Classroom Assessment as a Learning Experience

NCME SPECIAL CONFERENCE ON CLASSROOM ASSESSMENT
SEPT. 18-19 | UNIVERSITY MEMORIAL CENTER, CU BOULDER

Conference Information, Schedule and Program

NCME national council on measurement in education

CADRE

School of Education
UNIVERSITY OF COLORADO BOULDER

Aurora Public Schools

Cherry Creek Schools
Dedicated to Excellence

CDE CO

DENVER PUBLIC SCHOOLS

Center for Assessment

Table of Contents

University Memorial Center 2nd Floor Map

General conference information

Checking in and Name Badges

There are three check-in periods for the conference: 1) During the welcome reception at the Embassy Suites on Tuesday, September 17th from 5-7; 2) in front of the UMC Ballroom from 7:30am-5:00pm on Wednesday, September 18th; and 3) 7:30am-12:00pm on Thursday, September 19th. Please do not lose your name badge as you must have this in order to attend all keynotes, sessions, and meals. Kindly return your name badge at the registration table or in designated boxes located in the Ballroom before you leave the conference as the lanyards will be reused.

Wi-Fi

Visitors can connect to Wi-Fi anywhere on campus using "UCB Guest". No password is required. While we typically do not have connection issues at the UMC, we strongly recommend that presenters back up any slides on your laptop or tablet.

Conference URL

https://www.colorado.edu/cadre/classroom-assessment-learning-experience

Transportation and Parking

Hotel Shuttle
Due to the difficulty in finding open parking spots on campus, we strongly encourage attendees staying at the Embassy Suites or Hilton Garden Inn to take advantage of our complimentary 25-passenger shuttle via Green Ride, which will go directly to the University Memorial Center. The shuttle will pick up in the alley between the Embassy Suites and Hilton Garden Inn. It takes approximately 15 minutes to drive between the hotels and the UMC. See the shuttle schedule below:

Wednesday, September 18th Hotel to UMC
7:00, 7:15, 7:30, 7:45, 8:00, 8:15, 8:30

Wednesday, September 18th UMC to Hotel
4:45, 5:15, 5:45, 6:15, 6:45, 7:00

Thursday, September 19th Hotel to UMC
7:00, 7:15, 7:30, 7:45, 8:00, 8:15, 8:30

Thursday, September 19th UMC to Hotel
2:15 (only one shuttle returning to the hotel on Day 2)

Traveling between CU and the Airport
If you are traveling directly between the airport and CU campus, and have some flexibility with time, we recommend that you take the AB/AB1 bus that drops off and picks up within a 5-minute walk from the UMC at the Broadway-Euclid Stop- check here for the schedule. Otherwise, we recommend either booking through Green Ride Boulder or your favorite ride sharing app.

Parking at the Hotel and on Campus
If you are driving, be prepared to pay for parking at your hotel and on campus. Parking at the Hilton/Embassy Suites is $16/night. If you are driving to CU Campus, we recommend that you park at the Euclid Parking Garage in the CASE building, located between Broadway and 18th Street, east of the UMC building, and build in at least 20 minutes to park and walk to the UMC. You can pay for parking using the ParkMobile app. See here for details.

Media Permission

By registering for this conference, you grant permission to the National Council on Measurement in Education (NCME) and Achievement and the University of Colorado Boulder School of Education (SOE) to photograph, videotape, audiotape, and otherwise record images as part of the program identified below. You release NCME, SOE, and the University from any and all liability resulting from the publication of said photos, videos, or remarks. Publications may include but are not limited to website, publications, social media, and presentations.

Print services

If you require printing services we recommend going to the Inkspot located on the first floor of the UMC Building. Their hours are 7:30 a.m. to 4:30 p.m. M-F.

Workspace and Personal Belongings

We recognize that some participants may need to step out to send an email or take a call during the conference. There are sitting areas in the Student Engagement and Collaboration Area (SECA) on the 2nd floor, and if the weather is nice, you can also head outside to the South Terrace.

If you need a space to store your luggage or large items like a poster, we recommend leaving them in the back of the Aspen Room. The University is not responsible for lost personal items.

Lost and Found

Smaller misplaced items may be dropped off at the registration desk in front of the Ballroom.

Lactation Room

Room 132 on the first floor of the UMC is a dedicated lactation room. Please go to the reception desk for a key.

All Gender Restrooms

There are two all gender restrooms on the 2nd floor of the UMC. One is located between Rooms 245 and 247 and the other is next to the West Ballroom.

Certificates

If you requested a certificate of attendance for the conference then please sign the check-out sheet at the registration desk before departing the conference. If you are attending both days, you only need to sign out on the second day.

Preventing Altitude Sickness

The elevation of the city of Boulder is 5,328 ft above sea level and altitude sickness is a serious issue for many visitors coming from lower elevations. We strongly recommend hydrating frequently, using sunscreen if you decide to head outdoors, and being mindful of alcohol intake.

What to Do in Boulder

Boulder is a fun, active and exciting city with a lot to offer for a variety of interests. Here is a helpful list of options.

Robert L. Linn Memorial Lecture

All conference attendees are welcome to attend the lecture by award recipient Lydia Liu, Senior Research Director at ETS following the September 19th Lunch Panel. The lecture and reception will be at Old Main Chapel from 3-4:30pm. For more details, click here.

Sustainability commitment

Please take a moment to review our statement on minimizing our environmental impact during the conference.

Before traveling to Boulder, we recommend taking the following actions so that you can help us fulfill our commitment:

- Save the link to this final program document on your laptop, tablet, or phone so you can easily access it before and during the conference. If you prefer having a paper program, please print out your own personal copy prior to the conference.
- Bring a reusable water bottle to keep hydrated during the conference. Hydration is critical for participants coming from lower elevations and there are several refilling stations around the UMC.
- Bring your own note-taking devices.
- Carpool as often as possible and use our Green Ride shuttle.

Conference Schedule with Presentation Summaries

Link to Schedule At-A-Glance

Link to Full Conference Schedule

The presentation summaries listed in the program include a list of co-authors.

Presenters' names are printed in bold.

SCHEDULE AT-A-GLANCE

TUESDAY, SEPT. 17

5:00 – 7:00 P.M. **WELCOME RECEPTION AND EARLY CHECK-IN**
Embassy Suites Boulder Patio, 2601 Canyon Blvd, Boulder, CO 80302

WEDNESDAY, SEPT. 18 | DAY ONE | UNIVERSITY MEMORIAL CENTER

7:30 - 5:00 P.M. **CONFERENCE CHECK-IN** | 2nd floor in front of Ballrooms

7:30 – 8:30 A.M. **BREAKFAST**

8:30 – 10:00 A.M. **WELCOME AND KEYNOTE 1: ASSESSMENT TO SUPPORT TEACHING AND LEARNING**
Welcome: CADRE Director Derek Briggs, Dean Kathy Schultz and NCME President Steve Sireci | Keynote 1: Deborah Ball and Lorrie Shepard

10:00 – 10:15 A.M. **TRANSITION TIME**

10:15 – 11:45 A.M. **BREAKOUT SESSION 1**

11:45 – 12:00 P.M. **TRANSITION TIME**

12:00 – 1:30 P.M. **LUNCH AND KEYNOTE 2: ASSESSMENT FOR CURRICULAR EFFICACY**
Angela DeBarger, Erin Furtak and Bill Penuel with discussant Margaret Heritage

1:30 – 1:45 P.M. **TRANSITION TIME**

1:45 – 2:45 P.M. **BREAKOUT SESSION 2**

2:45 – 3:15 P.M. **COFFEE BREAK AND REFRESHMENTS**

3:15 – 4:45 P.M. **BREAKOUT SESSION 3**

4:45 – 5:00 P.M. **TRANSITION TIME**

5:00 – 6:30 P.M. **ROUNDTABLES, POSTERS, HORS D'OEUVRES AND DRINKS**

THURSDAY, SEPT. 19 | DAY TWO | UNIVERSITY MEMORIAL CENTER

7:30 - 12:00 P.M. **CONFERENCE CHECK-IN** | 2nd floor in front of Ballrooms

7:30 – 8:30 A.M. **BREAKFAST**

8:30 – 10:00 A.M. **KEYNOTE 3: MAKING ASSESSMENT RESPONSIVE TO CULTURALLY AND LINGUISTICALLY DIVERSE IDENTITIES**
Megan Bang and Kalehua Krug with discussant Margaret Heritage

