



# DEVELOPMENT AND VALIDATION OF THE RUSSIAN VERSION OF THE IPIPS FIRST GRADERS STUDY

**Alina Ivanova**  
**National Research University**  
**Higher School of Economics**  
**Moscow, Russia**

*ITC 2016*



# Russian education: primary level

45,107 preschool institutions  
across Russia

Not compulsory

The majority of children do attend

1,500 000 first graders in Russia

1st of September – day X

Children start school at the age of 7



The primary education in Russia is regulated by Federal state educational standard. The Standard includes requirements regarding:

- learning outcomes
- core curriculum
- conditions for implementing the core curriculum

# Background

*No standardized LSA at the primary school level in Russia*

Author tests and diagnostics

*Assessment practices throughout the world*

EDI (Early Development Instrument), Canada

Monitoring and Evaluation system for Primary Schools, The Netherlands

PIPS (Performance Indicators in Primary Schools), UK



Writing

Vocabulary

Phonological awareness

- Repeating words
- Rhyming words

Reading

- Ideas about Reading
- Letters
- Words
- Reading short story
- Reading passages

Mathematics

- Counting
- Simple sums
- Digits
- Number manipulations
- Math problems

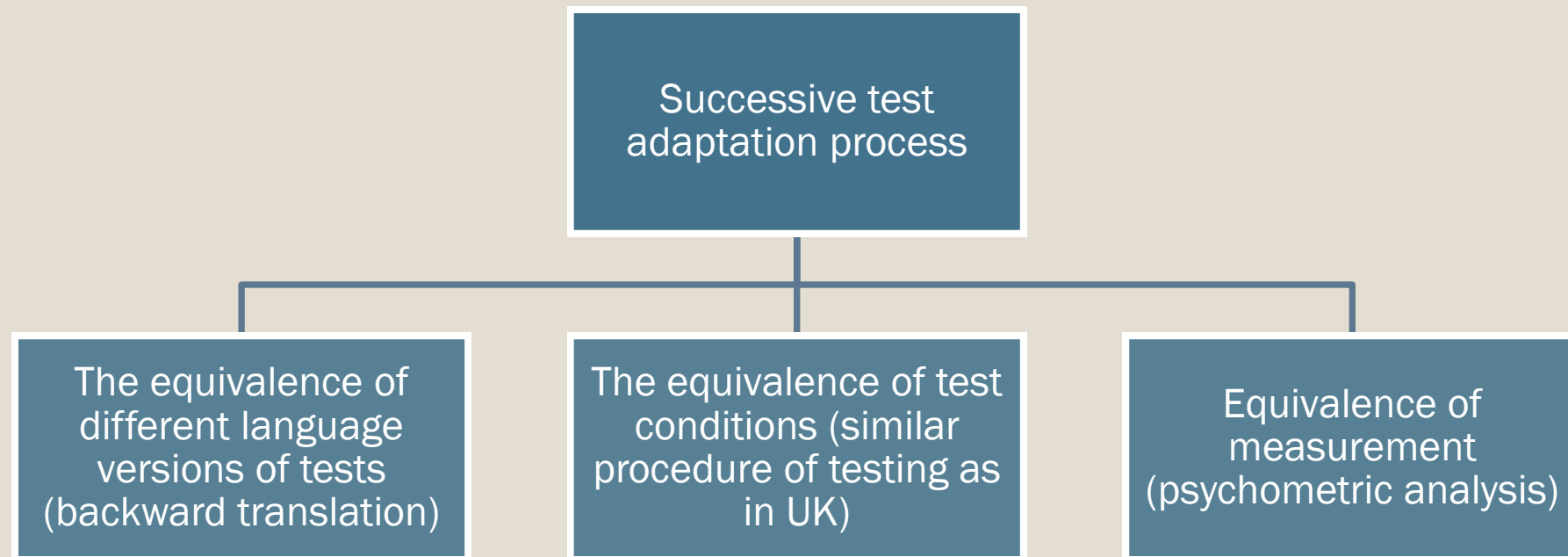
Behavior  
survey

PSED survey

The Russian  
iPIPS version

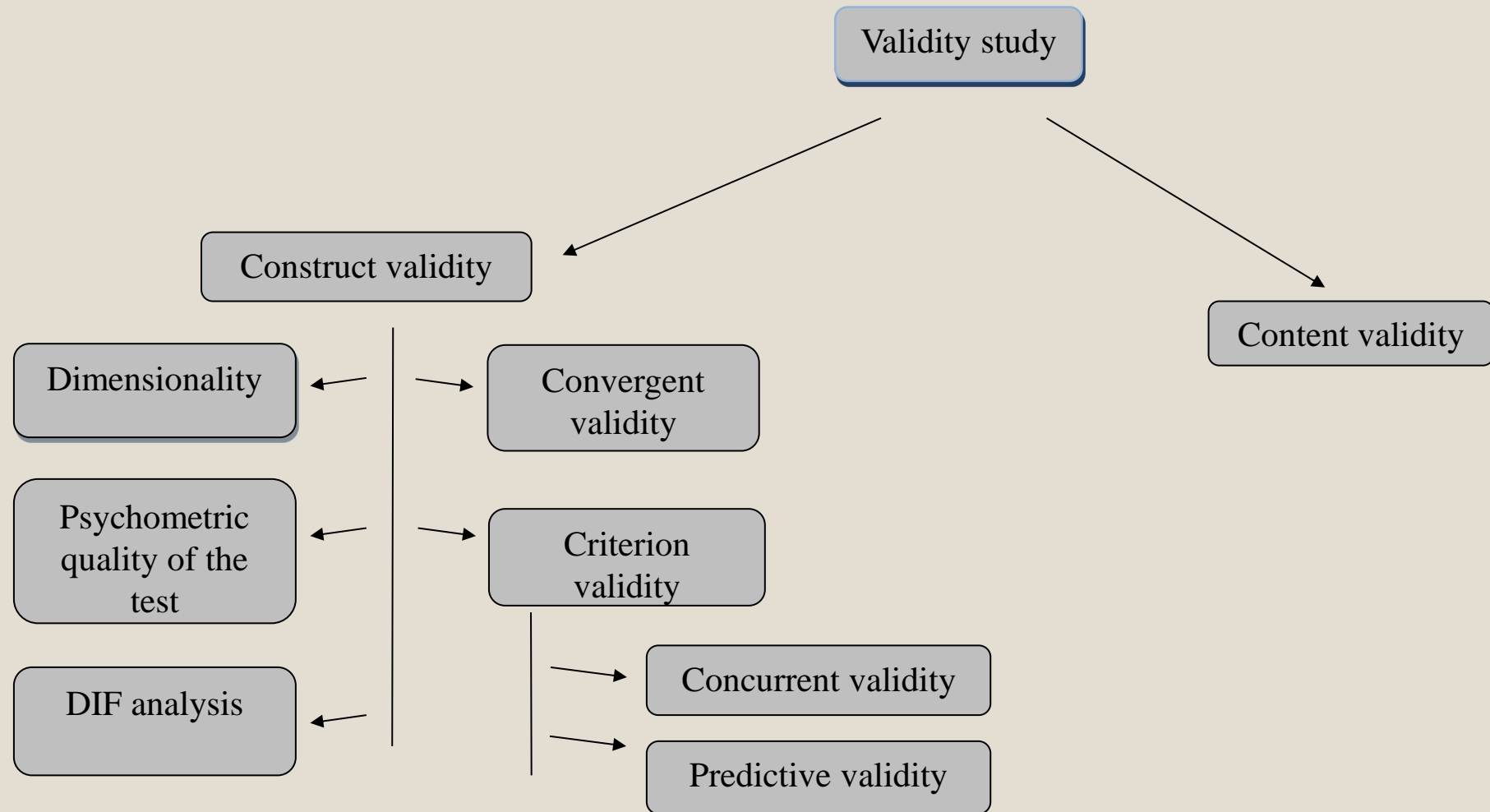
# Adaptation process

Test adaptation includes many activities from deciding whether or not a test could measure the same construct in a different language and culture, to checking equivalence of the initial and adapted test versions (Hambleton, 2005).



Validity refers to the degree to which evidence and theory support the interpretations of test scores entailed by proposed uses of tests . . . In the current validity conception, different forms of evidence on the validity of tests should not be considered to represent distinct types of validity, but validity should be considered a “unitary concept” (American Educational Research Association et al., 1999, p. 9-11).

The Dutch rating system was chosen as a basis for conducting the validity study (Evers, A., 2010).



