

[Institute of Education](#) of HSE
announces

**V International Summer School
“APPLIED PSYCHOMETRICS IN PSYCHOLOGY AND EDUCATION”**

School experts:

[Mary Pitoniak](#) (PhD, Educational Testing Service)

[Carol Myford](#) (PhD, University of Illinois at Chicago)

[Linda Kook](#) (PhD, The National Center for the Improvement of Assessment)

[Lidia Dobria](#) (PhD, Wilbur Wright College)

Dates: 29 July – 04 August 2018

Summer School will be hosted by

[“Palace of Prince Kochubey”](#) study center (Saint Petersburg, Russia)

Application [form](#)

School courses will be taught in English

The participants can opt for one of the two learning tracks:

Track 1

Standard Setting Course (Dr. Mary Pitoniak) is designed to provide participants with information about a wide range of considerations relevant to setting standards on educational assessments. These include how to choose a standard setting method, which methods are currently being used, and how to know if the cut scores set for an assessment yield valid interpretations within the context of a particular testing program. The fundamentals of standard setting will be presented, including the steps required in all methods. Vertically moderated standards and adjusting committee-recommended cut scores will also be discussed. The course will give a thorough consideration to the validity of standard setting procedures and the resulting cut scores.

Fairness in Educational Testing Course (Dr. Linda Kook) is dedicated to test fairness, a topic of central importance to test developers, test takers, and those who have been using test scores for many decades. The vision of fair assessment has evolved over time and has psychometric, societal, and legal foundations. With the conceptual framework for a fairness discussion developed, the course will focus on key practical implications of this framework. Fairness in test design and development, as well as test administration will be discussed. The comparison of scores across different tests, modes of administration, grade levels, and across different languages and different populations will be considered.

Track 2

Course “Analyzing Rating Data Using Many-Facet Rasch Measurement and Multilevel Rater Modelling Approaches” by Dr. Carol Myford and Dr. Lidia Dobria is designed to introduce participants to two approaches to analyzing the rating data: a many-facet Rasch measurement (MFRM) approach, and a multilevel rater model (MRM) approach. In performance assessment settings, raters who evaluate students’ performances or products may introduce errors (rater effects) into the assessment process. The approaches discussed in the course help assessment administrators learn how various “facets” (e.g., students, raters, rating criteria) of their assessment systems are performing. This discussion can be helpful in determining to what extent their systems are under statistical control and introducing

meaningful changes to improve their systems. The course will include two parts. The first part will focus on how to use a many-facet Rasch measurement (MFRM) approach to analyzing the rating data, while the second part will feature a multilevel rater model (MRM) approach to analyzing the rating data.

Prerequisites for the School participants:

- Spoken English
- Basics knowledge of Classical Test Theory and IRT
- Your own computers (laptops)

The participants of Track 2 are expected to have prior experience running regression analyses and analyses of variance and interpreting their findings.

Learning outcomes

Courses of our Summer School will equip you with the relevant skills to conduct more objective measurements and assessments in education and psychology. Professional and scientific networking is a helpful fringe benefit. Please, scroll down to see the lists of concrete learning outcomes.

Registration and participation fee

The cost of participation in the Summer School is 700 Euros. It includes accommodation in a beautiful historical [“Palace of Prince Kochubey”](#) study center (Saint Petersburg, Russia), meals and handouts. The early bird fee makes 600 Euros and is valid before March, 15th. Payment is effected only after you receive the confirmation of your participation in a chosen Track (not after the registration). This payment DOES NOT cover transportation costs.

Application deadline is April 30th, 2018.

To apply, please, fill in the [form](#), and attach your CV.

Head of the Summer School [is Elena Kardanova](#), head of the Center for monitoring the quality of education and Master Program “Measurement in Education and Psychology”.

