scores for each ACT score

ACT	SAT	SAT Range
36	1590	1570-1600
35	1540	1530-1560
34	1500	1490-1520
33	1460	1450-1480
32	1430	1420-1440
31	1400	1390-1410
30	1370	1360-1380
29	1340	1330-1350
28	1310	1300-1320
27	1280	1260-1290
26	1240	1230-1250
25	1210	1200-1220
24	1180	1160-1190
23	1140	1130-1150
22	1110	1100-1120
21	1080	1060-1090
20	1040	1030-1050
19	1010	990-1020
18	970	960-980
17	930	920-950
16	890	880-910
15	850	830-870
14	800	780-820
13	760	730-770
12	710	690-720
11	670	650-680
10	630	620-640
9	590	590-610

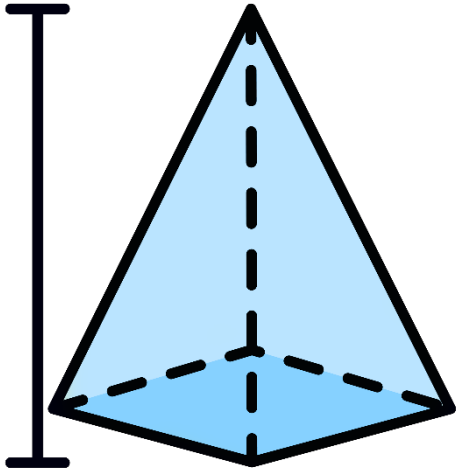
Overview of LaNA

- IEA's new less difficult Literacy and Numeracy assessment, with links to TIMSS and PIRLS
  - *Addresses the same constructs, with shorter, easier items*
- Current LaNA instrument the result of several waves of piloting – now ready for full scale implementation
- IEA seeking 5-10 countries to participate in launch of LaNA
  - Establish the LaNA scale metric
  - Construct concordance tables for LaNA vs TIMSS and PIRLS
  - Establish new "Basic" benchmark, below TIMSS/PIRLS Low International benchmark

LaNA Assessment Framework and Items

Numeracy – based on TIMSS mathematics 4th grade

- Number, Measurement & Geometry, Data
- Mostly Knowing, Applying, some reasoning
- TIMSS items too difficult, so all LaNA items newly developed



Literacy – based on PIRLS reading 4th grade

- Literary and Informational purposes for reading
- Mostly retrieval and straightforward inferencing, some integrating
- All passages and items adapted from PIRLS Literacy – shorter and easier

Extensive Piloting of LaNA Prototypes

- Early versions (LaNA 1.0 and 2.0) piloted in Haiti and Pakistan (2016, 2017)
- Large scale pilots (LaNA 3.0) in Serbia, North Macedonia, Nigeria, and Haiti (2019, 2020)

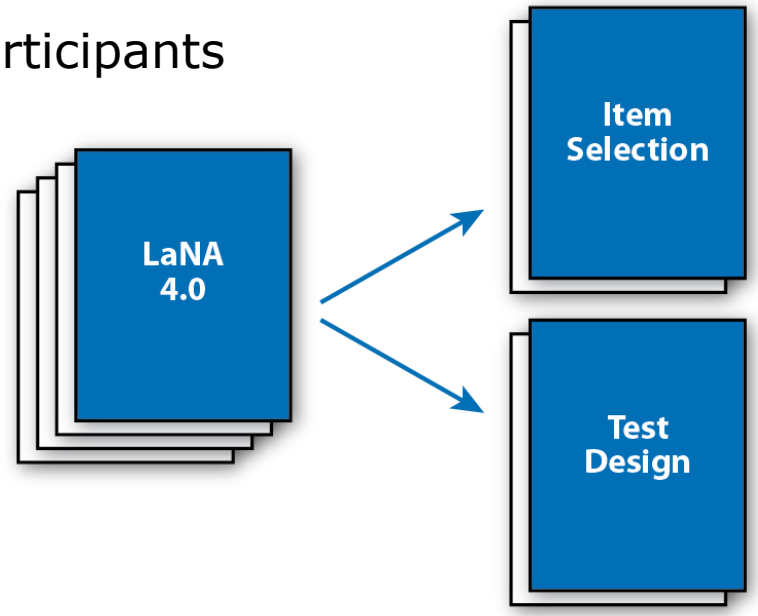
	Countries	N (used for analysis)
LaNA Pilot	Serbia (Grade 3)	1,295
	North Macedonia (Grade 3)	1,196
	Nigeria (Grade 4)	884
	Haiti (Grade 4)	952
	Haiti (Grade 6)	706

