



NATIONAL RESEARCH
UNIVERSITY

SETTING BENCHMARKS OF CHILDREN READING DEVELOPMENT FOR POTENTIAL COMPARISONS ACROSS DIFFERENT COUNTRIES AND CULTURES

Alina Ivanova, Elena Kardanova

National Research University Higher School of Economics, Moscow, Russia

Seventh International Conference
on Probabilistic Models for Measurement
Perth, Australia, 2018

Overview

- ▶ Background and purpose of the study
- ▶ Method and analytical approach
- ▶ Preliminary results
- ▶ Conclusions

Challenges in international reading assessment in primary school

Challenges of young children assessment (Merrell, 2017)

- Language acquisition
- Maturity to take part in assessment
- Reading ability
- Short concentration span

Challenges of ILSA in reading:

- Hard but possible to compare reading internationally when children *can already read* (PIRLS)
- Language differences greatly affect the literacy of students when students *only learn to read* (Ercikan, Roth, Asil, 2015)

The current study

What are the possibilities for international comparisons of what children know and can do at the start of schooling?
In particular, in reading.

Russian education: primary level

**Around 45,000 preschool institutions
across Russia**

Preschool is not compulsory

The majority of children do attend



1,500 000 first graders in Russia

1st of September – start of school year

Children start school at the age of 7

Sample

Place	N of schools	N of classes	N of students	Girls
Moscow	16	140	3173	48%
Naberezhny Chelny	41	94	2379	52%
Sevastopol	22	59	1283	49%
Tambov	5	37	943	49%

In total: 7778 first graders

The iPIPS (international Performance Indicators in Primary School) instrument

Vocabulary

Phonological awareness

Reading

Mathematics

- Development of PIPS (Performance Indicators in Primary School) in Durham University, UK, since 1994
- Baseline and Follow-up Assessment in the first year of schooling:
 - Measurement of individual progress
- High Reliability: Internal and test/re-test
- Links to later attainment: followed 45,000 children up to age 16 in England
- Translated into German, Chinese, Serbian, Slovenian, Portuguese, Russian and other languages

Tymms, 1999; Tymms, Merrell, Wildy, 2015

Development of the Russian version of iPIPS: mathematics

Translation

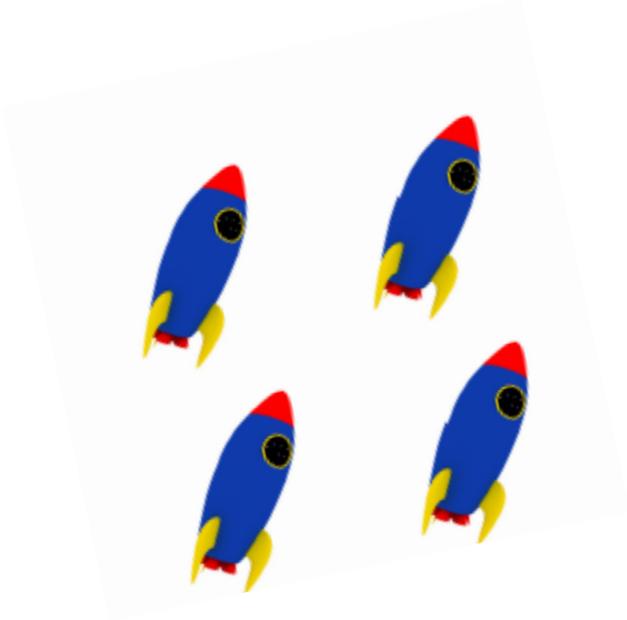
- Two independent translators
- Reconciliation
- Back translation

Selection of items

- Experts' evaluation
- Clinical approbation
- Items correction

Development of additional harder items

Examples of math items



“ Here are four rockets. If you put another three rockets in the picture how many would there be? “

Посмотри на этот список покупок. Сколько всего ты заплатишь за все эти продукты?