10:00 – 10:15 A.M. **TRANSITION TIME**

10:15 – 11:45 A.M. **BREAKOUT SESSION 4**

11:45 – 12:00 P.M. **TRANSITION TIME**

12:00 – 2:00 P.M. **LUNCH PANEL: WHAT'S THE FUTURE FOR THIS CONFERENCE?**
Heidi Andrade, Dale Angney, Derek Briggs, Susan Brookhart, Kristen Huff, Mark Wilson, Caroline Wylie and facilitated by Lorrie Shepard

FULL CONFERENCE SCHEDULE

TUESDAY, SEPT. 17

7:30 - 5:00 P.M. **CONFERENCE CHECK-IN** | 2nd floor in front of Ballrooms

5:00 – 7:00 P.M. **WELCOME RECEPTION AND EARLY CHECK-IN**
Embassy Suites Boulder Patio, 2601 Canyon Blvd, Boulder, CO 80302
Hors d'oeurves and cash bar

WEDNESDAY, SEPT. 18 | DAY ONE

Location: University Memorial Center at the University of Colorado Boulder

7:30 – 8:30 A.M. **BREAKFAST** | *East-Central Ballroom*

8:30 – 10:00 A.M. **WELCOME AND KEYNOTE 1** | *East-Central Ballroom*
ASSESSMENT TO SUPPORT TEACHING AND LEARNING

Welcome: **Derek Briggs**, Director, Center for Assessment, Design, Research and Evaluation, University of Colorado Boulder; **Kathy Schultz**, Dean, School of Education, University of Colorado Boulder; and **Steve Sireci**, President, National Council on Measurement in Education

Keynote 1: **Deborah Ball**, William H. Payne Collegiate Professor and Arthur F. Thurnau Professor of Education, University of Michigan, and **Lorrie Shepard**, Distinguished Professor, University of Colorado Boulder.

10:00 – 10:15 A.M. **TRANSITION TIME**

10:15 – 11:45 A.M. **BREAKOUT SESSION 1**

An Innovative Model to Help Local Educators Learn to Use Formative Assessment Practices
Ellen Vorenkamp[1], Kristy Walters[2], Tara Kintz[3], Margaret Heritage[4] | *West Ballroom*

(1) Wayne RESA; (2) Corunna Public Schools (MI); (3) Michigan Assessment Consortium; (4) Heritage Consulting, Inc.

Several states engage educators to learn about formative assessment practices. These efforts include workshops, print and video resources, online learning opportunities, and websites. Michigan's FAME program offers these resources, but in a different manner: it is based on the belief that significant change in professional practice requires several years of supported learning. No other state provides this long-term learning opportunity for teachers to improve their practice. New FAME learning teams (comprised of a coach and several teachers) begin each year. After attending a full-day Launching into Learning session, they meet regularly for the next three school years for collaborative inquiry about formative assessment practices. FAME work is supported by regional Leads; documents (e.g., FAME Learning Guide); public and secure FAME websites with substantial print and video resources to assist Coaches in conducting their Learning Team; and, MDE- provided program direction, with research & development support provided by the MAC.

Context Matters: The Promise of Cultural and Community Validity in Assessment
Pohai Kukea Shultz[1], Kerry Englert[2], Kalehua Krug[3], Keli'i Ruth[1], Lindsey Miwa Ching[1], Johnalyn Leilani Franco[4] | *Aspen Room*

(1) University of Hawai'i at Manoa; (2) Seneca Consulting; (3) Hawai'i Department of Education; (4) Hawai'i Public Schools (HI)

In this presentation, we will discuss the development of the Kaiapuni Assessment of Educational Outcomes (KĀ'EO), a summative assessment for the Hawaiian language immersion program, and describe the work with teachers to build a joint understanding of cultural validity that embodies key priorities for Hawaiian immersion education. We will also discuss the process of working with teachers to strengthen their understanding of academic content standards and item development so they could return to their classrooms to build assessments that were aligned to the standards and items that reflected important cultural elements. The goal of the presentation is to highlight the power and promise of co-constructing knowledge of culture and worldview with teachers and how those discussions can support building cultural validity in both formative and summative assessments.

Formative Assessment with Interactive STEM Simulations: Strategies Informed by Research
Kathy Perkins[1], David Webb[1], Jonathan Massey-Allard[2], Karen Wang[3], | *Room 247*

(1) University of Colorado Boulder; (2) University of British Columbia; (3) Stanford University

Interactive simulations have become a popular learning technology among science and math teachers, and are particularly known for their ability to actively engage students with dynamic representations that support learning of both STEM content and practices. In this session, we will examine how simulation design and classroom implementation can also create new opportunities for formative assessment and feedback, including with hard-to-assess STEM practices. The session will include 1) a brief introduction to PhET Interactive Simulations, a collection of 150+ free, online math and science simulations, and the key pedagogical affordances of this technology, 2) discussion of new research findings and their implications for sim-based formative assessment strategies, and 3) significant opportunity for participants to collaboratively apply and extend these ideas by examining, discussing, and drafting sim-based learning materials. To explore the simulations in advance, visit https://phet.colorado.edu.

Integrating Learning Trajectories and Formative Assessment into Mathematics Instruction – The Ongoing Assessment Project (OGAP)
Marjorie M. Petit[1], Elizabeth Hulbert[1], Robert Laird[1] | *Room 235*

(1) OGAP Math, LLC

The Ongoing Assessment Project (OGAP) is a formative assessment project founded on mathematics education research on how students learn concepts related to additive reasoning, multiplicative reasoning, fractions, and ratios and proportions. This session will focus on the OGAP Multiplicative Reasoning Framework/Learning Trajectory and how learning trajectories and formative assessment can provide a research-based lens for teachers to understand student thinking, make 'on time' instructional decisions, and provide actionable feedback to students. In addition to the OGAP Multiplicative Framework/ Learning Trajectory, participants will be introduced to aspects of the research that underpins OGAP, learn about the OGAP professional development model and see results of the OGAP Philadelphia Study that showed OGAP impacted both student and teacher knowledge (Supovitz, J., Ebby, C., Remillard, J., &

Nathenson, R. (2018)). OGAP materials and resources are currently being used in classrooms in Alabama, Pennsylvania, South Carolina, New York, Nebraska, New Hampshire, Maryland, and Vermont.

Learning Progressions: Individual Paper Presentations
Moderated by Derek Briggs, University of Colorado Boulder | *Room 245*

Alison Bailey, University of California, Los Angeles; **Margaret Heritage**, Heritage Consulting, Inc.; Anne Blackstock-Bernstein, University of California, Los Angeles; Sandy Chang, Loyola Marymount University; Cindy H.Y. Lee, Stanford University; Eusebio Martinez, Greenfield Union School District; Beverly Eidmann, King City Union School District

Improving English Learner Instruction & Formative Assessment: A Pilot Study of Teacher Implementation of Language Learning Progressions
To support English learners' academic achievement, teachers need to be able to assist students in both language and content learning. Language learning progressions provide an interpretive framework, supporting teachers' inferences about students' current language status so they can take instructional action to advance both language and content learning (Bailey & Heritage, 2014; Black & Wiliam, 1998). We present a five-month pilot implementation study that involved teacher professional learning through a community of practice focusing on elementary students' oral and written explanations. Many teachers began with knowledge of formative assessment but few had knowledge of learning progressions. After the pilot, teachers reported more frequent opportunities for oral discourse, self-assessment, and peer assessment. Qualitative analyses revealed ways teachers changed practices. Classroom observations were consistent with teacher self-reports. While this study reflects themes found in the scalability literature, it also contributes to understanding the specificity needed to effectively teach and assess EL students.

Alison Castro Superfine, University of Illinois at Chicago; Kathleen Pitvorec, University of Illinois at Chicago

Learning-Trajectory-Based Formative Assessment: A Collaborative Inquiry Between Teachers and Researchers
In this session, we will report on our efforts at supporting elementary teachers' understanding of robust formative assessment practices as part of the iFAST Algebra Project. In particular, we will describe the evolution of our work with teachers as our professional learning experiences moved away from delivering professional development to teachers and evolved into professional learning with teachers as co-researchers and co-designers, a model of professional learning based on collaborative inquiry.

Jere Confrey, North Carolina State University; Emily Toutkoushian, North Carolina State University; Meetal Shah, North Carolina State University

Working at Scale on Diagnostic Classroom-Assessments on Learning Trajectories
Math-Mapper 6-8 (MM) is a classroom assessment system designed to measure students' progress along these learning trajectories(LTs), providing real-time data to students and teachers through user-friendly reports. Teachers embed these assessments within ongoing instruction based on their district's curricular sequence . Three years of data from MM's implementation in six middle schools have been analyzed. In this presentation, the team will discuss the results of the first round of validation studies. The presentation will summarize the extent and type of modifications made to the

LTs as a result of the analysis of over 35,822 assessments. Specific examples of the collaborative considerations of learning scientists, psychometricians, and practitioners will also be presented. These analyses provide a foundation for what future research around diagnostic classroom assessments conducted at scale can provide to the field of measurement.