LaNA Pilot Data Collection

- Four LaNA booklets
  - Five blocks of Numeracy items (one common)
  - Five literacy passages and items (one common)
- One linking booklet
  - Easy items from TIMSS 2015/2019 and PIRLS 2016
- Data collection design in pilot
  - Two-day model: LaNA and linking booklets on successive days, each student completes one of each

LaNA Pilot Data Analyses and Design

- Extensive analyses of pilot data using classical item anayses and IRT scaling
  - Comparability of the measured constructs across instruments
  - Comparability of the difficulty targets of LaNA and TIMSS/PIRLS
  - Reliability and construct coverage of the constructed scales
  - Fit of TIMSS and PIRLS item parameters for LaNA participants
- LaNA 4.0 assessment instrument finalized based on pilot results
  - Item selection and test design
- LaNA linking design strengthened and finalized



LaNA Booklet Design

- Four blocks Numeracy Items (N1-N4), chained across booklets
- Five Literacy passages (L0, L1-L4), common + unique

LaNA Booklets	Part 1 (40 min)		Part 2 (40 min)	
1	N1	N2	L0_Common	L1
2	L0_Common	L2	N2	N3
3	N3	N4	L0_Common	L3
4	L0_Common	L4	N4	N1

LaNA Linking Booklet Design

- Four blocks less difficult TIMSS Items (NL1-NL4), chained across booklets
- Four PIRLS Literacy passages (LL1-LL4), also chained across booklets

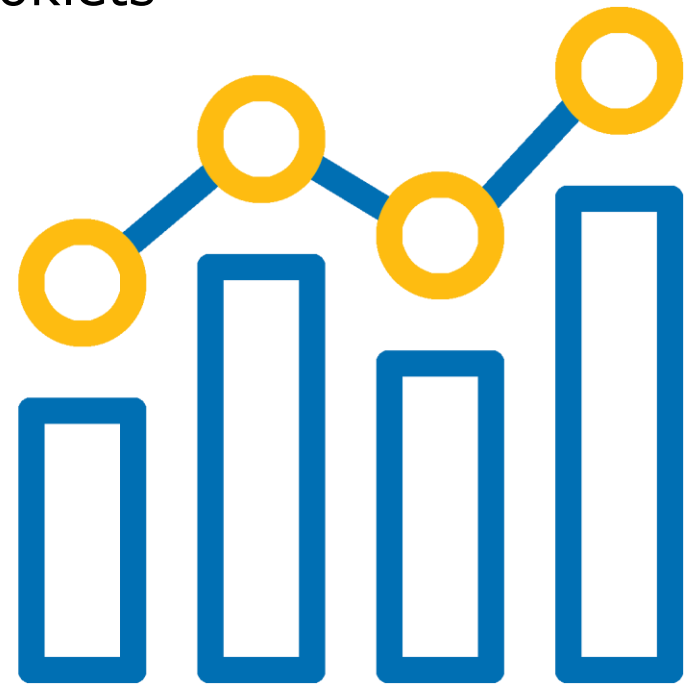
Linking Booklets	Part 1 (40 min)		Part 2 (40 min)	
1	NL1	NL2	LL1	
2	LL2		NL2	NL3
3	NL3	NL4	LL3	
4	LL4		NL4	NL1

LaNA Implementation

- The final LaNA instrument together with the linking booklets will be administered to 5-10 seed countries
- Instruments, administration manuals and scoring guides have been shared with IEA
  - French translations underway
- Senegal the first participating country
  - November 2021

LaNA Implementation (cont.)