1 апельсин	10 рублей
1 яблоко	10 рублей
1 банан	5 рублей

“ Look at this shopping list. How much does the food cost altogether? “

International comparison of iPIPS math results: England, Scotland, Russia

Rasch analysis	
Dimensionality analysis	PCA of standardized residuals
Fit analysis	Unweighted and weighted mean square statistics
DIF analysis	MH method, LR method
Reliability study	Person reliability index, separation index
Linking of measures	Simultaneous calibration

It is possible to construct a common Math scale across the three countries:

- Person reliability 0.94
- Person separation index 3.96

Despite -
different languages
different ages
partly different tests

Development of the Russian version of iPIPS: reading

Russian and English languages have very particular properties that lead to severe limitations in adaptation

- ✓ The language units are longer
- ✓ The number of words with identical and similar graphic and phonetic shells in Russian is many times less than in English
- ✓ No differentiation of the articles and auxiliary verbs in the Russian-language text
- ✓ Free order of words in the Russian sentence and no stable syntactic models of wording

▶ English version

“ This kind of frog only leaves its burrow when it / there / the is raining outside ”

▶ Russian version

“ Лягушка покидает свое жилище только тогда, когда пройдет, уйдет, выйдет дождь ”

To summarize: it is impossible to translate and slightly correct the original test items to create the Russian version of test items

Development of the Russian version of iPIPS: reading

Translation

- Translation of original items into Russian

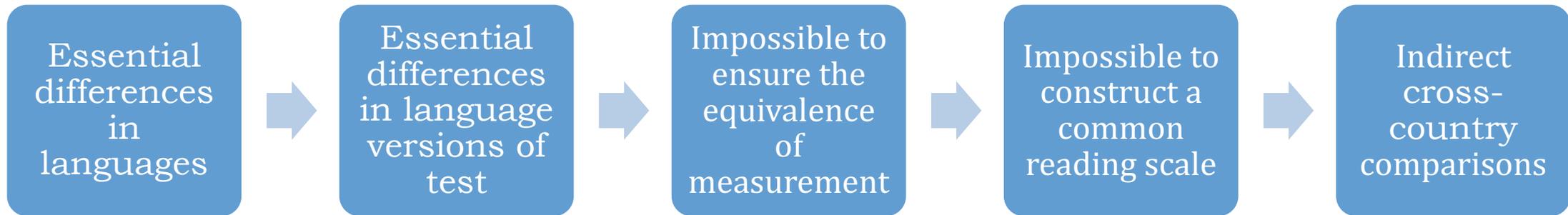
Selection of items

- Expert evaluation

Development of new items based on:

- Theoretical framework of iPIPS
- Structure of original test
- Type of original items