Frederick Peck, University of Montana; Jessica Alzen, University of Colorado Boulder; Derek Briggs, University of Colorado Boulder; Raymond Johnson, Colorado Department of Education

Using Learning Progressions to Support a Classroom Learning and Assessment System: The Learning Progression Framework

In a research-practice partnership, we developed a learning and assessment framework that could be integrated into teachers' practice as assessments for learning, while also serving to provide meaningful information on student growth for summative accountability purposes. Responding to calls to place models of learning at the center of assessment design, our framework is organized around learning progressions (LPs). LPs are the basis for the design and interpretation of assessment activities, they support teachers in using assessment to inform instruction, and they serve as a means for evaluating student learning over time. Participating teachers shifted their assessment and instructional practices to be more attentive to student reasoning, and they reported satisfaction with the framework as a means to infer student learning for accountability purposes. In the session we will share an overview of the framework, tools that support implementation, a summary of findings related to shifts for collaborating teachers, and lessons learned.

11:45 – 12:00 P.M. TRANSITION TIME

12:00 – 1:30 P.M. LUNCH + KEYNOTE 2: ASSESSMENT FOR CURRICULAR EFFICACY
East-Central Ballroom

Angela DeBarger, Program Officer, William and Flora Hewlett Foundation;
Erin Furtak, Professor of Science Education and Associate Dean of Faculty, University of Colorado Boulder; and **Bill Penuel**, Professor of Learning Sciences and Director of the National Center for Research in Policy and Practice, University of Colorado Boulder; with discussant **Margaret Heritage**, Consultant, Heritage Consulting Inc.

1:30 – 1:45 P.M. TRANSITION TIME

1:45 – 2:45 P.M. BREAKOUT SESSION 2

Classroom Assessment in Support of the Next Generation Science Standards
Alicia Alonzo[1], Amelia Wenk Gotwals[1], Kirsten Edwards[1] | *West Ballroom*

(1) Michigan State University

The Next Generation Science Standards (NGSS) challenges science teachers to shift from directing students to "learn about" facts to supporting students in "figuring out" phenomena through science and engineering practices. Classroom assessment is an essential component of this new vision for science teaching and learning. In particular, to support a coherent focus on a core concepts that are revised over time and the integration of those ideas with science and engineering practices, classroom assessment should:

1. Investigate students' use of science and engineering practices to engage with disciplinary core ideas and crosscutting concepts and

2. Attend to longitudinal progress along a learning pathway rather than snapshots of correct/incorrect responses.

In this session, we share how three research projects have partnered with teachers to use classroom assessment to engage students in three-dimensional tasks and obtain evidence of students' learning in terms of pathways. We also discuss challenges entailed in this work.

Constructing Items to Support Classroom Assessment Diagnosticity
Laine Bradshaw[1], Madeline Schellman[2] | *Room 247*

(1) University of Georgia/Navvy Education, LLC; (2) University of Georgia

This session will share efforts to design well-functioning, closed-form items for classroom assessments designed to yield instructionally-relevant and psychometrically-reliable diagnoses of what students do and do not understand. We will first share efforts from an IES-funded measurement project to write closed-form items with incorrect responses hypothesized to manifest due to different misconceptions. For this project, we developed 73 items and conducted 64 cognitive labs in four iterative rounds to collect response process validity evidence. We will then discuss efforts to design multiple-choice items that elicit higher-order reasoning for the Navvy standards-level, on-demand assessment system being piloted by 18 school districts in Georgia under the Innovative Assessment Demonstration Authority (IADA). In both contexts, using practical examples, we will illustrate successes and challenges in balancing efforts to maximize the items' performance—for measuring multiple dimensions and/or high levels of rigor—while minimizing construct irrelevant variance.

Developing Teacher Expertise in Formative Assessment: From Preservice to Teacher
Caroline Wylie[1], Margaret Heritage[2] | *Room 235*

(1) Educational Testing Service; (2) Heritage Consulting

We will explore what can constitute a professional learning continuum (from preservice to induction to in-service) for the development of teacher expertise in formative assessment, a critical lever in improving outcomes for students. To inform professional learning across all phases, we propose a framework to illustrate our thesis that the cultivation of teacher expertise to implement formative assessment effectively is grounded in three core, interrelated domains: formative assessment knowledge and skills, disciplinary knowledge, and habits of practice. The presenters will discuss the three phases of the continuum and situate the work of supporting teacher learning at each phase in the framework. We will engage participants in small-group discussions about how this continuum might apply to their own context, and how they can support coherence across different phases of teacher learning. We will conclude with a report-out session and capture responses to be shared with anyone interested after the session concludes.

Integrating Data from a Balanced Assessment System for Use with Individualized, Small Group, and Whole Class Instruction
Daniel F. Mix[1], Kristen Huff[1], Ellen Forte[2] (discussant) | *Room 245*

(1) Curriculum Associates; (2) edCount

In this session, we will explain how data is generated for the many assessment reports from a balanced assessment system. The assessments include an interim computer adaptive assessment, an assessment for monitoring growth that can be used in an RTI framework, assessments that demonstrate

student mastery of standards, and assessments that can be configured as homework assignments or lesson quizzes. Then we will detail how these assessment data can be used to inform instructional recommendations for individual students, small groups, or whole classes in a core mathematics curriculum throughout the year to ensure that all students have as much remediation as necessary to be prepared for the on-level material, and to help close the proficiency gaps. Finally, there will be a summary of the presentation with a question and answer session.

Standing on the Shoulders of Giants: A Framework for Authentically Involving Students in their Learning and Assessment
Scott Marion[1], Carla Evans[1] | *Aspen Room*

(1) Center for Assessment

There is a renewed interest in authentically involving students in collecting and evaluating evidence of their learning, but we are concerned that current initiatives have not attended to previous research and theory. Additionally, we have expanded the conceptualization from a simple dichotomy of student-led or teacher-directed to continua of multiple dimensions including decisions about the learning goals, learning approaches, demonstrations of learning, and evaluations of learning. We have translated this conceptualization into practice in two projects focused on different grade spans and different use cases, but overlapping in that both are working to incorporate student agency throughout the design and implementation of the curriculum and assessment projects. Our paper, on which this session is based, Standing on the shoulders of giants: A conceptual framework for student-involved assessment and learning, was written to summarize the research literature, the multi-dimensional conceptualization, and current practice.

2:45 – 3:15 P.M. COFFEE BREAK AND REFRESHMENTS | *East-Central Ballroom*

3:15 – 4:45 P.M. BREAKOUT SESSION 3

A Sociocultural Approach for Building Classroom Assessment Literacy Across Multiple Contexts
Scott Marion[1], Jeri Thompson[1], Nathan Dadey[1], Carla Evans[1], Kadie Wilson[2], Monica Ousley[3], Wendy Stewman[4] | *Room 235*

(1) Center for Assessment; (2) SAU 9 School District (NH); (3) Alabama Mathematics and Science Teachers Initiative; (4) Polk County Public Schools (FL)

The Center for Assessment relies on Lave and Wenger's (1991) concept of "legitimate peripheral participation" in our work to improve assessment literacy at scale. Legitimate peripheral participation describes how novice members become fully participating and eventually expert members of a community of practice. This approach is notably different than the more intuitive, but much less successful, train-the-trainer model.

We first used this approach to professional learning with Wyoming's Body of Evidence system 20 years ago and in several states and districts. We continue to refine this approach to support various assessment literacy initiatives including: developing science assessment expertise with New Hampshire's Performance Assessment of Competency Education (PACE), Polk County, FL., the Alabama Mathematics and Science Teachers Initiative, and the Arkansas Department of Education. This session will involve a conceptual

overview, reports from several projects, and an interactive panel discussion among Center for Assessment experts and key practitioners from these sites.

Learning from the Arts to Advance Meaningful Assessment Experiences for Students
Capucine Chapman[1], Barth Quenzer[1], Amy Martinson[1], Janelle Constance[1], Debra Rhoer[1], Jennifer Smith[1], Elena Diaz-Bilello[2] | *Room 247*

(1) Denver Public Schools (CO); (2) University of Colorado Boulder

This organized session panel will present ongoing efforts undertaken by both curriculum specialists and teacher leaders at the Denver Public Schools (DPS) to deepen instructional practices and student learning in Music and Visual Arts. Curriculum specialists and teacher leaders intentionally designed curricular materials and assessments to draw on the diverse socio-cultural backgrounds of students to motivate and help shape student engagement with the creative process in the Arts. Curriculum specialists and two teacher leaders involved with efforts to develop common units and tasks will discuss the challenges and strengths of using this approach to better define curriculum and instructional expectations in these disciplinary areas. The panel will be moderated by the Director of the Arts and PE Department, and a long-time partner at CADRE, University of Colorado Boulder will offer reflections on the presentations as a discussant.