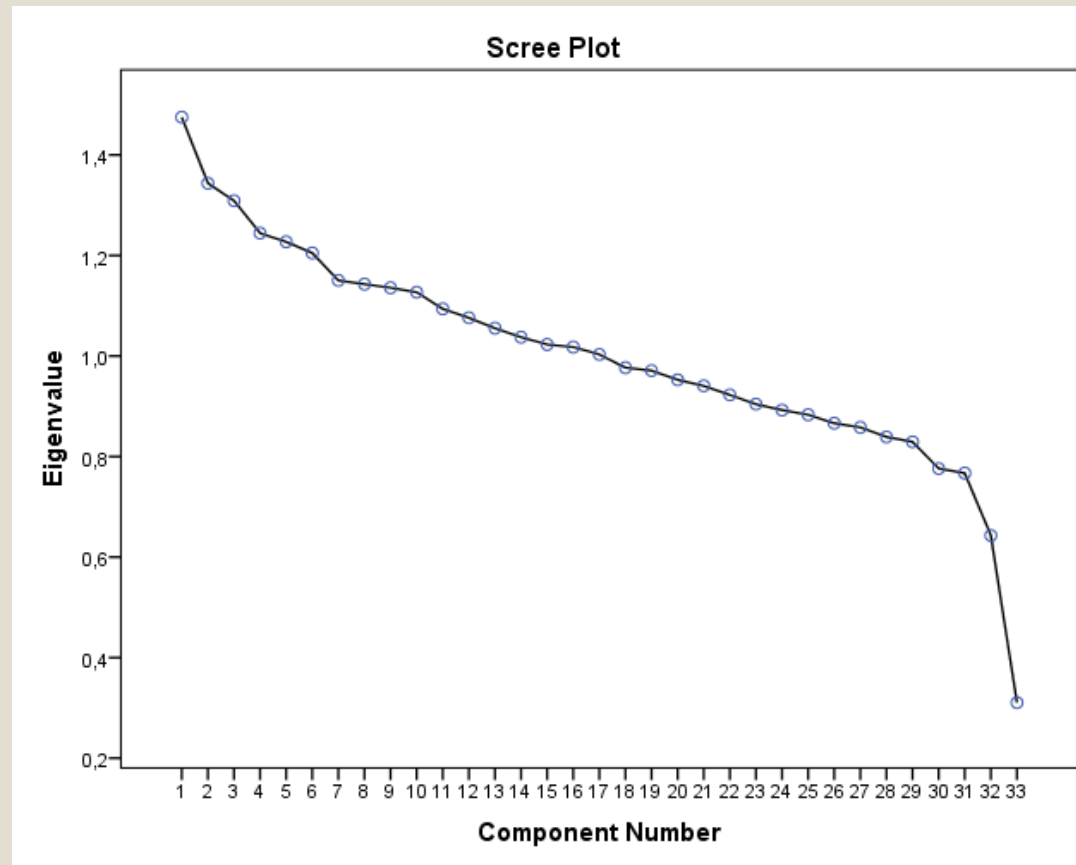
# Validation study: Math scale

Section	Description: Russian math scale
Sums	Simple sums presented informally using pictures
Number	Number manipulation – the child is asked how many more or less a number is than a target
Sums A	Sums – addition and subtraction problems presented without symbols
Sums B	More difficult mathematics problems including sums presented with formal notation

Sample	Velikiy Novgorod, 2013-2014	Krasnoyarsk and Kazan cities, 2014-2015
% of female	49%	50,4%
% of comprehensive schools	62%	56%
% of urban schools	72%	100%
in total	310	2741