- Data collection design
  - One-day model: all 8 (LaNA and linking) booklets distributed in each class
- Data from the seed countries will be used to
  - Establish LaNA scale and scale metric
  - Create LaNA-TIMSS/PIRLS concordance tables
  - Create new "Basic" International Benchmark



Benefits for LaNA Countries

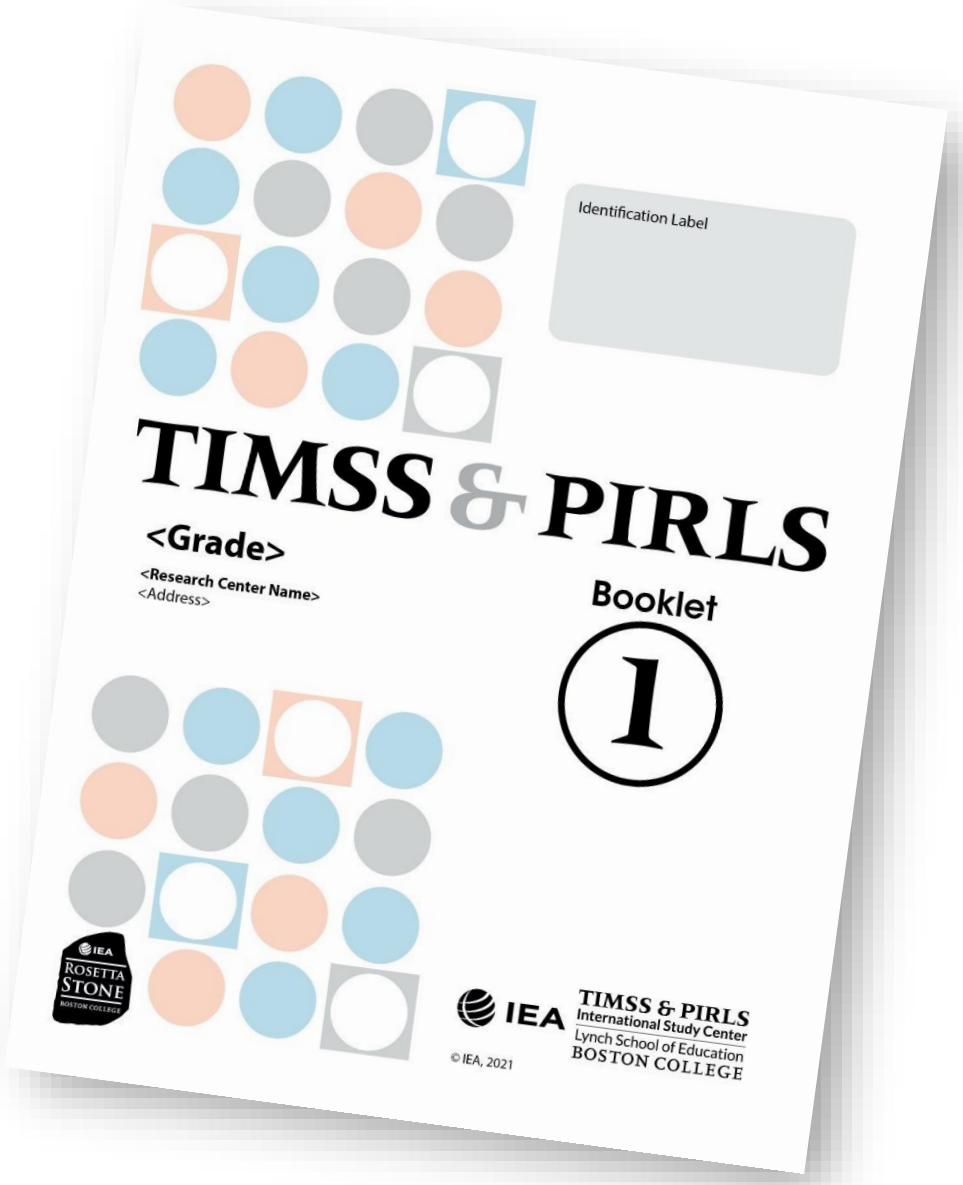
- Reliable, valid, and internationally comparable indicators of student literacy and numeracy
  - Ideal for measuring progress toward SDG 4
- Concordance tables provide LaNA countries' projected scores on TIMSS and PIRLS
  - Also estimates of percentages of students reaching TIMSS and PIRLS International Benchmarks
- Progress toward suitability for participation in TIMSS and PIRLS