Learning Progressions and NGSS
Derek Briggs[1], Jason Buell[1], Rajendra Chattergoon[1], Clarissa Deverel-Rico[1], Samantha Duwe[2], Ryann Patrick-Stuart[2], Sandy Student[1], Amy Burkhardt[1], Kate Henson[1], Kelsey Tayne[1] | *West Ballroom*

(1) University of Colorado Boulder; (2) Aurora Public Schools (CO)

This presentation describes the early results from the development of an NGSS-aligned assessment system. The system is comprised of a learning progression for modeling energy flows, assessment items and tasks that can be used for both formative and summative purposes, and tools and routines for teachers to apply when making sense of student assessment responses. A core feature is the use of a "cross-cutting" learning progression as a framework for drawing connections between multiple NGSS performance expectations that span the disciplines of physics, chemistry and biology. We describe some of the novel features of our approach to assessment development, the nature of our collaboration with staff and teachers at our partner school district, and the challenge of making "three-dimensional" inferences about student learning that are valid and generalizable.

The Use of Learning Progression-Based Assessments with Students and Teachers
Aurora Graf[1], Sarah Ohls[1], Chad Milner[2], and Bill Crombie[3] | *Room 245*

(1) Educational Testing Service; (2) The Young People's Project, Inc.; (3) The Algebra Project, Inc.

Understanding the concept of function is key to learning higher mathematics. While traditional approaches to mathematics instruction emphasize functions, they tend to provide a brief introduction to the concept, and then focus on real-valued functions. In our project, we are developing a learning progression-based assessment that addresses the function concept through three strands: traditional, finite-to-finite, and geometric transformations. The first speaker will discuss the goals of the project and summarize work conducted thus far. Prior to collecting data from large numbers of students on computer, we conducted cognitive interviews and focus groups with students. Our second speaker will discuss the goals and

procedures for the cognitive interviews and focus groups, and our third speaker will discuss observations about the language used in the assessment and how students in the cognitive interviews and focus groups responded to it. Our final speaker will focus on teacher professional development related to the work.

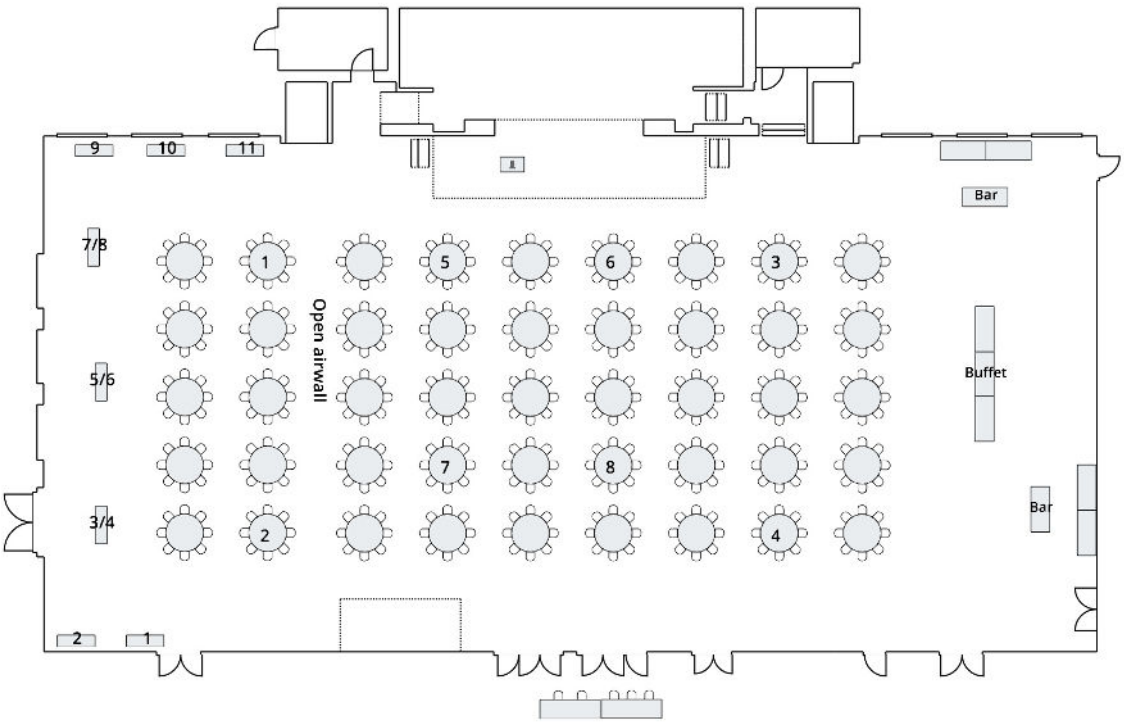
Using Design Principles to Measure Enactment in Project Based Learning
Alison Gould Boardman[1], Robert Laurie[2], Ashley Potvin[1], Karla Scornavacco[1], Kristina Stamatis[1]
| *Aspen Room*
(1) University of Colorado Boulder; (2) Poudre School District (CO)

In project based learning (PBL) educators develop practice around projects and inquiry, connecting learning to real world challenges, and engaging youth in activities that feel authentic and meaningful. Yet, assessing learning in PBL classrooms can be difficult as skills such as 'critical thinking' are not easy to measure. In this session we will share our PBL design principles in the areas of authentic making, collaboration, feedback and revision, reflection, social and emotional learning, and universal design for learning. These design principles were used to develop a variety of classroom enactment and student assessment measures and to ensure coherence across curricular materials, professional development, and instruction in high school language arts. We will also discuss ways to engage students in creating assessment criteria and reflecting on learning. This session invites educators to consider ways to apply design principles across grade levels and contexts.

4:45 – 5:00 P.M. **TRANSITION TIME**

5:00 – 6:30 P.M. **ROUNDTABLES, POSTERS, HORS D'OEUVRES, AND DRINKS**
Combined Ballroom
Bartenders will be checking photo IDs for individuals ordering alcoholic drinks

Ballroom Diagram: Numbered roundtables in the middle and numbered posters along the side



Roundtable 1: Chair and Discussant, Kristen Huff, Curriculum Associates

Jonathan D. Bostic, Bowling Green State University; **Hilary Steinmiller**, Perrysburg Schools (OH); Gabriel Matney, Bowling Green State University; Toni Sondergeld, Drexel University; Gregory Stone, Metriks Amerique

Gathering Response Process Validity Evidence and Influencing Classroom Instruction in the Process

Assessment development can function as a form of teacher education and influence teachers' instruction and students' outcomes. This presentation provides evidence of how partnerships for assessment development have potential to strengthen classroom assessment development and teachers' classroom instruction. Assessment development for a grant-funded project led to a mutually beneficial experience for all, including school administrators, teachers, and students. The presenters for this session are a university faculty member and a school administrator where the assessments have been piloted. We will engage with attendees around ways to facilitate school partnerships for assessment development and in turn, leverage activities around assessment development to impact classroom instruction. Similarly, school personnel provide necessary and helpful feedback about assessments and ways to construct meaningful tools for teaching and learning. This presentation has implications for school-university partnerships.

Sarah M. Bonner, Hunter College; **Peggy P. Chen**, Hunter College, Brandon Milonovich, Ardsley High School (NY); Kristi Jones, High School of Hospitality Management (NY)

Self-Regulated Learning and Classroom Assessment: An Integrated Framework with Proof of Concept

We present results of a pilot study of assessment design using the CA:SRL framework, which infuses a well-known model of self-regulated learning (SRL) into classroom assessment (CA). Working in high school computer science classrooms, we added explicit prompting in SRL forethought and reflection to tasks that required application of computational thinking concepts. We sought evidence that a) students were learning from assessment, demonstrated by SRL responses, and b) the CA tasks were an appropriate model for assessment of computational thinking. We analyzed the work of 12 students in terms of performance outcomes, SRL, and cognitive process evidence gathered through think-alouds. We found that students increased in self-awareness of computational thinking concepts immediately following task completion. We found considerable variation in computational thinking practices that was not captured by performance scores. Implications for CA are discussed, with particular attention to the need for process evidence in measurement of computational thinking.