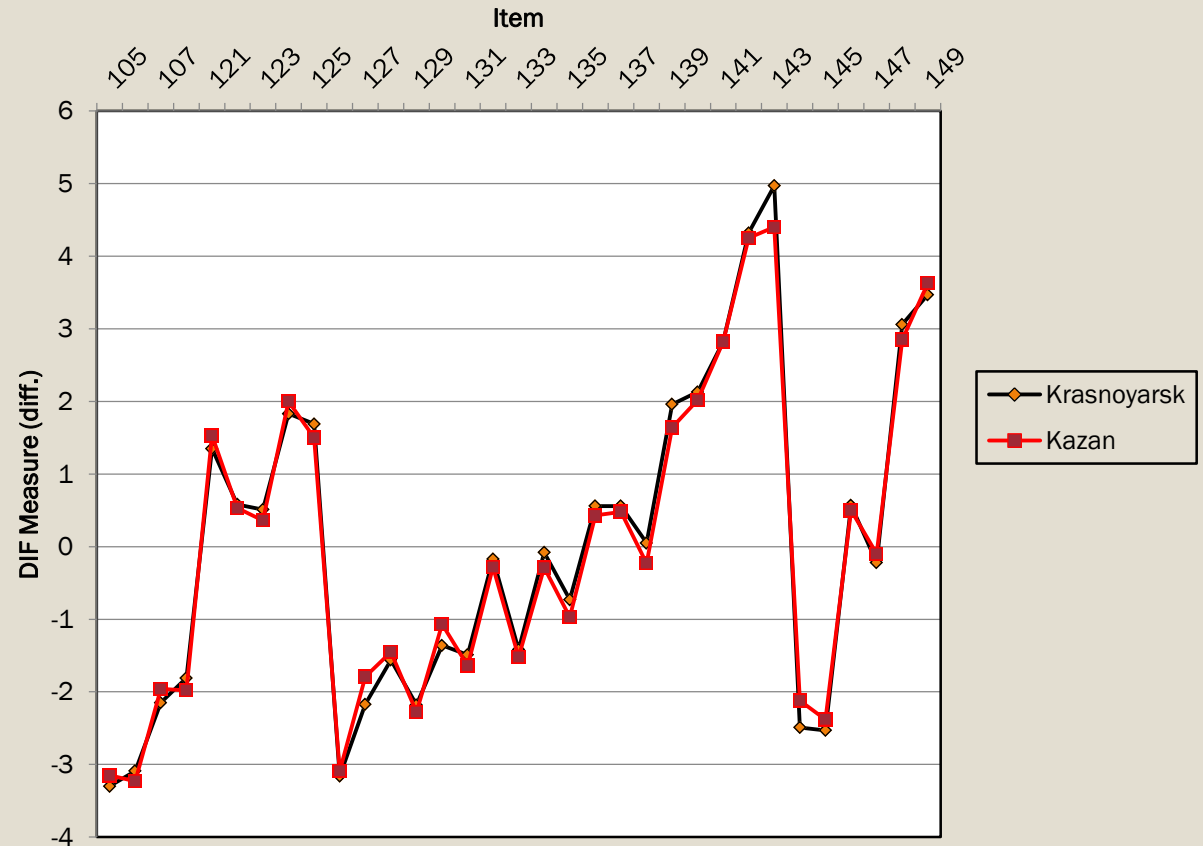
## Dimensionality study

Table of STANDARDIZED RESIDUAL variance (in Eigenvalue units)				
		-- Empirical --		Modeled
total raw variance in observations	=	57.1	100.0%	100.0%
Raw variance explained by measures	=	24.1	42.3%	41.3%
Raw variance explained by persons	=	10.1	17.7%	17.3%
Raw Variance explained by items	=	14.0	24.6%	24.0%
Raw unexplained variance (total)	=	33.0	57.7%	58.7%
Unexplnd variance in 1st contrast	=	1.7	2.9%	5.1%
Unexplnd variance in 2nd contrast	=	1.5	2.6%	4.5%