The organizing committee of the Summer School will be happy to provide you with additional information and answer your questions. Please, contact:

[Alina Ivanova](#) (e-mail: aeivanova@hse.ru)

[Inna Antipkina](#) (e-mail: iantipkina@hse.ru)

[Denis Ferediakin](#) (e-mail: dafederiakin@hse.ru)

Agenda for Track 1

Note: Time allotment to specific topics may shift as course materials are being developed. The purpose of this document is to show structure of daily schedule and division of time across courses (fairness and standard setting).

Time/# Hours	Topic	Instructor
Day 1—Monday, July 30 (Half-Day)		
9:00–11:00 (2 hours)	Fairness: Introduction and History Different Perspectives of Fairness	Dr. Cook
11:00–11:30	<i>Coffee Break</i>	
11:30–13:00 (1.5 hours)	Fairness: Test Design and Development	Dr. Cook
13:00–14:00	<i>Lunch</i>	
14:00–16:00	Standard Setting: Introductions, Overview, and Review of	Dr. Pitoniak

(2 hours)	Measurement Concepts Key Concepts	
16:00–16:30	<i>Coffee Break</i>	
16:30–18:00 (1.5 hours)	Standard Setting: Developing Performance-Level Descriptors	Dr. Pitoniak
18:00–19:00	<i>Dinner</i>	
Day 2—Tuesday, July 31		
9:00–11:00 (2 hours)	Fairness: Test Design and Development (<i>continued</i>)	Dr. Cook
11:00–11:30	<i>Coffee Break</i>	
11:30–13:00 (1.5 hours)	Fairness: Test Administration	Dr. Cook
13:00–14:00	<i>Lunch</i>	
14:00–16:00 (2 hours)	Standard Setting: Currently-Used Standard-Setting Methods	Dr. Pitoniak
16:00–16:30	<i>Coffee Break</i>	
16:30–18:00 (1.5 hours)	Standard Setting: Selecting and Training Panelists	Dr. Pitoniak
18:00–19:00	<i>Dinner</i>	
Day 3—Wednesday, August 1		
9:00–11:00 (2 hours)	Fairness: Comparability of Score Interpretations	Dr. Cook
11:00–11:30	<i>Coffee Break</i>	
11:30–13:00 (1.5 hours)	Standard Setting: General Steps for Standard-Setting Studies: Training in Modified Angoff Method	Dr. Pitoniak
13:00–14:00	<i>Lunch</i>	
13:00–18:00	Summer School sight-seeing tour	
18:00–19:00	<i>Dinner</i>	
Day 4—Thursday, August 2		
9:00–11:00 (2 hours)	Fairness: Comparability of Scores	Dr. Cook
11:00–11:30	<i>Coffee Break</i>	
11:30–13:00 (1.5 hours)	Fairness: Comparability of Scores (<i>continued</i>)	Dr. Cook
13:00–14:00	<i>Lunch</i>	
14:00–16:00 (2 hours)	Standard Setting: Hands-On Modified Angoff Method	Dr. Pitoniak
16:00–16:30	<i>Coffee Break</i>	
16:30–18:00 (1.5 hours)	Standard Setting: Training in Bookmark Method	Dr. Pitoniak
18:00–19:00	<i>Dinner</i>	
Day 5—Friday, August 3		
9:00–11:00 (2 hours)	Fairness: Fairness Standards	Dr. Cook
11:00–11:30	<i>Coffee Break</i>	
11:30–13:00 (1.5 hours)	Fairness: Fairness Standards (<i>continued</i>)	Dr. Cook
13:00–14:00	<i>Lunch</i>	
14:00–16:00 (2 hours)	Standard Setting: Hands-On Bookmark Method	Dr. Pitoniak
16:00–16:30	<i>Coffee Break</i>	

16:30–18:00 (1.5 hours)	Standard Setting: Documenting and Evaluating Validity Evidence	Dr. Pitoniak
18:00–19:00	<i>Dinner</i>	
Day 6—Saturday, August 4		
9:00–11:00 (2 hours)	Wrap-Up	Drs. Pitoniak and Cook
11:00–11:30	<i>Coffee Break</i>	
11:30–13:00 (1.5 hours)	Final session, feedbacks, and certificates	All instructors