Roundtable 2: Chair and Discussant, Sue Brookhart, Duquesne University

Jessica Alzen, University of Colorado Boulder; **Casey Kilpatrick**, Colorado Education Association, **Catherine Lee**, Adams 12 Star School (CO); Brittany Osborn, Poudre School District (CO)

Peer-Led Professional Development about Classroom Assessment—Examples from the Colorado Professional Institute for Learning Online

In this presentation, we will describe the Colorado Professional Institute for Learning Online Together (COpilot) model for professional development supported by the Colorado Education Association (CEA). The COpilot Director will discuss the program generally and its adoption across the state of Colorado

as well as specific courses designed to help teachers with classroom assessment. In addition, we will discuss results from a descriptive evaluation of COpilot. Findings from the evaluation include evidence of all elements of effective professional development as outlined in the research literature, participant satisfaction with COpilot, and changes made to classroom practices generally and assessment practices specifically because of participation in COpilot courses. Discussion will center on how online peer-led PD can foster change in assessment practices by supporting collaboration across grade-level as well as district and geographic region.

Roundtable 3: Chair and Discussant, Alicia Alonzo, Michigan State University

Aurora Graf, Educational Testing Service; Peter W. van Rijn, Educational Testing Service:

Development and Empirical Recovery for a Learning Progression-Based Assessment of the Function Concept

In this presentation, we will discuss findings based on student responses to an LP-based assessment for the Concept of Function in mathematics. Each task developed for the assessment consists of several items around a common scenario. The LP consists of three strands, "Traditional," which focuses primarily on real-to-real functions, "Finite-to-Finite," which focuses on mappings from finite sets to finite sets, and "Geometry," which focuses on geometric transformations. We collected cognitive interview data from students responding to the tasks, in preparation for the development of a computer-delivered pilot assessment. We will discuss highlights from the cognitive interviews, and the implications of those findings for both the soundness and the validity of the LP. In addition, we will discuss preliminary findings from the pilot, in which data from 1102 students were collected. The pilot included technology-enhanced task types that allowed students to enter equations, Cartesian graphs, arrow diagrams, directed graphs, and matrices.

Emily Toutkoushian, North Carolina State University; Jere Confrey, North Carolina State University; Meetal Shah, North Carolina State University

Exploring the Relationships Between Learning Trajectory-Aligned Classroom Assessments and State-Wide Summative Assessments

Investigating the relationship between conceptually rich, theory-based classroom assessments and high-stakes assessments is both critical for establishing evidence of the larger impact of these assessments and a difficult task due to the embedded nature of classroom assessments within instruction. This presentation utilizes longitudinal data from four years of use of a learning-trajectory aligned middle school mathematics diagnostic classroom assessment system, MathMapper 6-8 (MM), as well as state-wide summative data, to explore different ways of defining and interpreting the connections among these assessments. The scale and scope of MM presents a unique opportunity to track teacher practices and student understanding over the curriculum in an academic year and across years. Regression analysis suggests significant relationships among performance and usage variables in MM, including scores and counts of assessments and years using MM, and growth on end-of-course tests. Further analysis looks at growth related to performance in specific content areas in MM.

Amy Elizabeth Cardace, Cornell University

Applying a Learning Progression Framework to Classroom Assessment in Elementary Science
This paper examines the design process for creating a classroom assessment system that can serve both summative and formative purposes. The context was a newly-developed science curriculum teaching concepts of micro-evolution to second- and third-graders at two public schools. Instruction was based on a learning progression that integrated findings from the research literature, classroom observations, and structured interviews. We consider two assessments: first, and considered the "gold standard," a structured pre-post interview, and second, a written alternative. Our analyses show that a written assessment in this context can elicit meaningful information, and we discuss how it compares to the summative information from the interview. Both assessments map to the learning progression, providing a framework in which student responses can signal a need for particular instructional support. We will describe the underlying learning progression, assessment design, data analyses, and examples of how student responses could be used to inform teaching.

Roundtable 4: Chair and Discussant, Erin Furtak, University of Colorado Boulder

Heidi Andrade, University at Albany-SUNY

A Critical Review of Theory and Research on Student Self-Assessment
This presentation will provide an updated overview of theory and research on student self-assessment. The treatment of theory will include the introduction of a refined definition, taxonomy, and operationalization that is designed to ensure that self-assessment serves the purposes of learning. The review of empirical research will provide a critical perspective in the interest of provoking new investigations into neglected areas that might ensure that self-assessment can be used effectively in educational contexts to support learning and its self-regulation. Recommended areas of inquiry include the influence of the standards or criteria used by learners during self-assessment, the cognitive and affective mechanisms of self-assessment and their influences on student decision-making about next steps, and the key components of effective self-assessment, especially social-emotional components related to power and trust.

Andrew Martin, University of Colorado Boulder

Using Two or Multistage Formative Assessments: Effects of Cognitive Challenge, Student Preparation, and Group Interaction
Multistage exams provide an opportunity to assess different aspects of student learning gains. In this roundtable, I introduce a simple quantitative framework for partitioning the variance in student performance on exams into three different components: the effect of individual student knowledge (individual effect), the effect of students working in small groups (group effect), and the effect of the cognitive challenge of individual questions (question effect). The data are generated using two-stage exams implemented using IF-AT formatted exams and the data visualized and analyzed in ways that help instructors assess student gains in their content understanding and in their collaborative abilities. The roundtable will go through the different stages of implementing this type of assessment, examine some data from a two-stage exam, and provide a context for discussions about the power and pitfalls of multi-stage exams for simultaneous formative and summative assessment.

Aarti P. Bellara, University of Connecticut; Alexandra Stone, University of Connecticut

"Oh, I Get It Now". Closing the Research to Practice Gap in Assessment Literacy

Being assessment literate (Popham, 2011) is a vital component to effective teaching. However, there has been little to no research on the efficacy of teacher training in assessment (Greenberg & Walsh, 2012). This empirical case study examines the pedagogical practices that aim to close the research to the practice gap in assessment for preservice teachers. More specifically, this study analyzes an assignment that places the preservice teacher in the role of a researcher, where they were to make connections between classroom assessment theory and research, and classroom assessment practices. Preservice teachers made sense of the interviews by identifying the (mis)alignment between what teachers do in practice and what they are taught in an assessment course in a teacher preparation program. In this study, we as the researchers, analyzed this assignment, identifying broad themes and areas of (mis)alignment between theory and practice the preservice teachers reported.

Roundtable 5: Chair and Discussant, Caroline Wylie, Educational Testing Service

Diana Akhmedjanova, University at Albany-SUNY; Heidi Andrade, University at Albany-SUNY

Feedback as an Integral Part of the Self-Regulated Writing Intervention

The mixed design, combining the single case quasi-experimental design with semi-structured focus group interviews, was used to examine the effects of the instruction of writing persuasive essays and self-regulating behaviors to international freshmen (n=8) in a medium-sized research university. Students were exposed to three types of formative assessment during the study: self- , peer-, and teacher feedback both in and outside of the classroom. The results based on focus group interviews and SRL journals suggest that students valued the peer-assessment aspect of the course the most along with the feedback from the teacher, which is not surprising and replicates previous research findings (Goldstein, 2016; Wiliam; 2013). Hence, feedback from different sources encouraged students to set detailed proximal learning goals and employ refined writing and progress monitoring strategies than at the beginning of the course. Self-assessment was not as instrumental with this group of students, possibly due to the lack of practice.

Angela M. Lui, University at Albany-SUNY; Heidi Andrade, University at Albany-SUNY

Inside the New Black Box: Operationalizing and Measuring Student's Responses to Teacher Feedback

The purposes of this project were to develop and validate the Responses to Feedback (RtF) survey and examine how students respond to feedback as measured by the survey. Responses to feedback is operationalized as the internal mechanisms of feedback processing, which is built on three assumptions: (1) the reception of assessment information involves both cognitive and affective mechanisms, and decisions on how information is used to inform next steps; (2) initial motivational states influence how students respond to feedback; (3) responses to feedback can be indirectly accessed through surveys and prompted self-reports. Considering this operationalization, think alouds and RtF survey administration were conducted with students during their initial exposure to their teacher's feedback. For the roundtable, we will present on the validity and reliability of the RtF survey, including evidence based on test content, response processes, and internal structure.

Zachary Feldberg, University of Georgia; Shelbi Kuhlmann, University of Georgia

Enhancing Monitoring Accuracy and Learning with Various Forms of Feedback on Learner-Generated Drawings

This study investigated how various forms of feedback on student-generated drawings impacts students monitoring of their learning throughout a one-week science lesson and how this effect ultimately helped or hampered their learning. One hundred ninety-seven eighth grade students were randomly assigned to a formative or summative feedback condition. Students either received elaborative or corrective (formative) feedback on each of their drawings throughout the week or received corrective (summative) feedback on their drawings at the end of the week. All students completed pre- and post-tests that assessed retention and transfer.