# Math scale

Evidence for fair test use: cross-regional comparison



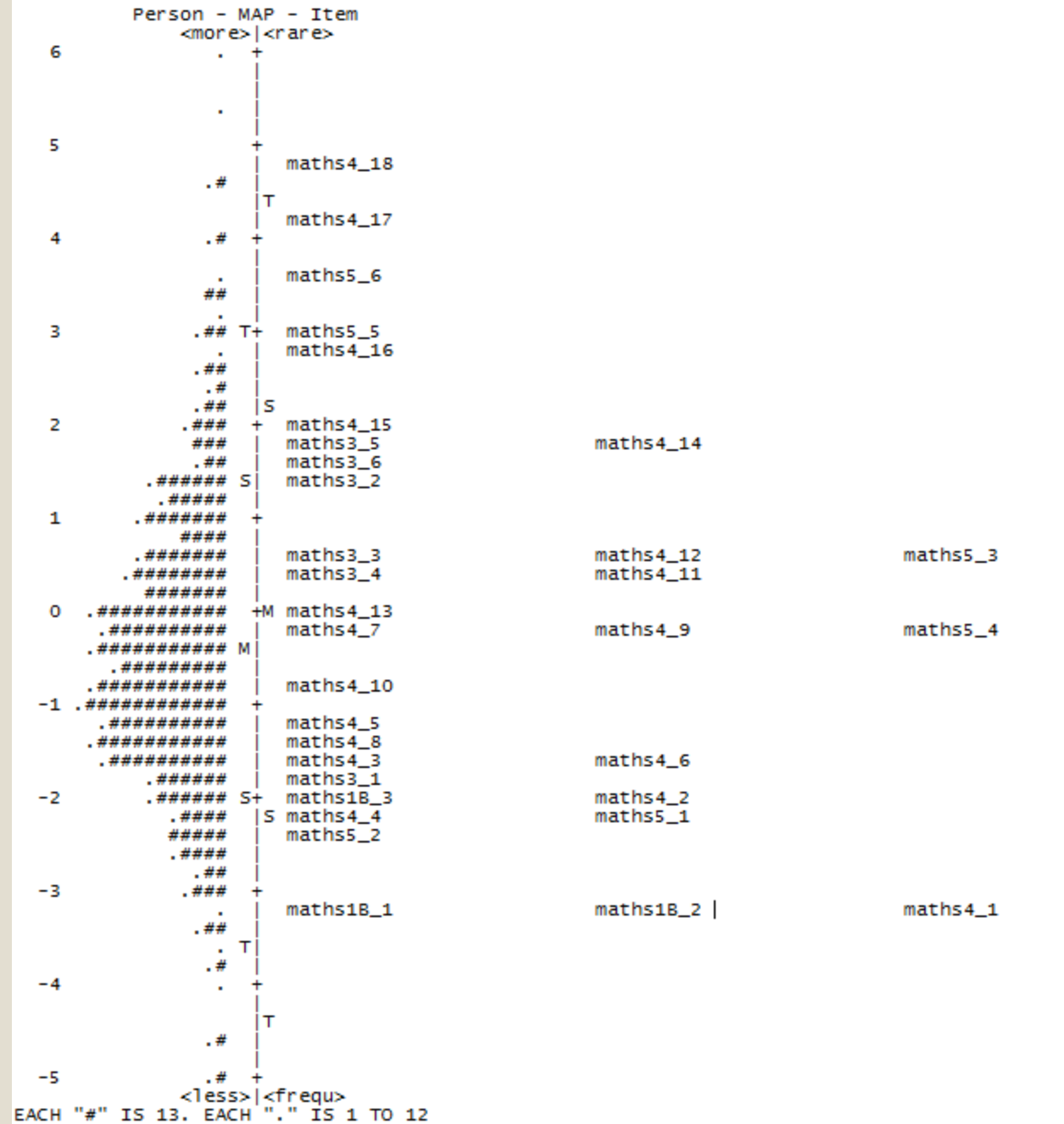


# The psychometric quality of the Math scale

33 items math scale

Person reliability 0.87

Person separation index 2.57



## The predictive validity of Math scale

- 310 children of the Velikiy Novgorod sample:
  - ✓ Their results on the iPIPS math scale
  - ✓ Their school grades in the beginning of the 2nd and of the 3d grades

The correlation of children's iPIPS math scores and their school math grades

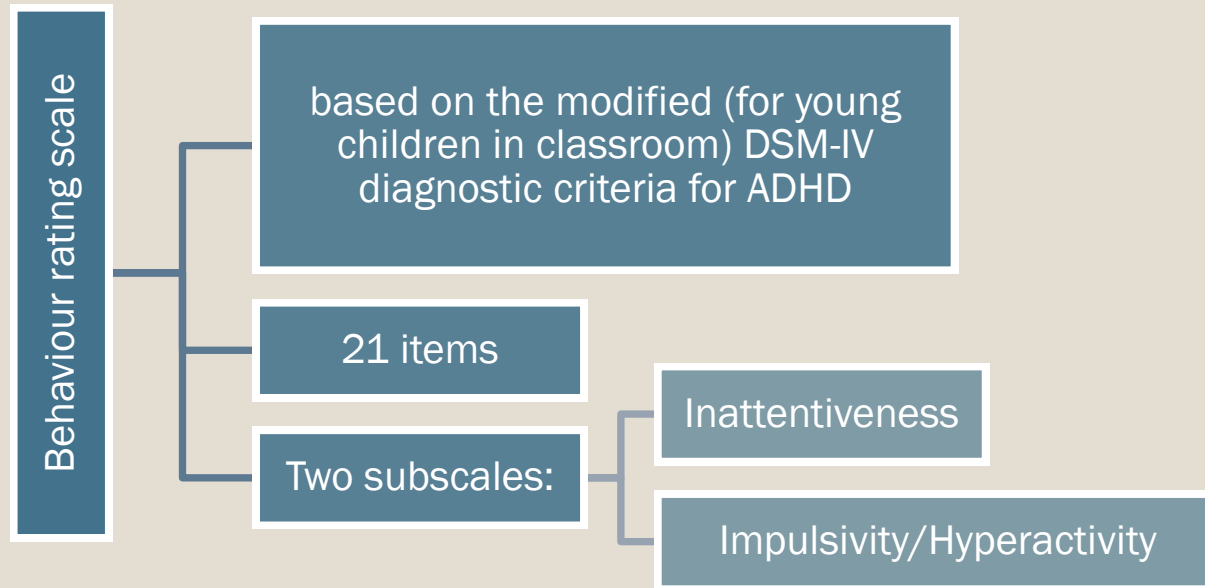
Spearman correlation	2nd grade of school, math	3d grade of school, math
iPIPS: math	.510**	.471**

\*  $p < 0.05$ , \*\*  $p < 0.001$

## The content validity

10 independent experts (psychology, pedagogy, educational science, educational policy) confirmed the feasibility of the iPIPS instrument for assessment of children starting school and for measurement of their progress during the first year of schooling