Overview of Rosetta Stone

- Objective: "Translate" or link to TIMSS/PIRLS from regional assessments – ERCE, PASEC, SACMEQ, SEA-PLM, PILNA
- Funded by UNESCO, representative countries from two assessments
  - ERCE: Guatemala, Colombia (Chile out due to COVID-19)
  - PASEC: Burundi, Guinea, Senegal
- TIMSS/PIRLS items administered to students participating in ERCE or PASEC
  - Provides psychometric link to establish concordance
- Can report "likely" performance on TIMSS/PIRLS based on ERCE or PASEC results

Rosetta Stone Data Collection

- Students participating in the ERCE or PASEC assessments at sixth grade also were administered TIMSS and PIRLS booklets in a separate session
  - Equivalent sample design
- Eight linking booklets
  - For ERCE, medium and easy items from TIMSS 2015/2019 and PIRLS 2016
  - For PASEC, all easy items from TIMSS 2015/2019 and PIRLS 2016



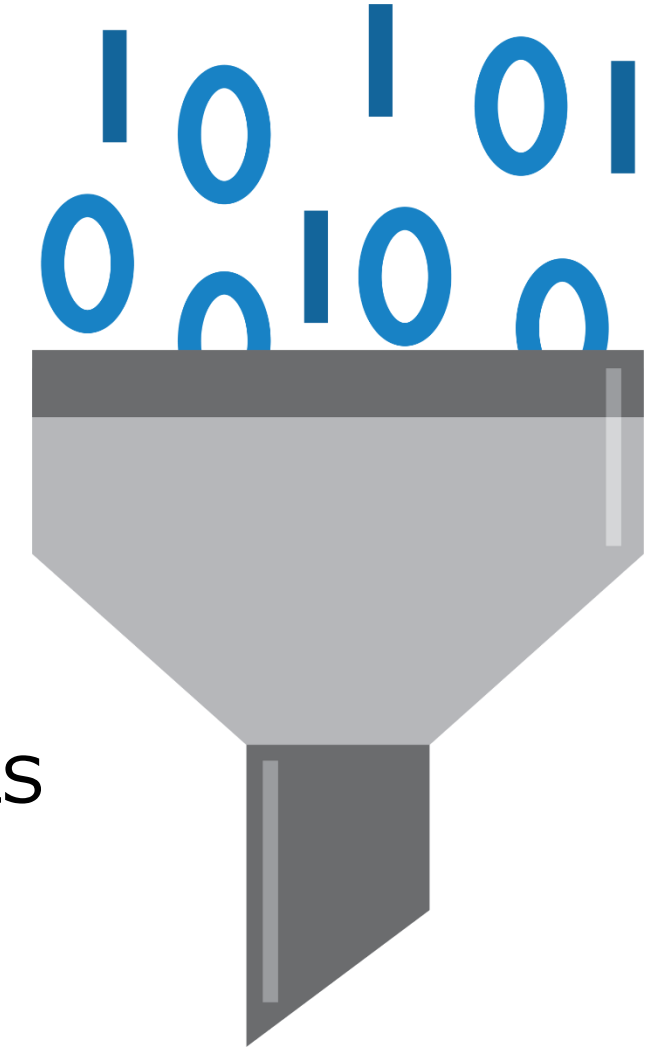
Rosetta Stone Project Status

- Successful administration of Rosetta Stone in 2020 (in spite of delays due to Covid-19)
- Data received from five countries and processed for analysis