Kerri Wingert, University of Colorado Boulder; Kris Kilibarda, Iowa Department of Education

A Framework for Assessing Student Talk

Opportunities for youth to make sense of their learning should be the cornerstone of classroom life (e.g., Bang et al., 2016), and science education has been shifting to focus on this sensemaking as central. This roundtable session takes up the issues of validity and equity in the assessment of sense-making discourse in classroom life. What do we already know about assessing talk? What cases of use might inform the design of tools? What design principles seem to emerge around assessment of student talk from science education literature? Attendees of this session will engage in shared conversation about the future for assessment of student talk in practical contexts.

Roundtable 6: Chair and Discussant, Rick Stiggins, Assessment Training Institute, Pearson

Jessica Weber, Denver Public Schools and **Jennifer Smith**, Denver Public Schools

Transformation in Assessing Movement

An ongoing concern in Physical Education has been the issue of defining authentic assessment experiences. In the past, widely used assessments in the field stressed the importance of evaluating mainly the physical gains made by students in various fitness or sport-based activities at two different time points.

The Arts and Physical Education Department contend that authentic assessment in these areas need to expand beyond the traditional focus of physical gains and allow other opportunities for students to display their learning. This session will focus on different and innovative ways in which teachers use assessment to engage students through activities that require them to: 1.) display cognitive understanding of cues necessary for skill acquisition 2.) engage in self- and peer-assessments to develop skills in providing constructive feedback; 3.) strategize and plan with other students; 4.) understand how and why particular activities connect back to important principles outside of the school setting.

Sonya T. Stephens, Wake County Public Schools (NC); **Spencer Ziegler**, Wake County Public Schools; Mellisa Smith, Wake County Public Schools; Serena Halstead, Wake County Public Schools

Using Performance Tasks to Measure the 4Cs

Equipping students with the necessary skills to be successful in the 21st century is a challenge that

most school districts are currently facing. As a result educators are seeking best supports for students to practice the 4Cs (critical thinking, collaboration, communication, and creativity). This discussion will focus on how one large urban school district is using performance tasks designed to directly measure aspects of the 4Cs, specifically communication and critical thinking, which are part of how the district defines readiness for college, career, and citizenship. During the round-table discussion the process used for developing, piloting, administering, and scoring the performance tasks will be shared. Additionally participants will engage in a discussion about the findings and the methods used to communicate the results. The district will also seek feedback on effective uses of the data collected from the performance tasks and additional ways to socialize the results.

Roundtable 7: Chair and Discussant, Scott Marion, Center for Assessment

Brian Gong, Center for Assessment

Matching Models for Instructional Uses with Assessment Designs

Assessments are most powerful and useful when designed intentionally for particular purposes and uses. I discuss five instructional uses as examples of the need to specify "classroom assessment use" in order to inform assessment design and validation. The illustrations are: a) predicting end of year performance on a state assessment; b) declaring "mastery" under a competency-based education model; c) eliciting "pre-assessment" information about student knowledge prior to instruction; d) using information from a comprehensive assessment system to perform "increasingly focused diagnosis to inform instructional remediation," and e) evaluating instruction and curriculum to improve for the next instructional cycle. My main points are that "classroom assessment" would benefit from more specific claims and use case scenarios, and that different assessment designs will be needed to provide appropriate information. I also show sample interpretive arguments and validation arguments, similar to what have been advocated for large-scale and high-stakes assessments.

Michael W. Klymkowsky, University of Colorado Boulder; Melanie M. Cooper, Michigan State University

Characterizing Exam Questions Using the 3D Learning Assessment Protocol as a Way to Improve Course Design and Outcomes Assessment

Often missing from the evaluation of course design and efficacy is an examination of the measures, typically exam questions, used to assess students' learning. An AAU funded project at Michigan State University recognized that exam questions provide useful insights into what a particular course/instructor values and measures. These are messages not lost on students. Their approach was to analyze questions on the basis of whether their answers involved the components of "three-dimensional (3D) learning", that is the inclusion of scientific practices, cross-cutting and discipline-specific concepts. They developed the "3D learning assessment protocol" (3D-LAP)(Laverty et al., 2016). A 3D-LAP analysis allows us to visualize what types of learning is valued in a particular course (see Matz et al., 2018). As part of a parallel project at UC Boulder, we plan to characterize exam questions given in the context of a "transformed" and "conventional" courses in order to provide feedback to instructors. We will discuss the basic process of applying a 3DLAP analysis together with insights garnered, specifically the challenges and compromises involved.

Philipp Sonnleitner, University of Luxembourg; Carrie Kovacs, University of Applied Sciences Hagenberg

Improving Fairness of Classroom Assessment Through a Self-Administered Questionnaire: the Fairnessbarometer

Although classroom assessment is at the core of every teacher's daily activities, teachers' assessment literacy is often in need of improvement. Consequences of suboptimal assessment practices in turn have a strong impact on students' thinking, feeling, and their actions. The present study aims to explore whether a psychometric sound, self-administered questionnaire could be used as basis to establish problem awareness among teachers concerning their classroom assessment habits. The administered questionnaire, the Fairnessbarometer, covers informational, as well as procedural facets of fairness. Teachers and their students answer the same aspects of assessment practices but from different angles. Exploration of the resulting discrepancy-profiles revealed certain problem types with some teachers differing from their students' perception in almost every rated aspect, some showing single aspects that require improvement, and others demonstrating almost identical responses to the addressed fairness aspects. Results clearly indicate the potential of the Fairnessbarometer for teacher training and self-development.

Pamela Ann Harvey, University of Colorado Boulder

Assessment of Course-Based Undergraduate Research Experiences: What Do Students Learn?

Students who participate in apprentice-based research activities experience increased graduation rates and persistence in science careers, and invest in science as a life-long learning process. However, it is not possible to scale these one-on-one experiences to serve the needs of all undergraduate students. A solution is to formalize research experiences within departmental curricula. Course-based Undergraduate Research Experiences (CUREs), an innovative approach to teaching scientific experimentation, can efficiently achieve this goal. However, despite the increasing interest in CUREs over the past two decades, a cohesive description of the successful approaches for CUREs across STEM fields has only recently been presented (Corwin, et al., 2015). Using this model, we will discuss a range of instruments that are aligned with common learning goals in CUREs. We will also use large-scale CUREs at University of Colorado Boulder as a framework for exploring customization of these assessments.

Roundtable 8: Chair and Discussant, Jim Popham, University of California Los Angeles

Christopher D. Andrews, Indiana University; **Edward Roeber**, Michigan Assessment Consortium; **Angela Landrum**, Colorado Department of Education; and **Jeri Thompson**, Center for Assessment

Creating Sustainable and Scalable Approaches for Enhancing Educators' Assessment Literacy

Efforts addressing assessment literacy for education practitioners in the United States has become more and more critical. Assessment is an essential part of the educational ecosystem and influences essential educational decisions ranging from education policy affecting an entire nation or state to teachers deciding whether additional educational support is needed for a particular student. Despite assessment's prevalence, understanding which assessments to use, how and when to use them, and what decisions can be appropriately made based on their results is a difficult task. K-12 assessment literacy efforts need to work with educators to address the relevant and context-specific issues they

are currently facing in their schools and classrooms and have an explicit goal of building capacity and sustainability. This roundtable session brings together diverse efforts of university researchers, non-profit organizations, and public education administration toward designing and implementing effective and sustainable assessment literacy for K-12 educators.

Poster Presentations | *West Ballroom*
Opening remarks to poster session by Douglas Watkins, Denver Public Schools

Poster 1: P. Shawn Irvin, University of Oregon, Leilani Sáez, University of Oregon

Assessment-guided Decision-making in Prekindergarten: A Reconceptualization of Seminal Practices

Current classroom assessment and instructional practices across preschool and early elementary systems are often misaligned, lacking a coherent approach to ensuring all children begin kindergarten equitably. Improving between-systems alignment, by introducing developmentally appropriate, yet systematic, approaches to literacy learning could reduce children's risk for early reading difficulties by strengthening foundational competencies in prekindergarten. In this poster, we present a preschool model of classroom-based assessment for instruction comprised of early childhood adapted K-12 evidence-based practices reconceptualized as an assessment-guided approach to decision-making for prekindergarten classrooms. We operationalize the model in an assessment-curriculum system that complements preschool teachers' typical practices and supports their capacity to measure, evaluate, and monitor learning to inform and adjust their instruction. We present promising results that show enhanced growth in pre-reading skills and learning behaviors for children in implementing classrooms, and a substantial end-of-year reduction in the number of children identified as at-risk for reading disability.