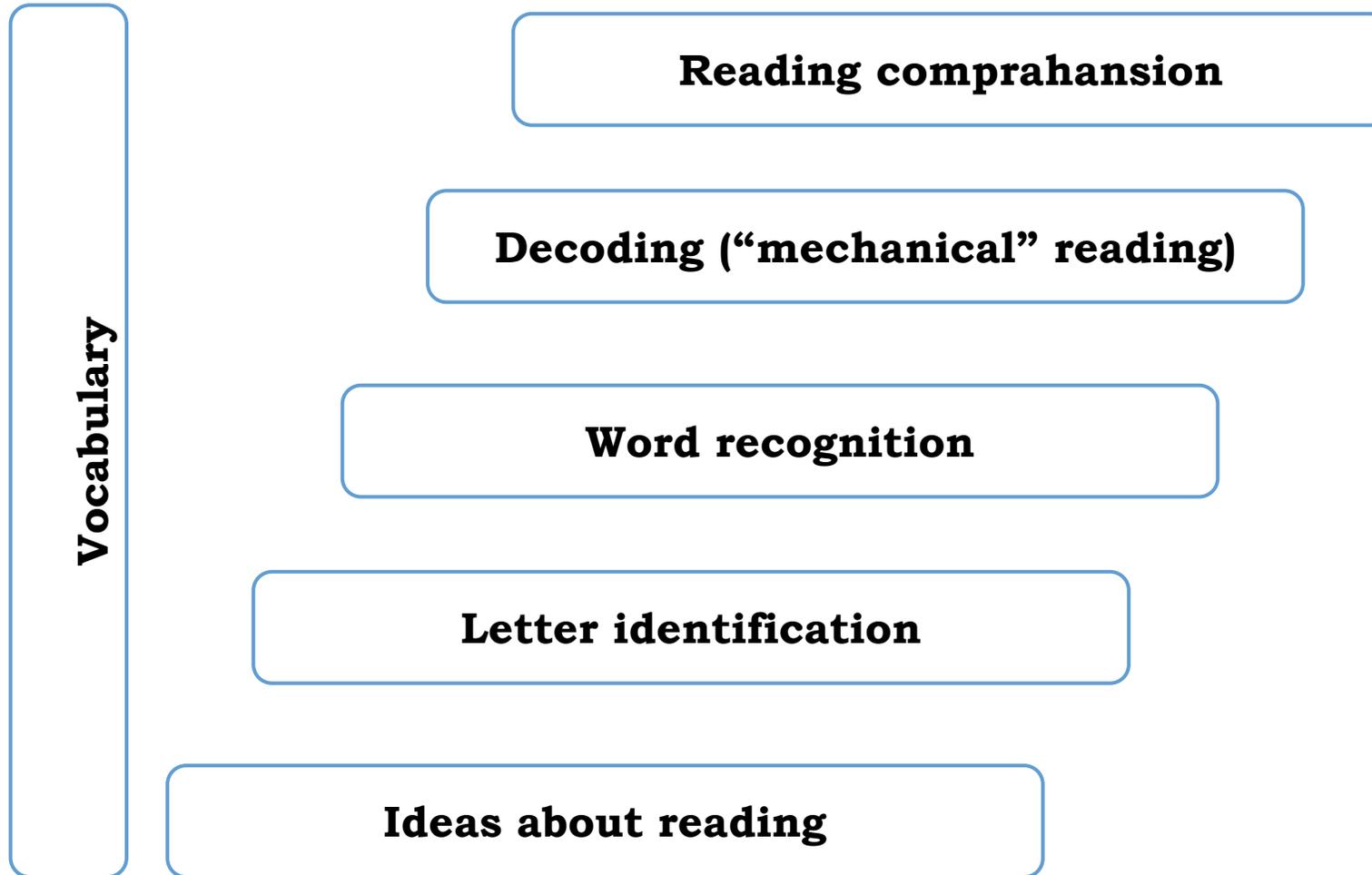
Possibilities for common scale in reading



Our solution: to develop internationally comparable benchmarks based on iPIPS theoretical framework for reading assessment and Rasch approach

We want to extend our understanding of the possibilities to compare the young children reading development at the start of schooling across countries.