# Validation study: Behavior survey

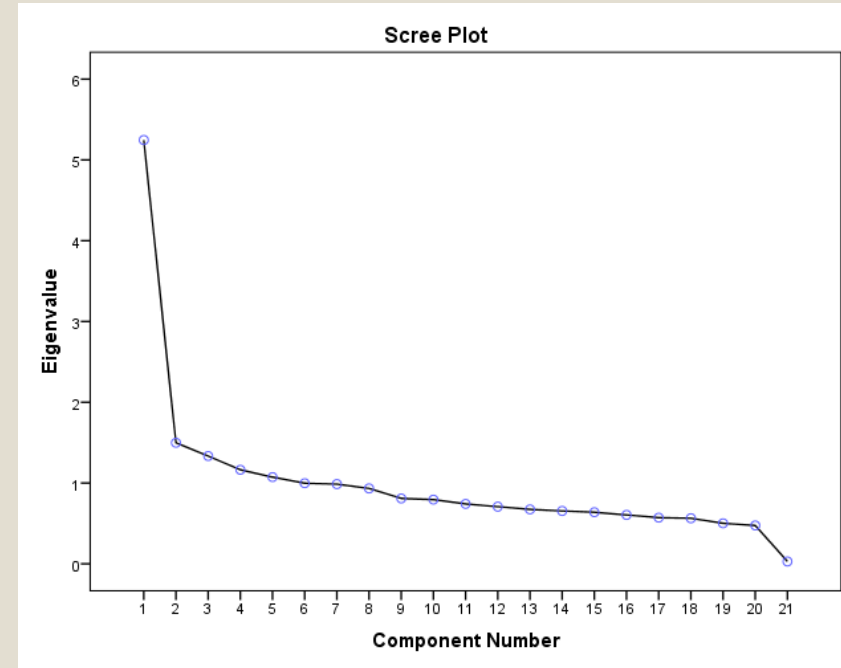


Sample	Krasnoyarsk, 2014-2015	Moscow, 2015-2016*
% of female	49%	48%
% of comprehensive schools	63%	100%
% of urban schools	100%	100%
Total	1303	2625

# Dimensionality study

		-- Empirical --		Modeled
Total raw variance in observations	=	48.5	100.0%	100.0%
Raw variance explained by measures	=	27.5	56.7%	57.5%
Raw variance explained by persons	=	17.5	36.0%	36.5%
Raw Variance explained by items	=	10.0	20.7%	21.0%
Raw unexplained variance (total)	=	21.0	43.3%	42.5%
Unexplnd variance in 1st contrast	=	5.3	10.8%	25.0%
Unexplnd variance in 2nd contrast	=	1.5	3.1%	7.1%
Unexplnd variance in 3rd contrast	=	1.3	2.7%	6.4%
Unexplnd variance in 4th contrast	=	1.2	2.4%	5.5%
Unexplnd variance in 5th contrast	=	1.1	2.2%	5.0%

