

## NCME VIRTUAL ANNUAL MEETING SCHEDULE AT A GLANCE

**WEDNESDAY, SEPTEMBER 9**

11:00 AM – 12:30 PM ET	Psychometric Innovations and Advances in Medical Educational Assessments	<p><u>Participants:</u></p> <p>Helping Item Writers by Making NLP-Based Suggestions for Item Distractors <b>Peter Baldwin</b>, National Board of Medical Examiners (NBME); <b>Victoria Yaneva</b>, National Board of Medical Examiners; <b>Janet Mee</b>, National Board of Medical Examiners; <b>Brian Clauser</b>, National Board of Medical Examiners; <b>Le An Ha</b>, University of Wolverhampton</p> <p>On the Utility of Using Transfer Learning to Predict Item Characteristics <b>Kang Xue</b>, University of Georgia; <b>Victoria Yaneva</b>, National Board of Medical Examiners; <b>Christopher Runyon</b>, NBME</p> <p>Exploring Automated Assessment of Spoken English Proficiency for Medical Licensure Exams <b>Su G. Somay</b>, National Board of Medical Examiners (NBME); <b>Jessica Salt</b>, Education Commission for Foreign Medical Graduates (ECFMG)</p> <p>An Automated Scoring Routine for Constructed Responses on a Medical Licensure Exam <b>Christopher Runyon</b>, NBME; <b>Polina Harik</b>, National Board of Medical Examiners; <b>Abeed Sarker</b>, Emory University School of Medicine; <b>Graciela Gonzalex-Hernandez</b>, University of Pennsylvania</p> <p>Identifying Response Pattern in Clinical Notes with NLP Feature Identification <b>Janet Mee</b>, National Board of Medical Examiners; <b>Ravi Pandian</b>, National Board of Medical Examiners; <b>Andrew Houriet</b>, National Board of Medical Examiners; <b>Christopher Yang</b>, Drexel University</p> <p>IRTTree Response Style Modeling for Improving Feedbacks about Cognitive Biases <b>Lale Khorramdel</b>, National Board of Medical Examiners; <b>Matthias von Davier</b>, National Board of Medical Examiners; <b>Ann King</b>, National Board of Medical Examiners; <b>Andrew Houriet</b>, National Board of Medical Examiners</p> <p><u>Session Organizers:</u></p> <p><b>Lale Khorramdel</b>, National Board of Medical Examiners</p> <p><b>Christopher Runyon</b>, NBME</p> <p><u>Chair:</u></p> <p><b>Lale Khorramdel</b>, National Board of Medical Examiners</p> <p><u>Discussant:</u></p> <p><b>Isaac I Bejar</b>, Educational Testing Service</p>
1:00 – 2:30 PM ET	Fireside Chat with Classroom Assessment Task Force: Making Measurement Meaningful with Classroom Assessment	<p><u>Presenters:</u></p> <p><b>Sue Brookhart</b>, Duquesne University</p> <p><b>Heidi Andrade</b>, SUNY Albany</p> <p><b>Mark Wilson</b>, University of California, Berkeley</p>