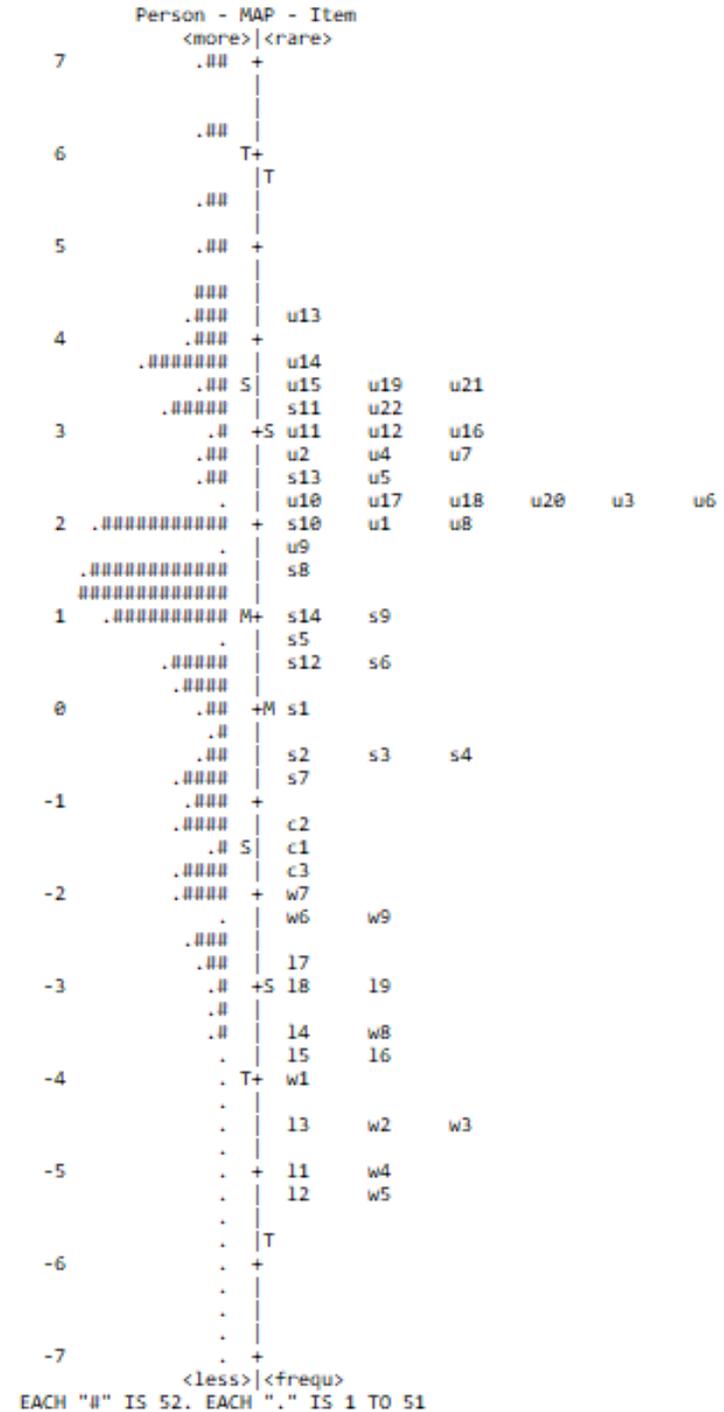
Model* of Reading assessment



*Tymms, 2009

Analysis of reading scale

- ✓ Scale is essentially unidimensional
- ✓ Good overall item fit
- ✓ High Reliability:
 - ✓ Alpha 0.97
 - ✓ Person Reliability 0.94
 - ✓ Separation index 3.94
- ✓ Item distribution meets theoretical expectations