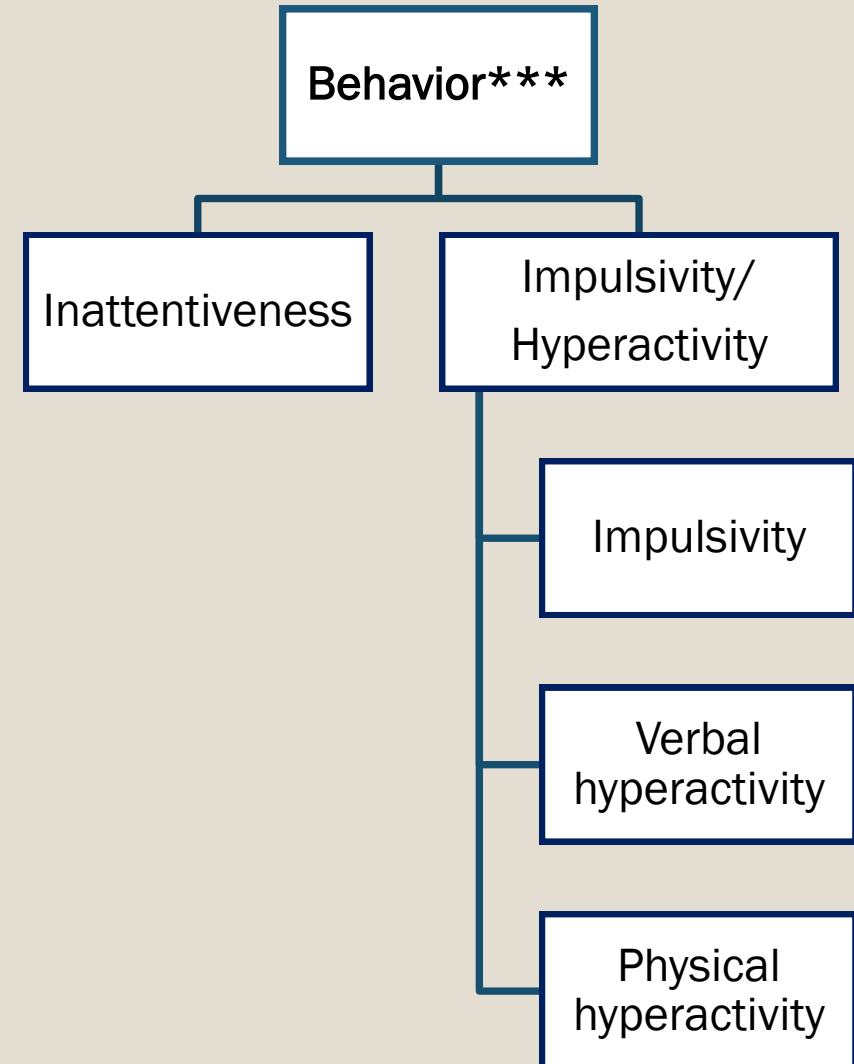
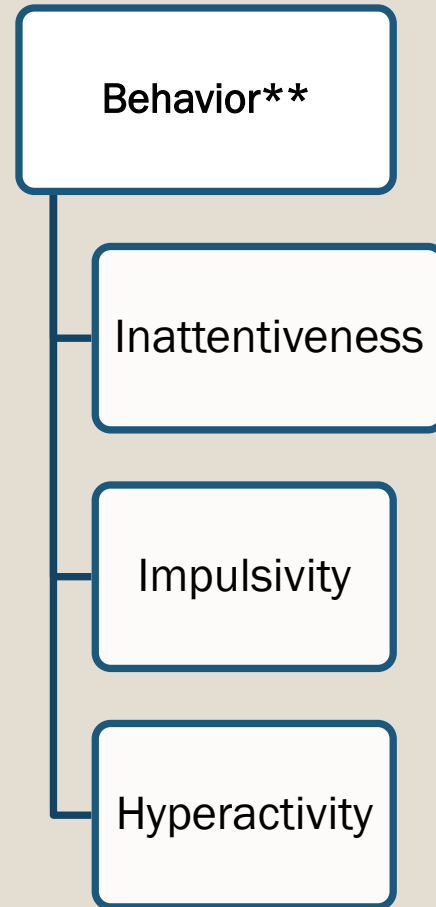
CON-TRAST	LOADING	MEASURE	INFIT MNSQ	OUTFIT MNSQ	ENTRY NUMBER
1 1	.68	-.13	.99	.92	A 15
1 1	.64	-.13	1.20	1.20	B 9
1 1	.59	-.01	1.23	1.28	C 8
1 1	.58	-.42	1.06	1.11	D 16
1 1	.48	-.09	.91	.98	E 4
1 1	.46	.26	1.03	.92	F 6
1 1	.41	-.70	1.33	1.64	G 5
1 1	.32	-.14	1.15	1.30	H 1
1 1	.31	-.50	.66	.74	I 10
1 1	.29	.17	1.08	.99	J 19
1 2	-.63	.22	1.05	.98	a 13
1 2	-.63	.34	1.00	.87	b 11
1 2	-.61	-.04	1.01	.95	c 17
1 2	-.60	-.13	.98	1.06	d 20
1 2	-.56	.33	1.37	1.39	e 3
1 2	-.54	.01	.85	.81	f 14
1 2	-.44	.12	.81	.76	g 21
1 2	-.43	.42	1.02	.87	h 2
1 2	-.39	.05	.70	.70	i 18
1 2	-.37	.72	1.34	1.25	j 7
1 2	-.12	-.33	.75	.76	K 12



Preliminary analysis indicated at least two dimensions:

- Related to inattentiveness
- Related to impulsivity/hyperactivity

# Theoretical models



\*Smith, E. V. , Johnson, B. D. (2000) Attention Deficit Hyperactivity Disorder: Scaling and Standard Setting using Rasch Measurement. *Journal of Applied Measurement* , 1 (1), 3-24.

\*\* Sukhotina, N., Egorova, T. (2008) ADHD scales. *Social and Clinical Psychiatry*, 18 (4) (in Russian)

\*\*\*Merrell, C., Tymms, P. (2005) Rasch Analysis of Inattentive, Hyperactive and Impulsive Behaviour in Young Children and the Link with Academic Achievement, *Journal of Applied Measurement* 6(1) p1 - 18.