		<p><b>E Caroline Wylie</b>, Educational Testing Service</p> <p><b>Jade Caines Lee</b>, University of New Hampshire</p> <p><b>Alison Bailey</b>, UCLA</p> <p><b>Neal Kingston</b>, University of Kansas</p> <p><b>Dale Whittington</b>, Retired</p> <p><u>Session Organizer:</u></p> <p><b>Kristen Huff</b>, Curriculum Associates</p>
5:00 – 6:30 PM ET	Using Psychometric Approaches to Improve Test Development	<p><u>Participants:</u></p> <p>Developing Useful Test Blueprints: An Evaluation Study <b>Andrew C. Dwyer</b>, American Board of Pediatrics; <b>Robert Furter</b>, American Board of Pediatrics</p> <p>One Blueprint Fits All? Using Person-centered Methodological Approaches in Practice Analysis Studies <b>Pamela Kaliski</b>, American Board of Internal Medicine; <b>Kelly Foelber</b>, ABIM; <b>Jerome Clauser</b>, ABIM</p> <p>The Effect of Item Writing Experience on Item Quality <b>Bozhidar M. Bashkov</b>, American Board of Internal Medicine; <b>Jerome Clauser</b>, ABIM</p> <p>Modeling Text Complexity Judgments and Their Association with Item Difficulty <b>Brian F. Patterson</b>, Curriculum Associates; <b>Pamela Seastrand</b>, Curriculum Associates</p> <p>Using Item Text to Predict Item Survival: Can It Go Beyond Linguistic Characteristics <b>Victoria Yaneva</b>, National Board of Medical Examiners; <b>Le An Ha</b>, University of Wolverhampton; <b>Peter Baldwin</b>, National Board of Medical Examiners (NBME); <b>Janet Mee</b>, National Board of Medical Examiners</p> <p><u>Session Organizer:</u></p> <p><b>Bozhidar M. Bashkov</b>, American Board of Internal Medicine</p> <p><u>Discussant:</u></p> <p><b>Kurt F. Geisinger</b>, Buros Center for Testing</p>
7:30 – 9:00 PM ET	NCME Social Event: Trivia <i>Pre-registration required;</i> <i>N&lt;= 150</i>	<p>Join us for a virtual game night with “The Go Remote: Virtual Game Show!”</p> <p>HOW IT WORKS: Grab some headphones, a beverage of your choice, and your game face (yes-you will want to turn on your laptop video for this!) for an action-packed Happy Hour Game Show. Compete in live games such as: Trivia, Pictionary, Name that Movie/Tune, and GIF Battles. This session will be hosted outside of the event platform. Pre-registration required. <a href="#">Sign up here.</a></p>

THURSDAY, SEPTEMBER 10

9:00 – 10:30 AM ET	Assessing Collaborative Problem Solving at Scale: The Status Quo and the Next	<p><u>Participants:</u></p> <p>CPS assessment in ATC21S - learning from the past <b>Esther Care</b>, <i>Brookings Institution</i></p> <p>Updates on CPS Assessment in PISA 2015 <b>Art Graesser</b>, <i>University of Memphis</i></p> <p>Assessing CPS: Conceptualization and Methodological Differences between the CSCL and Assessment Communities <b>Nancy Law</b>, <i>University of Hongkong</i></p> <p>Learning &amp; Assessment of Collaborative Problem Solving at ACT <b>Alina von Davier</b>, <i>ACTNext</i>; <b>Kristin Stoeffler</b>, <i>ACTNext by ACT</i>; <b>Benjamin Deonovic</b>, <i>ACTNext by ACT</i>; <b>Michael Yudelson</b>, <i>ACTNext by ACT</i>; <b>Pravin Chopade</b>, <i>ACTNext by ACT</i>; <b>David Edwards</b>, <i>ACTNext by ACT</i>; <b>Saad Khan</b>, <i>ACTNext by ACT</i>; <b>Yigal Rosen</b>, <i>ACTNext by ACT</i></p> <p>ETS' Efforts towards Learning and Assessments of CPS at Scale <b>Jiangang Hao</b>, <i>Educational Testing Service</i></p> <p><u>Session Organizer:</u></p> <p><b>Jiangang Hao</b>, <i>Educational Testing Service</i></p> <p><u>Discussant:</u></p> <p><b>Patrick Kyllonen</b>, <i>Educational Testing Service</i></p>
11:00 AM – 12:30 PM ET	Assessment of Mathematical and Scientific Reasoning: An alternative to Machine-scoring Open-Ended Items	<p><u>Participants:</u></p> <p>Investigating an Alternative to Machine-scoring of Open-Ended Items <b>Mark Wilson</b>, <i>University of California, Berkeley</i></p> <p>Developing selected response items for a data-based decision-making instrument <b>Amy Arneson</b>, <i>Education Northwest</i></p> <p>An Exploration of Selected-Response Items Compared to Constructed-Response Item Types in Science Education <b>Linda Morell</b>, <i>University of California, Berkeley</i>; <b>Weeraphat Suksiri</b>, <i>University of California-Berkeley</i>; <b>Sara Dozier</b>, <i>Stanford University</i>; <b>Jonathan Osborne</b>, <i>Stanford University</i>; <b>Mark Wilson</b>, <i>University of California, Berkeley</i></p> <p>Comparing Selected Response and Constructed Response Items in Mathematical Problem Solving <b>Yukie Toyama</b>, <i>University of California</i>; <b>Jerred Jolin</b>, <i>University of California, Berkeley</i>; <b>James Mason</b>, <i>University of California, Berkeley</i>; <b>Mark Wilson</b>, <i>University of California, Berkeley</i></p> <p><u>Session Organizer:</u></p> <p><b>Mark Wilson</b>, <i>University of California, Berkeley</i></p> <p><u>Chair:</u></p> <p><b>Karen Draney</b>, <i>University of California Berkeley</i></p> <p><u>Discussants:</u></p>