Track 1 Learning objectives

Standard Setting

Upon completion of the course, participants will understand:

- Terminology used in standard setting
- Different types of standards that are used (e.g., relative vs. absolute, criterion-referenced vs. norm-referenced) and the advantages of each
- Role of judgment in standard setting and how it affects the design and implementation of standard-setting studies
- Meaning and use of classification errors
- Roles of different stakeholders in standard-setting process
- Categorizations of standard-setting methods
- General steps for standard setting studies
- Key aspects of training panelists in standard setting
- General features of many standard-setting methods
- In-depth features of modified Angoff and Bookmark methods
- Concept of definition of borderline (minimally competent) candidates
- Framework for evaluation validity of standard-setting studies
- Relevance of *Standards for Educational and Psychological Testing* to standard setting

Participants will be able to:

- Choose among standard-setting methods
- Set targets for panelist recruitment
- Develop performance-level descriptions
- Run standard setting study using modified Angoff and Bookmark methods
- Document and evaluate the validity of the results of a standard setting study

Fairness in Testing

Following completion of the course, students should have a basic understanding of:

- Different concepts of fairness in educational testing
- The ideas represented by selected underlying philosophies of fairness
- How to design, develop and evaluate a fair test
- Key issues associated with comparing test scores across modes of administration, grade levels, languages, and populations
- The Standards found in the Fairness chapter of the 2014 *Standards for Educational and Psychological Testing*

Participants will be able to:

- Explore the literature on fairness in educational testing with a basic level of understanding
- Evaluate whether or not principals of fairness have been adequately addressed in the design, development and administration of a test
- Evaluate whether comparisons of scores from tests given in different modes of administration, across grade levels, to different populations and in different languages, can be interpreted fairly

Agenda for Track 2

Note: Time allotment to specific topics may shift as we develop course materials. The purpose of this document is to show the structure of the daily schedule and the division of time across the two course segments (**Segment #1:** Learning to use a many-facet Rasch measurement (MFRM) approach to analyze rating data; **Segment #2:** Learning to use a multilevel rater model (MRM) approach to analyze rating data)

Time/# Hours	Topic	Instructor
Day 1—Monday, July 30 (Half Day)		
9:00–11:00 (2 hours)	Introductions and Overview of the Training Raters and Their Influence in Open-Ended Assessments	Dr. Myford and Dr. Dobria
11:00–11:30	<i>Coffee Break</i>	
11:30–13:00 (1.5 hours)	Conceptual and Mathematical Foundations of Many-Facet Rasch Measurement (MFRM) Creating a Data File	Dr. Myford
13:00–14:00	<i>Lunch</i>	
14:00–16:00 (2 hours)	Building a Specification File: The “Basics”	Dr. Myford
16:00–16:30	<i>Coffee Break</i>	
16:30–18:00 (1.5 hours)	Running MFRM Analyses and Reformatting Output Creating Judging Plans: Making Informed Choices	Dr. Myford
18:00–19:00	<i>Dinner</i>	
Day 2—Tuesday, July 31		
9:00–11:00 (2 hours)	Making Sense of Output from MFRM Analyses	Dr. Myford
11:00–11:30	<i>Coffee Break</i>	
11:30–13:00 (1.5 hours)	MFRM Anchoring Procedures Building a Specification File: Beyond the Basics	Dr. Myford
13:00–14:00	<i>Lunch</i>	
14:00–16:00 (2 hours)	Conceptual and Mathematical Foundations of Bias Interaction Analyses and Hybrid Models	Dr. Myford
16:00–16:30	<i>Coffee Break</i>	
16:30–18:00 (1.5 hours)	The Partial Credit Form of the MFRM Approach Rating Scale Optimization	Dr. Dobria
18:00–19:00	<i>Dinner</i>	
Day 3—Wednesday, August 1		
9:00–11:00 (2 hours)	Guided Practice Session: Running a Full MFRM Analysis (Rating Scale & Partial Credit Analyses)	Dr. Dobria and Dr. Myford
11:00–11:30	<i>Coffee Break</i>	
11:30–13:00 (1.5 hours)	Interpreting Output Formatting Results	Dr. Dobria and Dr. Myford
13:00–14:00	<i>Lunch</i>	
14:00–16:00 (2 hours)	Summer School sight-seeing tour	Summer school team
18:00–19:00	<i>Dinner</i>	
Day 4—Thursday, August 2		
9:00–11:00 (2 hours)	Conceptual and Mathematical Foundations of the Multilevel Modeling Framework	Dr. Dobria
11:00–11:30	<i>Coffee Break</i>	
11:30–13:00 (1.5 hours)	Reformulating the MFRM Model as a Multilevel Rater Model (MRM)	Dr. Dobria
13:00–14:00	<i>Lunch</i>	
14:00–16:00	Preparing Data Files for a Multilevel Analysis	Dr. Dobria