	Countries	N (used for analysis)
ERCE	Guatemala (Grade 6)	3,144
	Colombia (Grade 6)	5,340
PASEC	Burundi (Grade 6)	2,271
	Guinea (Grade 6)	2,207
	Senegal (Grade 6)	2,059

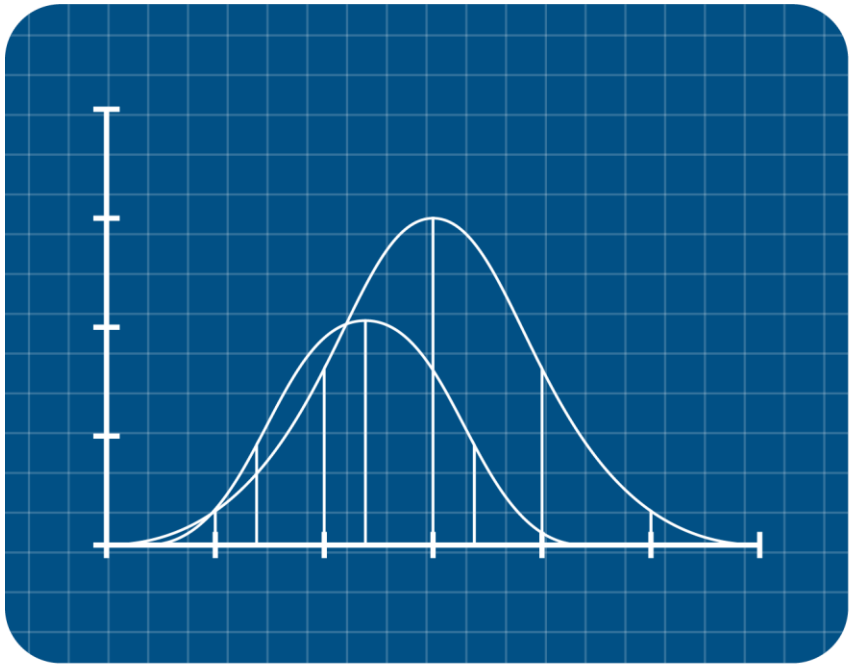
Data Quality and Analysis

- Rosetta Stone requires careful analysis and review of the data in order to determine:
  - The quality of the data base for comparison between national and international assessments
  - The psychometric quality of the items
  - The measurement accuracy (measurement error)
- This is done by applying standard TIMSS/PIRLS data quality control procedures, statistical analysis and psychometric models



Psychometric Analyses

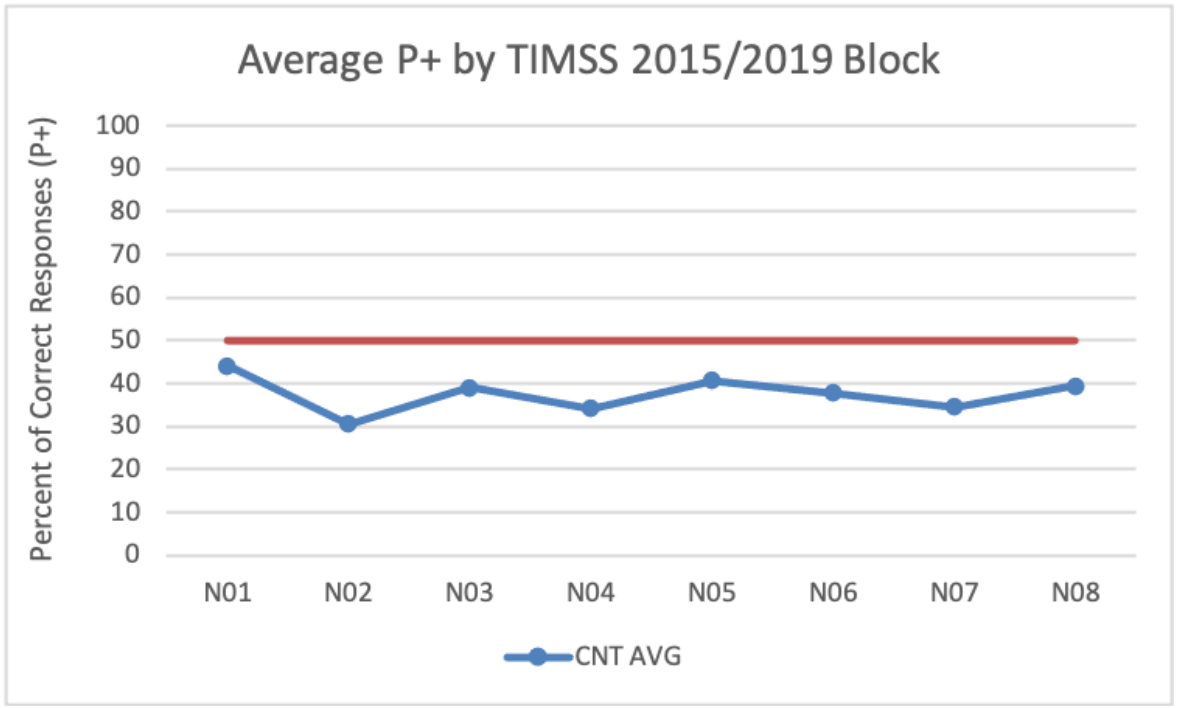
- Establishing a concordance between the ERCE/PASEC data and the TIMSS/PIRLS international scales/benchmarks requires the following Psychometric analysis step:
  - Establishing comparability through IRT scaling
  - Applying TIMSS/PIRLS item parameters to ERCE/PASEC data
  - Producing plausible values through population modeling
  - Validating and replicating the plausible values received in the ERCE/PASEC data sets
  - Building a concordance table based on derived plausible values