Poster 2: Richard Lambert, University of North Carolina at Charlotte; Bryndle Bottoms, University of North Carolina at Charlotte

Examining the Fidelity of Implementation of Formative Assessment for Young Children

Teachers of young children often have difficulty implementing formative assessment systems to fidelity (Ferrara & Lambert, 2015; 2016). Four strategies for identifying teachers who are struggling with implementation were developed: 1.) an implementation fidelity measure, 2.) a teacher interview protocol, 3.) criteria for evaluating evidences of child progress, and 4.) criteria for evaluating the validity of placements made on developmental progressions. The fidelity measure includes 18 rubrics covering preconditions for successful implementation, and five components of formative assessment (Heritage, 2013): 1.) selecting learning targets, 2.) developing criteria for success, 3.) eliciting evidence of learning, 4.) interpreting evidence of learning, and 5.) adapting and responding to learning needs. Examples are provided, based on statewide data, that demonstrate how the strategies were used to help coaches identify struggling teachers, and differentiate support for those teachers based on the areas of belief and practice with the greatest potential for growth.

Poster 3: Bill Penuel, University of Colorado Boulder; Kerri Wingert, University of Colorado Boulder

Collaborations to Promote Validity and Equity in Classrooms and Large-Scale Assessments

This poster showcases a newer form of practical measurement (Yeager et al., 2013) : that of student

experiences throughout our curriculum. We demonstrate our process and the ways that our data gathering allows for disaggregation along racial markers to understand whether our curriculum is working to include all students in science assessment, in addition to supporting all students to learn.

Poster 4: Joshua Wilson, University of Delaware; Gaysha Beard, Red Clay Consolidated School District (DE); Yue Huang, University of Delaware; Charles MacArthur, University of Delaware

A Research-Practice Partnership Aimed at Improving Writing Outcomes via Implementation of Automated Writing Evaluation

This study examined whether the large-scale implementation of Automated Writing Evaluation software in elementary schools had a positive association with students' writing quality, writing attitudes and writing self-efficacy, and performance on state English Language Arts and writing tests. The participants were approximately 2000 students in Grades 3-5 and 185 writing teachers from 14 elementary schools in a mid-Atlantic school district. Findings from hierarchical linear models indicated that only a few of AWE-usage predictors had statistically significant effects on students' writing outcomes. However, identified relationships are promising. Findings from this study inspire implications for other school districts that are considering implementing AWE systems to improve students' writing outcomes.

Poster 5: Jonathan Dings, Boulder Valley School District (CO)

Impact of a Complex, Performance-Based Algebra Assessment on Classroom Assessment

States and districts adopt or develop assessments with an aim to influence instructional practice, including classroom assessment. This case study examines the impact on classroom assessment of a district developed performance based algebra assessment, as judged through the classroom assessments selected or developed by teachers in response to a new requirement. Teacher response provides a jumping off point for an exploration of the effectiveness of using assessment to drive practice, as well as the question of whether teaching to a complex performance based assessment mitigates concerns usually associated with teaching to the test.

Poster 6: Holly Rutledge, Cobb County School District (GA)

Assessment Auditing for Improved Teaching & Learning

The Cobb County School District (CCSD), the second largest public school district in Georgia, developed an easy-to-use, research-based, school-level assessment auditing process to monitor use of their balanced assessment system. This process includes reflection and planning, an assessment inventory, identification of which assessments are administered as well as their frequency of administration, and identification of how assessment data are actually used. School leadership and district staff then review and discuss the completed assessment inventory, identifying areas of strength and areas of needed improvement, and develop a plan of action. In addition, a single assessment audit process was developed which focuses on the overall quality of a specific assessment. Both auditing processes will be discussed, and reflections from a pilot implementation will be shared.

Poster 7: Douglas G. Wren, Virginia Beach City Public Schools (VA)

Performance Tasks for Deeper Learning: Meeting the Challenges

Well-designed performance tasks not only provide opportunities for students to hone critical-thinking and problem-solving skills in scenarios that mirror real life, they also yield evidence so teachers are able to assess their students' proficiency in the skills. The pedagogical potential of performance tasks is widely recognized; however, from a psychometric standpoint, they can be messy.

Developing performance tasks that measure higher-order thinking and support deeper learning requires both higher-order thinking and high degree of assessment literacy on the part of the developers. This poster presentation illustrates best practices for designing, implementing, and scoring such performance tasks by highlighting the major challenges of each step and describing how to overcome specific obstacles. The presenter, a former classroom teacher and current educational measurement & assessment specialist, has led a large Virginia school district's efforts to assess students' critical thinking, problem solving, and written communication for the past 10 years.

Poster 8: Mark Hudson, Colorado Christian University

Effective Assessment in Music Education: Student Self-Assessment of Progress Towards Student-Friendly Learning Objectives

Assessment of learning and assessment for learning are of significant importance in today's educational culture. Student involvement in the learning process figures prominently in the Colorado State Model Evaluation System for Teachers, required in Level 4 and 5 Practices. Student self-assessment is arguably essential to quality assessment in music, and has been identified through various research studies as a means to enrich musical understanding, aesthetic sensitivity, critical-listening skills, and engagement. However, student self-assessment in the traditional ensemble setting has been little explored, particularly with regard to measures of individual student learning. This poster presentation outlines one method by which students in a traditional music ensemble class may be directly and authentically involved in monitoring and regulating their own progress. This method has been field-tested by the presenter, with positive results. These results, along with the relevant background and methodology, will be displayed and discussed during the presentation.

Poster 9: Chelsea Nehler, University of Kansas

Teacher Perspectives on Student Mastery: Implications for Diagnostic Assessment Use and Design

The Dynamic Learning Maps alternate assessment based on alternate achievement standards serves 90,000 students with the most significant cognitive disabilities in 18 states. DLM assessments are scored using a diagnostic classification model (DCM), which determines the student's mastery of discrete skills and concepts. In addition to their accountability purposes, the assessment results are intended to be used by educators to facilitate instructional decision-making. How teachers themselves view the students' skills and knowledge, in particular what constitutes mastery, may affect how they interpret and use assessment results. In this study, we asked teachers to share how they thought about mastery while completing a skill ratings task for a specific student. In some cases, teachers' views of mastery aligned with mastery as defined by the DCM; in other cases, teachers

provided more nuanced pedagogical views of mastery. The implications of these differences are discussed in light of their students' assessment results.

Poster 10: Adelina Valiquette, Queen's University

Exploring Elementary School Teachers' Approaches to Mathematics Assessment in Response to Diverse Groups of Students

Classroom assessment has become an integral part of teaching and learning. While student achievement continues to be measured in relation to academic standards, the central purpose of classroom assessment has shifted from a focus on measuring past learning, to a focus on supporting future learning (Cowie, 2013). This poster reports on a year-long case-study of 14 mathematics teachers (Grades K-8) in Ontario, Canada. The purpose of this research was to explore how these teachers vary their assessment practices based on their students' unique learning needs and how these different approaches are influenced by contextual factors both inside (e.g. classroom culture) and outside (e.g. school policies) the classroom. Results of this study highlight the need for ongoing collaboration and support for teachers as they strive to align their pedagogical approaches with our contemporary view of classroom assessment as a responsive, ongoing and context-dependent social practice.

Poster 11: Brent Duckor, San José State University; **Carrie Holmberg**, San José State University; **Mark Wilson**, University of California, Berkeley

One Framework, Two Cases, Many Pathways for Research

It is largely taken for granted that teachers must become more assessment literate. We are interested in identifying teachers' understanding and use of classroom assessment knowledge along a continuum, or "teacher learning progression" (TLP) from an IRT-based psychometric perspective.

We present findings from a preservice "assessment literacy" construct modeling exploration in three major domains (NRC, 2001). We examine reliability and validity evidence of the CAL instrument based on Testing Standards and its diagnostic potential for evaluating preservice teachers (Duckor, Draney, & Wilson, 2017).

We also present findings from a construct modeling exploration of preservice teachers' formative assessment "moves" with three TLPs: posing, pausing, probing. Construct maps, video-based episodes, and progress guides are used to evaluate student teachers in clinical placements (Duckor & Holmberg, in press).

By articulating the "4 Building Blocks" approach to assessment literacy research, we invite conversation around evidence-centered design for scholars and practitioners interested in advancing classroom assessment practice.