(2 hours)	Creating MDM Files Running MRM Analyses	
16:00–16:30	<i>Coffee Break</i>	
16:30–18:00 (1.5 hours)	Guided Practice Session: Running a MRM Analysis	Dr. Dobria
18:00–19:00	<i>Dinner</i>	
Day 5—Friday, August 3		
9:00–11:00 (2 hours)	Interpreting Output from a MRM Analysis Formatting Output from a MRM Analysis (Tables & Graphs)	Dr. Dobria
11:00–11:30	<i>Coffee Break</i>	
11:30–13:00 (1.5 hours)	Research Questions Addressed by a MRM Analysis	Dr. Dobria
13:00–14:00	<i>Lunch</i>	
14:00–16:00 (2 hours)	Extending the MRM with Covariates of Interest Rater Drift Differential Facet Functioning	Dr. Dobria
16:00–16:30	<i>Coffee Break</i>	
16:30–18:00 (1.5 hours)	Model Selection Advantages and Disadvantages of using MFRM and MRM Approaches for Analyzing Rating Data	Dr. Dobria
18:00–19:00	<i>Dinner</i>	
Day 6—Saturday, August 4		
9:00–11:00 (2 hours)	Wrap-Up	Drs. Myford and Dobria
11:00–11:30	<i>Coffee Break</i>	
11:30–13:00 (1.5 hours)	Final session, feedbacks, and certificates	All instructors

Track 2 Learning Objectives

Analyzing Rating Data

In this course, participants will work toward acquiring the following knowledge, skills, and understandings:

- Knowledge of different types of rater effects
- Knowledge of the conceptual and mathematical foundations of MFRM and MRM approaches for analyzing rating data
- Knowledge of the advantages and disadvantages of using MFRM and MRM approaches for analyzing rating data
- Knowledge of the kinds of group-level and individual-level questions that researchers and practitioners can address using MFRM and MRM approaches
- Knowledge of the kinds of information found in output from MFRM and MRM analyses
- An understanding of the role of raters in open-ended assessments
- An understanding of how raters can influence ratings
- An understanding of how researchers can use MFRM and MRM approaches to analyze rating data in order to help establish quality control over an assessment system
- An understanding of how MFRM and MRM approaches for analyzing rating data differ and when each type of approach is most appropriately used
- Skill in creating judging plans
- Skill in creating specification files for MFRM and MRM analyses
- Skill in preparing data files for MFRM and MRM analyses
- Skill in constructing multivariate data matrices
- Skill in running analyses of rating data using the Minifac (Linacre, 2017) and HLM v. 7.03 for Windows (Raudenbush, Bryk, & Congdon, 2013) computer software packages
- Skill in printing and reformatting output from MFRM and MRM analyses

- Skill in interpreting the various statistical indices that appear in output from MFRM and MRM analyses
- Skill in troubleshooting when analyses will not run