# Behavior rating scale

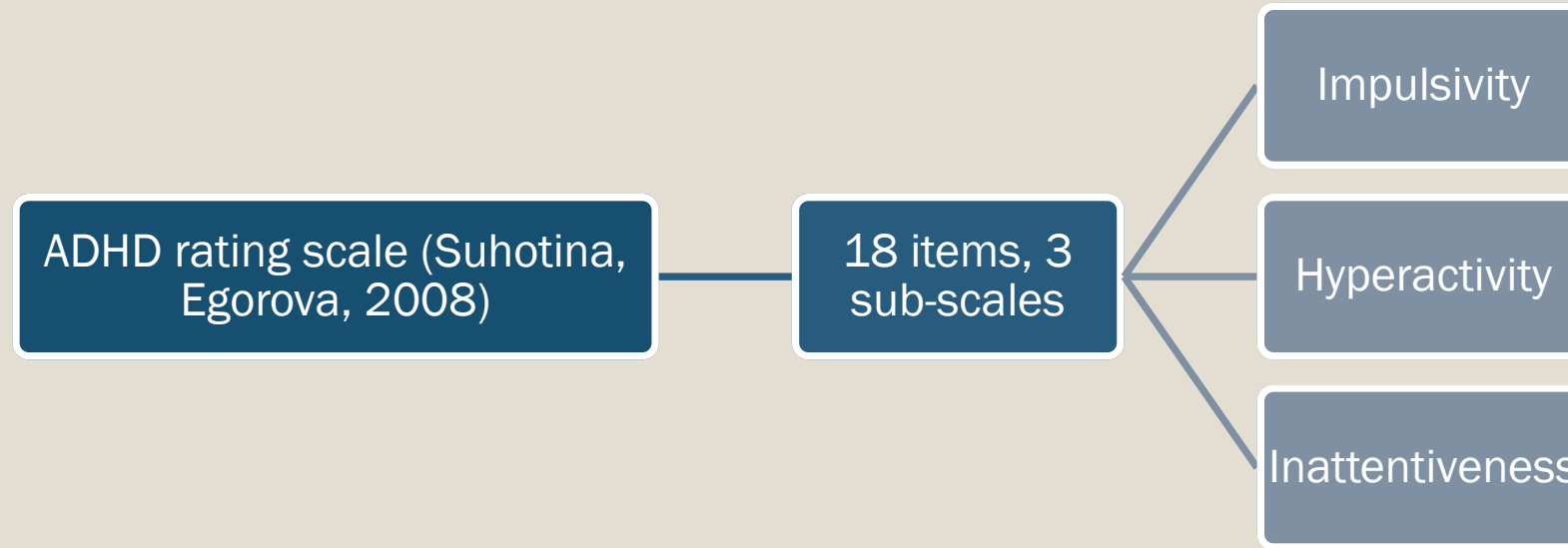
## 3-dimentional model

	Reliability	Impulsivity	Hyperactivity
Impulsivity	0.93		
Hyperactivity	0.89	0.93	
Inattentiveness	0.95	0.79	0.65

## 2-dimentional model

	Reliability	Impulsivity/ Hyperactivity
Impulsivity/ Hyperactivity	0.93	
Inattentiveness	0.94	0.75

# Convergent validity study



Sub-sample  
Moscow 2015-2016  
723, 1-graders

Pearson correlation	Imp.\ Hyper. iPIPS	Inattentiveness iPIPS	Impulsivity	Hyperactivity
Inattentiveness iPIPS	,798**			
Impulsivity	,876**	,690**		
Hyperactivity	,884**	,667**	,987**	
Inattentiveness	,766**	,814**	,882**	,854**

\*\*p<0.01

# The predictive validity of the behavior rating scale

- 1303 children of the Krasnoyarsk sample:
  - ✓ Their results on the iPIPS math and reading scale at the end of the first school year
  - ✓ Their scores on the Behavior rating scale

*The correlation of children's iPIPS math and reading scores and their behavior rating scale scores*

Pearson correlation	Impulsivity/ Hyperactivity	Inattentiveness	Reading
Inattentiveness	,790**		
Reading	-,199**	-,384**	
Math	-,120**	-,305**	,548**

\* p <0.05, \*\* p <0.001

## The content validity

3 independent experts (clinical psychology) confirmed the feasibility of the iPIPS Behavior rating scale for usage by primary school teachers for children's' potential behavioral problems identification



# Conclusions

- ✓ Cognitive scales: *confirmed*
  - *Psychometric quality of the scales*
  - *Association with children' school grades (end of the 2nd and 3d school year)*
  - *Expert evaluation of the instrument content*
  
- ✓ Behavior rating scale: *confirmed*
  - *Psychometric quality of the scale*
  - *Association with children' math and reading results on iPIPS end of year assessment*
  - *Correlation with similar constructs*
  - *Expert evaluation of the scale*

# Future steps

- Investigate the association between the iPIPS scales' scores and children results of the assessment monitoring at the end of their 4th school year
- International comparative research of children iPIPS results from different countries and cultures
- Validation study of the PSED survey

Thank you