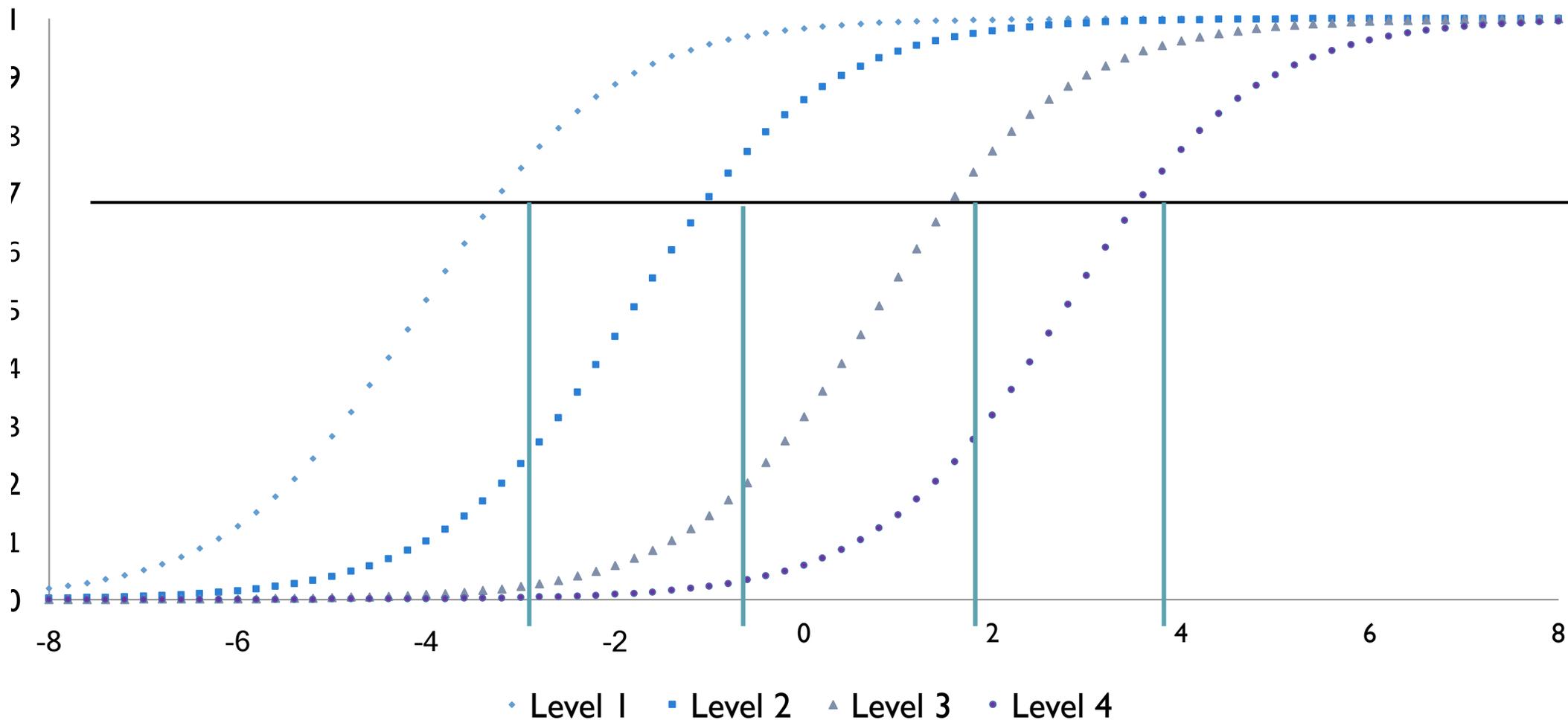


Proficiency levels of reading comprehension

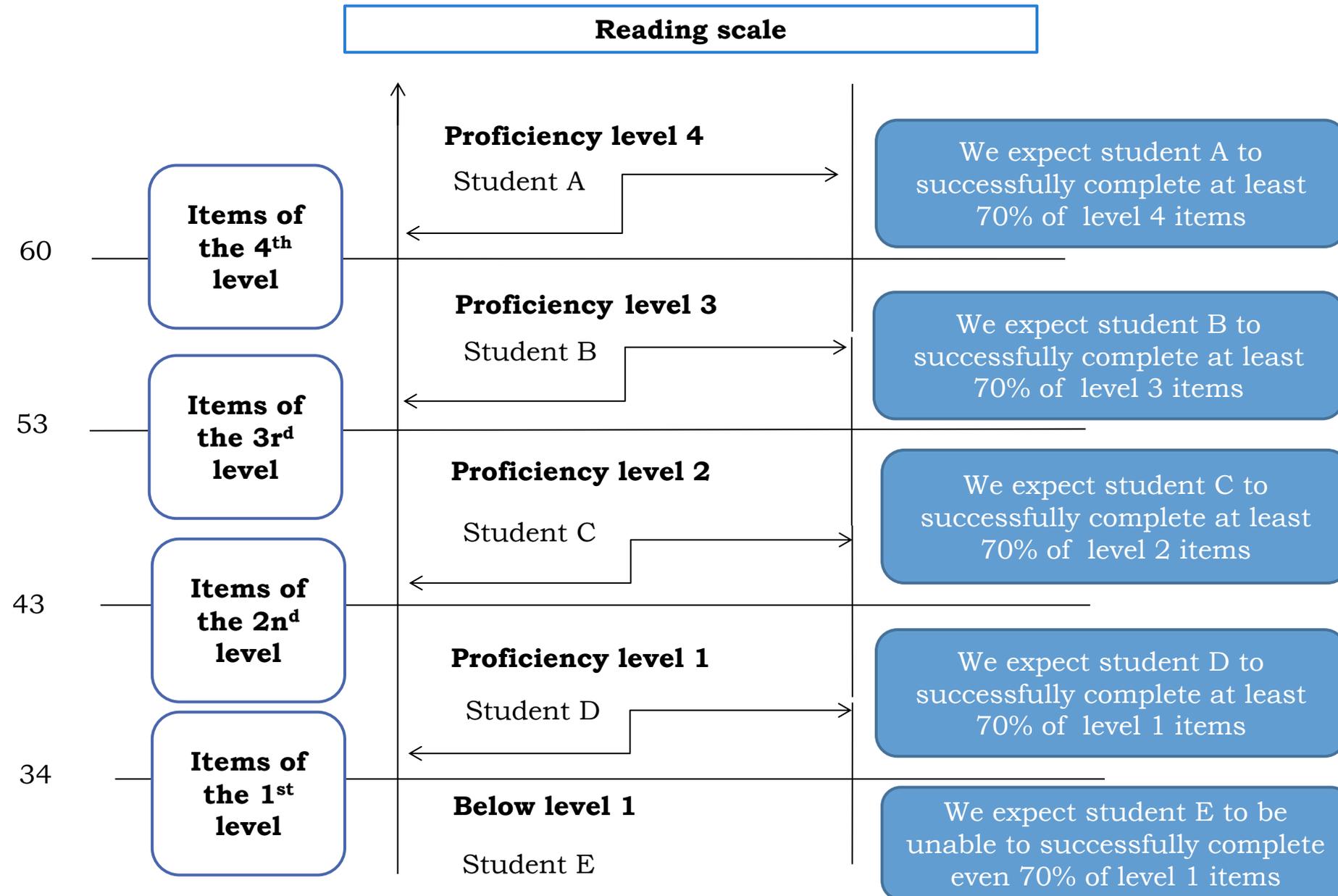


<i>Level</i>	<i>Description</i>
Level 4	Reading and analysis of the text
Level 3	Ability to read longer texts, partial understanding
Level 2	Ability to read short text without deep understanding
Level 1	Recognition of letters and high-frequency words
Below level 1	Very basic perception of ideas about reading

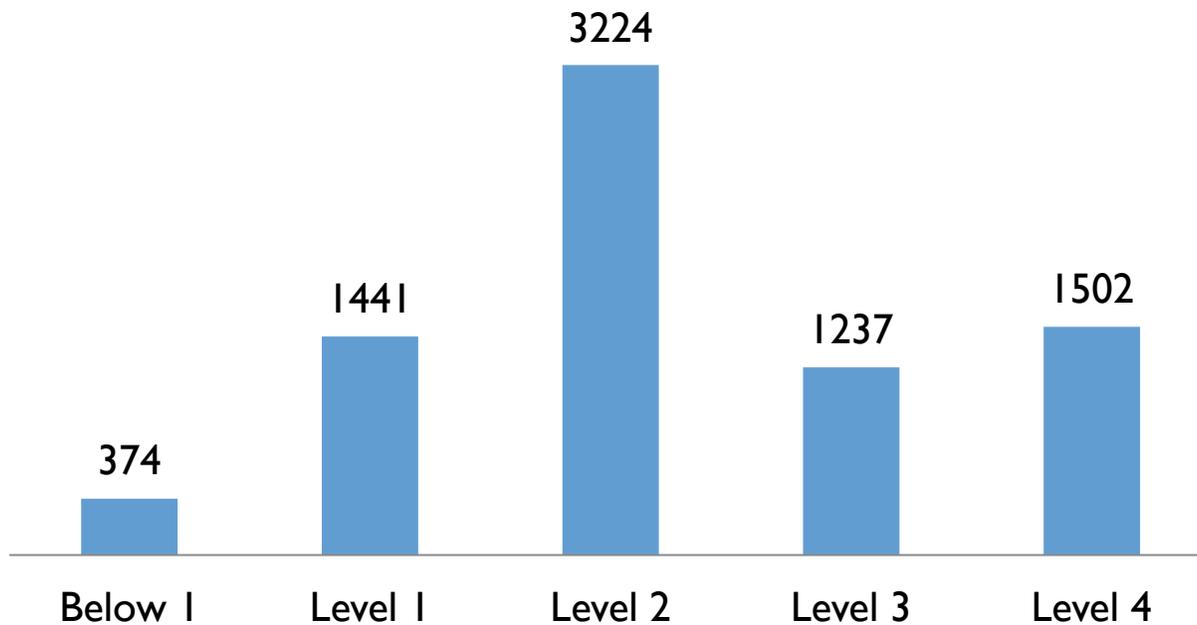
ICC for benchmarks



Proficiency levels of reading comprehension



Results



- ▶ Students are represented at different levels
- ▶ Distribution of students generally meets our expectations

Conclusions

- ▶ It is possible to use the instrument's theoretical framework for reasonable benchmarks setting
- ▶ These benchmarks allow us to see what children know and can do in reading when they start school
- ▶ The approach can be applied to measure reading progression of students for the first year of schooling
- ▶ The benchmarks can be used as reference for indirect comparisons of children reading among countries

Limitations

- ▶ The approach is illustrated on the Russian data only
- ▶ The indirect nature of our comparisons
 - ▶ Usefulness?
 - ▶ Appropriate?

Thank you