		<p><b>Richard Patz</b>, University of California-Berkeley</p> <p><b>Jonathan Osborne</b>, Stanford University</p>
1:00 – 2:00 PM ET	Presidential Address: “Psychometricians in the Hands of an Angry Mob.”	<p><b>Stephen G. Sireci</b>, University of Massachusetts Amherst</p> <p><b>Ye Tong</b>, Pearson</p>
2:30 – 4:00 PM ET	Beyond Accommodations: Intentional Design Methods for Improving Validity When Assessing Diverse Populations	<p><u>Participants:</u></p> <p>Improving Fairness: Theories of Action and Intentional Design Methods <b>Melissa L. Gholson</b>, Educational Testing Service</p> <p>Equitable Assessment through Twinning <b>Danielle Guzman-Orth</b>, Educational Testing Service; <b>Cary Supalo</b>, Educational Testing Service; <b>Cinda Parton</b>, WestEd</p> <p>Applying Universal Design Principles to the Accessible Test Development Process <b>Robert P Dolan</b>, Diverse Learners Consulting; <b>Cara Wojcik</b>, CAST; <b>Jenna Gravel</b>, CAST; <b>Allison Posey</b>, CAST; <b>Elizabeth Hartmann</b>, Lasell College; <b>Kimberly Ducharme</b>, CAST; <b>Jose Blackorby</b>, CAST</p> <p><u>Session Organizer:</u></p> <p><b>Robert P Dolan</b>, Diverse Learners Consulting</p> <p><u>Discussant:</u></p> <p><b>Sheryl Lazarus</b>, National Center on Educational Outcomes</p>
4:30 – 6:00 PM ET	Principled Item Design: State-of-the-Art	<p><u>Participants:</u></p> <p>Test and Item Development with Cognitive Design Systems: Some Examples <b>Susan Embretson</b>, Professor</p> <p>The Interplay of Validity and Efficiency <b>Isaac I Bejar</b>, Educational Testing Service</p> <p>Considerations for Training Subject Matter Experts to Write Item Models <b>Audra Kosh</b>, Edmentum</p> <p>The Challenges of Principled Item Design <b>Richard Luecht</b>, University of North Carolina</p> <p>Automated Item Generation Using Deep Learning <b>Matthias von Davier</b>, National Board of Medical Examiners</p> <p><u>Session Organizers:</u></p> <p><b>Richard Luecht</b>, University of North Carolina</p> <p><b>Isaac I Bejar</b>, Educational Testing Service</p> <p><u>Chair:</u></p> <p><b>Richard Luecht</b>, University of North Carolina</p>

		<p>Discussant:</p> <p><b>Steve Ferrara</b>, Measured Progress</p>
6:00 – 7:00 PM ET	DREAM ON! ( <b>DR</b> inking, Eating, And Measurement <b>ON</b> line Party)	<p><b>Stephen G. Sireci</b>, University of Massachusetts Amherst</p> <p><b>Bruno D. Zumbo</b>, University of British Columbia</p> <p><b>Saskia Wools</b>, CITO</p>
FRIDAY, SEPTEMBER 11		
11:00 AM – 12:30 PM ET	Feasibility of an Embedded Field-Test Model in the Enhancement of Preequating	<p><u>Participants:</u></p> <p>An Overview of the Embedded Field-Test Model <b>Wei Tao</b>, ACT</p> <p>Evaluating Impacts on Operational Item Performance in the Embedded Field-Test Model <b>Troy Chen</b>, ACT</p> <p>Comparing CTT Postequating and IRT Preequating in the Embedded Field-Test Model <b>Yi-Fang Wu</b>, ACT</p> <p>Investigating Potential Confounding Testing Mode Effects in Transition from Postequating to Preequating <b>Yong He</b>, ACT</p> <p><u>Session Organizer:</u></p> <p><b>Yi-Fang Wu</b>, ACT</p> <p><u>Chair:</u></p> <p><b>Yi-Fang Wu</b>, ACT</p> <p><u>Discussant:</u></p> <p><b>Deborah Harris</b>, University of Iowa</p>
1:00 PM – 2:30 PM ET	Separate but (Un)Equal? Measurement Expertise vs. Policy Knowledge in Testing for Accountability	<p><u>Participants:</u></p> <p>Does the Measurement Community ‘Own’ the Problems on Interpretation and Use of Test Scores? <b>Ellen Forte</b>, edCount, LLC</p> <p>Ensuring that Test Use Leads to the Intended, Positive Consequences, while Minimizing Unintended, Negative Consequences <b>Suzanne Lane</b>, University of Pittsburgh</p> <p>Towards Inclusive Assessments for all Students, and Inclusive Accountability Systems <b>Martha Lynn Thurlow</b>, NCEO/University of Minnesota</p> <p>Mapping the Relationship Between Policy and Measurement Considerations in Supporting Accountability Testing <b>Neal Kingston</b>, University of Kansas</p> <p>Separate and Unequal? Lessons from Selective High Schools Uses of Standardized Test Scores <b>Howard Everson</b>, City University of New York</p>