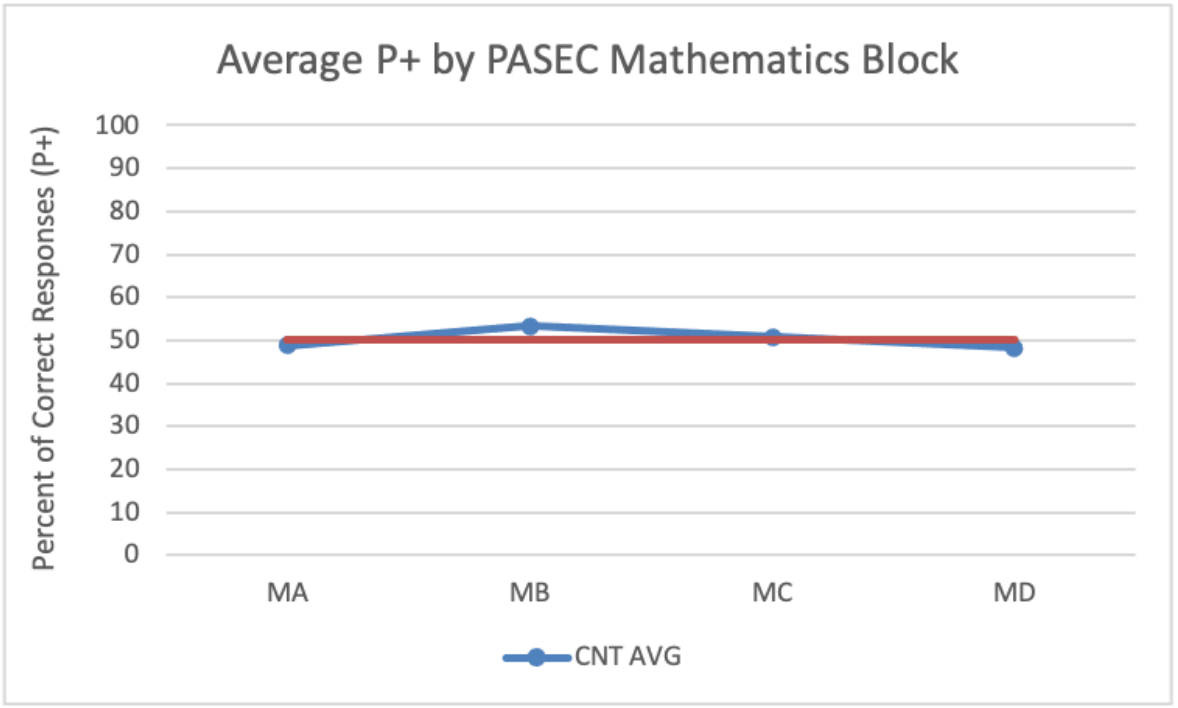


Rosetta Stone Project Status - PASEC

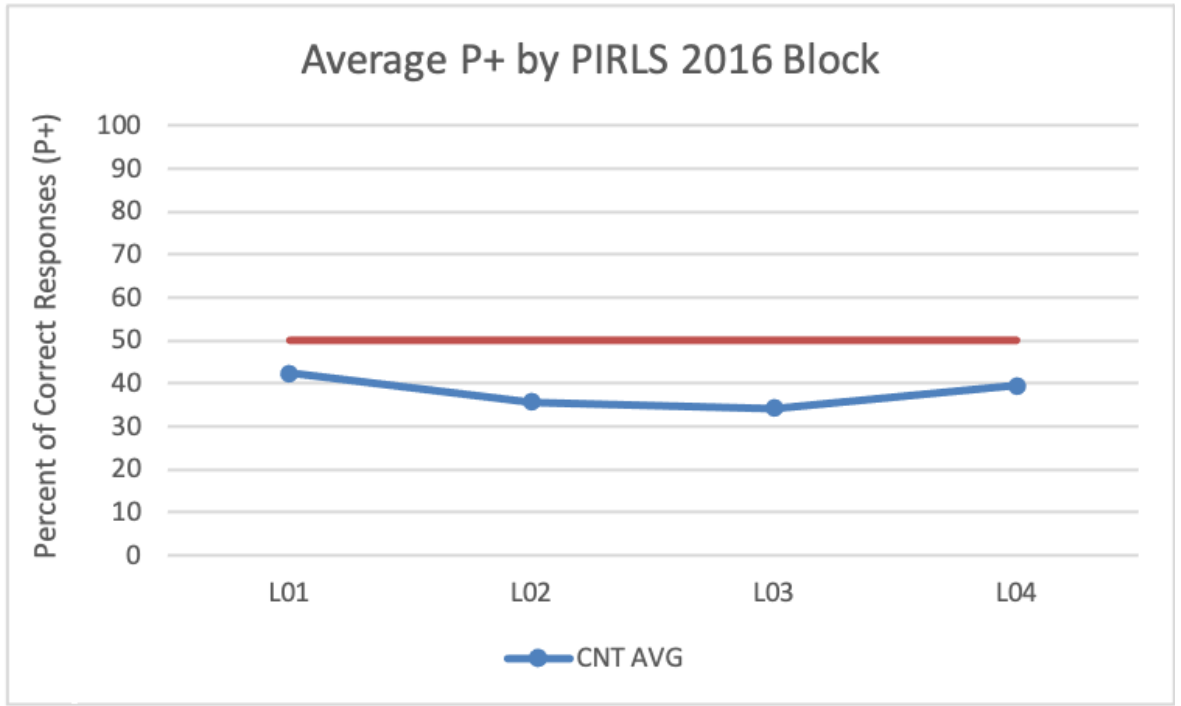
- Final data from Burundi, Guinea, and Senegal at TIMSS & PIRLS ISC, May 2021
  - Quality checks completed and almanacs produced, June 2021
- As expected, TIMSS and PIRLS item blocks are more difficult than PASEC Mathematics and Reading blocks
  - However, the difficulty level of the TIMSS and PIRLS items seems appropriate for Rosetta Stone analyses
- TIMSS and PIRLS item parameters show a good fit for the majority of link items
- Draft results of psychometric analyses presented to PASEC countries, September 23, 2021

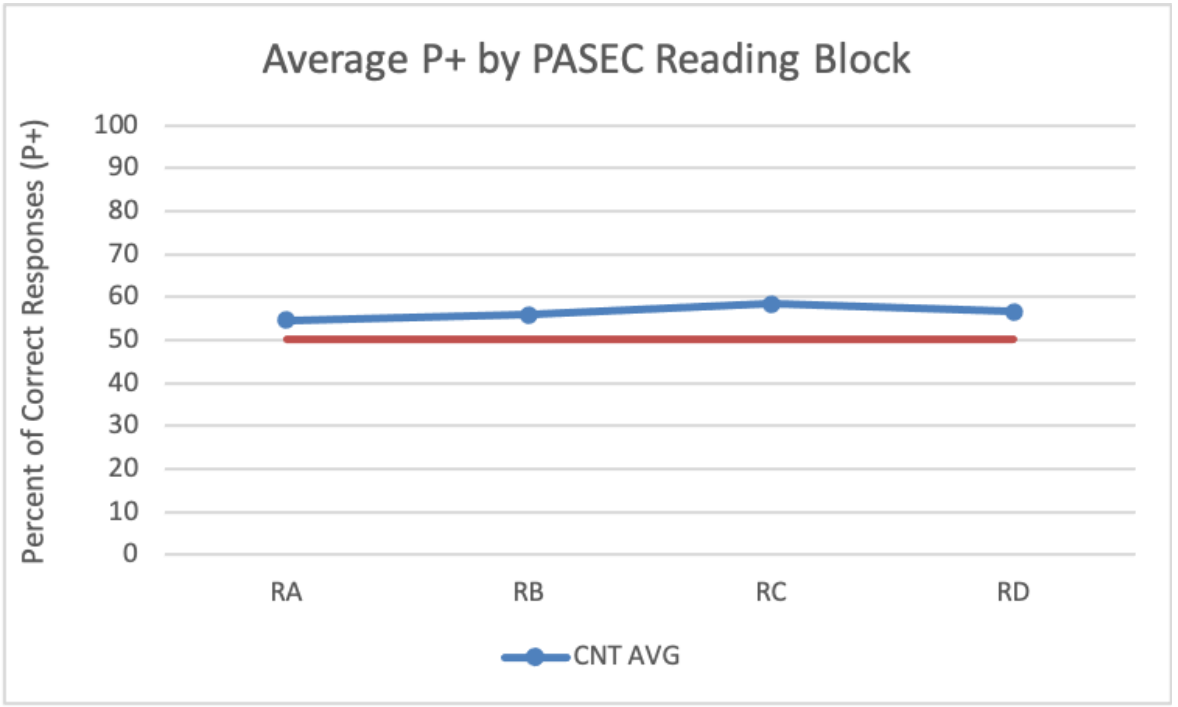
Average P+ by Item Block
TIMSS vs. PASEC Mathematics





Average P+ by Item Block
PIRLS vs. PASEC Reading





Rosetta Stone Status - ERCE

- Some delays due to COVID-19
- Classical item statistics and IRT scaling conducted on preliminary data
- Similar to PASEC, TIMSS and PIRLS items appear to be more difficult than ERCE Mathematics and Reading blocks
- Again, however, the difficulty level of the TIMSS and PIRLS blocks seems appropriate for the Rosetta Stone analyses
- Analysis will be finalized once the updated/final data files have been received

Reporting

- Concordance tables will be provided to translate between ERCE/PASEC and TIMSS/PIRLS achievement scales
  - Confidence intervals will be reported to account for measurement and linking error
- Using the concordance tables, it will be possible to
  - Estimate the expected TIMSS or PIRLS score corresponding to each ERCE/PASEC score
  - Determine the ERCE/PASEC scores corresponding to TIMSS/PIRLS International Benchmarks
  - Estimate the percent of students in each country that would have reached the TIMSS and PIRLS benchmarks
- Technical documentation of all results will be provided

IEA's LaNA and Rosetta Stone: Extending the Reach of TIMSS and PIRLS

Michael O. Martin, Matthias von Davier

& Ina V. S. Mullis

62nd IEA General Assembly

October 2021

IEA TIMSS & PIRLS International Study Center
Lynch School of Education
BOSTON COLLEGE