		<p>Why Assessment Professionals Should Cultivate Policy Analysis Skills <b>Aaron Pallas</b>, <i>Columbia University, Teachers College</i></p> <p>Conflicting Theories of Action? Reconciling Mindsets of Measurement vs. K-12 Stakeholders to Improve Validity and Consequences of Testing <b>Madhabi Chatterji</b>, <i>Columbia University, Teachers College</i></p> <p><u>Session Organizer:</u></p> <p><b>Madhabi Chatterji</b>, Columbia University, Teachers College</p>
3:00 PM – 4:30 PM ET	New Research Findings on Understanding and Managing Test-Taking Disengagement	<p><u>Participants:</u></p> <p>Aggregate-Level Ability Estimation Accuracy Under Varying Noneffortful Responding Types and Rates <b>Joseph A. Rios</b>, <i>University of Minnesota</i>; <b>Chelsey Legacy</b>, <i>University of Minnesota</i></p> <p>Using Retest Data to Evaluate and Improve Effort-Moderated Scoring <b>Steven L. Wise</b>, <i>NWEA</i>; <i>Megan R. Kuhfeld, NWEA</i></p> <p>How to Design a Drag-and-Drop Item for Motivated Responding <b>Blair Lehman</b>, <i>ETS</i>; <b>Burcu Arslan</b>, <i>ETS</i></p> <p>Test Value and Emotions: Predicting Examinee Effort and Performance on Low-Stakes Tests <b>Paulius Satkus</b>, <i>James Madison University</i>; <b>Sara J. Finney</b>, <i>James Madison University</i></p> <p>Do Students Rapidly Guess Repeatedly Over Time? A Longitudinal Analysis <b>James Soland</b>, <i>NWEA</i>; <b>Megan R. Kuhfeld</b>, <i>NWEA</i></p> <p><u>Session Organizer:</u></p> <p><b>Steven L. Wise</b>, NWEA</p> <p><u>Chair:</u></p> <p><b>Dena Pastor</b>, James Madison University</p>
5 PM ET – 6:30 PM ET	Insights on Text-to-Speech as a Universal Design Feature: NAEP Mathematics Process Data	<p><u>Participants:</u></p> <p>Evaluation of the Text-to-Speech (TTS) Tool Using NAEP Process Data <b>Soo Lee</b>, <i>American Institutes for Research</i></p> <p>Exploration of TTS Using the Differential Boost Framework <b>Juanita Hicks</b>, <i>American Institutes for Research</i></p> <p>Exploring Read-Aloud Sentences to Understand Learners' TTS Use Tendency <b>Ruhan Circi</b>, <i>AIR</i></p> <p><u>Session Organizer:</u></p> <p><b>Juanita Hicks</b>, American Institutes for Research</p> <p><u>Discussant:</u></p> <p><b>Heather Buzick</b>, Educational Testing Service</p>
6:30 PM – 7:30 PM ET	GSIC Social Event	<p><u>Hosts:</u></p> <p><b>Maura O’Riordan</b>, University of Massachusetts Amherst</p>

		<b>Scott Holcomb</b> , UNC Charlotte
--	--	--------------------------------------

For the full line up, including sessions descriptions and call in information, visit the online portal – [CLICK HERE!](#)