THURSDAY, SEPT. 19 | DAY TWO

Location: University Memorial Center at the University of Colorado Boulder

7:30 - 12:00 P.M. **CONFERENCE CHECK-IN** | 2nd floor in front of Ballrooms

7:30 – 8:30 A.M. **BREAKFAST** | *East-Central Ballroom*

8:30 – 10:00 A.M. **KEYNOTE 3: MAKING ASSESSMENT RESPONSIVE TO CULTURALLY AND LINGUISTICALLY DIVERSE IDENTITIES** | *East-Central Ballroom*

Megan Bang, Professor of Learning Sciences, Northwestern University;
Kalehua Krug, Educational Specialist, Hawai'i State Department of Education;
with discussant **Margaret Heritage**, Consultant, Heritage Consulting Inc.

10:00 – 10:15 A.M. **TRANSITION TIME**

10:15 – 11:45 A.M. **BREAKOUT SESSION 4**

Using Performance Assessment to Demonstrate Educators' Effectiveness
Edward Roeber[1], Heather Vaughn Southard[1], Kathy Dewsbury-White[1] | *Room 247*

(1) Michigan Assessment Consortium

Many states require student performance to be a significant portion of teachers' evaluations. The Michigan Arts Education Instruction and Assessment (MAEIA) program developed performance assessments in dance, music, theatre, and visual arts programs K-12 students that arts educators can use to demonstrate their effectiveness by using worthwhile, challenging, and engaging assessments.

Michigan Assessment Consortium (MAC) staff worked with volunteer arts educators and supervisors to select student exemplars, document instruction provided, and assemble a brief, coherent demonstration of their effectiveness. Five topic guides (displaying data, documenting instruction, and administrator "look fors" that show what good arts instruction looks like) were developed and piloted with arts educators and state arts education associations. These resources will be made available through short online courses for educators. In addition, in a MAEIA Institute, pairs of administrators and arts educators implement the MAEIA resources by instructing and assessing students, and supporting and evaluating educators.

Enhancing Assessment Literacy:
Two Textbook Authors Examine the Definition and Enhancement Tactics
Rick Stiggins[1], Jim Popham[2] | *Room 235*

(1) Assessment Training Institute, Pearson; (2) University of California, Los Angeles

Assessment is the process of gathering evidence of student achievement to inform educational decisions. Often, assessment systems have been guided by unsound policies leading to unsound practices resulting in inaccurate evidence and inappropriate decisions, all due to a lack of assessment literacy among policy makers and practitioners. Two leaders in the development of assessment literacy among teachers, school leaders and policy makers will interact with session participants on the keys to promoting assessment literacy. Both presenters have authored introductory textbooks widely used in teacher and school leader preparation. They will argue about differences in their approaches and affirm points of agreement. But their primary mission in this session is to engage the audience in discussions to seek clarity about such

matters as the definition of assessment literacy, assessment skills and competencies underpinning assessment literacy, barriers to its development, and actions we can take to remove these barriers.

Teacher, Assessment Director, and Researcher Perspectives on Implementing a Classroom Assessment Tool Based on LTs
Meetal Shah[1], Norman Alerta[3], Janice Noll[3], Derek Briggs[2] (Discussant), **Jere Confrey** (Moderator) | *West Ballroom*

(1) North Carolina State University; (2) University of Colorado Boulder; (3) Cherry Creek School District (CO)

Teachers, assessment directors, and researchers hold unique perspectives on data-driven instruction and rarely have opportunities to share and learn from each others' expertise. In this presentation, each panelist shares some of their experiences with data-driven practice, relating specifically to an ongoing use of data from a learning trajectory-aligned classroom assessment system, Math Mapper 6-8 (MM) to inform various levels of instructional decision making (e.g. class, grade-level professional learning communities (PLCs), and administrative). One of the fundamental purposes of MM is to provide teachers with valid and reliable learner-centered data on how students in their classes are progressing towards acertain mathematical concept rather than binary information about whether they can or cannot meet a content standard. The audience will gain an insight into data return processes through video examples and panelists' reflections on how data return processes interact with the use of measures of classroom assessment.

A Teacher Learning Progressions Approach to Understanding and Improving Teachers' Formative Feedback with English Learners: Empirical Studies with Preservice Teachers
Brent Duckor[1], Carrie Holmberg[1], Adria Patthoff[2], Nicole Barnes[3] | *Aspen Room*

(1) San José State University; (2) University of California, Santa Cruz; (3) Montclair State University

Researchers from public university teacher preparation programs will share findings about preservice teachers' (PSTs) conceptualizations of and practices regarding formative feedback (FF) with K-12 students in general and English learners in particular. The empirical studies involved 140+ PSTs serving in high needs, multilingual contexts. **San José State University** will present findings from the California Teacher Education Improvement Network (CTERIN) study of PSTs' uses of FF strategies in a single subject credential program. **UC Santa Cruz** will present findings from the CTERIN study of PSTs' uses of FF strategies in a multiple subject credential program. **Montclair State University** will present findings from a study examining what survey responses from elementary and secondary PSTs serving in high-poverty, urban contexts revealed about how PSTs prioritized assessment-related knowledge and skills, including FF practices. We invite conversation among researchers and practitioners interested in novice enactment of formative feedback in various contexts.

Math and Science: Individual Paper Presentations:
Moderated by David Webb, University of Colorado Boulder | *425*

Michele Carney, Boise State University; Katie Paulding, Boise State University

Assessing Cognition in Mathematics to Inform Instruction: Evidence Needed to Support Inferences

This paper in the session will provide background on research related to assessing students' cognition in mathematics (i.e., what students do and do not understand as opposed to correctness of a

response) and the type of evidence necessary to support inferences related to student cognition. This will be done through an example situated in proportional reasoning. Evidence from individual cognitive interviews and the framework for analysis will be presented. Once the background and examples are presented, session attendees with engage in a discussion around (a) the realities of assessing cognition with relatively efficient items in terms of administration, scoring, an interpretation, and (b) the type of evidence educators should require from test developers, or gather for themselves, to support valid inferences related to student cognition.

Caitlin Fine, University of Colorado Boulder

Asset- and Deficit-Oriented Language Use During Science Formative Assessment Co-Design Cycles

The Framework for K-12 Science Education calls for broadening access to and participation in K-12 science (NRC, 2012). This is critically important in the United States (US), where schools increasingly serve students from diverse linguistic backgrounds (NCELA, 2018). Research has indicated that teachers' expectancy of these students' outcomes informs their instructional and assessment decisions, and those, in turn, are related to students' achievement (Lopez, 2017; NASEM, 2018). This study presents a content and discourse analysis of monthly video data from biology teacher formative assessment co-design meetings at a public high school in the Western US. Through applying a coding system focused on the way teachers discuss diverse students' language use and participation in science learning, the paper illuminates the ways in which formative assessment co-design can surface - and have the potential to counteract - teachers' deficit framings of linguistically diverse students.

James McMillan, Virginia Commonwealth University

Development and Validation of the Perceptions of Science Classroom Assessment Student Self-Report Survey

The purpose of this project's research and development efforts is to create and validate a secondary student self-report measure of perceptions toward science classroom assessment tasks. The presentation will include an overview of the importance of student perceptions of assessment, and show how three sources of information impact the conceptual framework: 1) existing theories of formative assessment, feedback, and self-regulation; 2) existing measures of student perceptions of assessment; and 3) recommendations of Next Generation Science Assessments for authenticity and engagement, fairness, formative uses, and student autonomy and responsibility. Based on these sources of information, five triangulated components are recommended for the survey: alignment, authenticity, challenge, fairness, and feedback. Working in partnership with a local school division, initial teacher reactions to the conceptual framework, components, and sample survey items will be summarized.

Jeffrey B. Bush, University of Colorado Boulder; Krista Marks, Woot Math

Technology Mediated Classroom Assessment: Co-Design of Features and Tasks for Student Response Systems and Effective Pedagogy in Algebra

There is widespread agreement on the benefits of using formative assessment to inform teaching, and many teachers strive to create more student-centered learning environments where student ideas and voices drive the activity of the classroom. Still, dilemmas of practice exist that keep these goals from being realized. Using a framework of technology-mediated classroom assessment, this study seeks

to better understand the role of tasks and features of digital student response systems in helping high school algebra teachers implement reform-oriented pedagogy and process-oriented formative assessment. Teachers use a student response system to run co-designed formative assessments; analysis of these sessions reveals how design elements of technology can interact with assessment practices and pedagogy. The study draws on student interview data to capture their perspective on the interactions between technology, learning algebra, formative assessment and dispositions to learn.

11:45 – 12:00 P.M. **TRANSITION TIME**

12:00 – 2:00 P.M. **LUNCH PANEL: WHAT'S THE FUTURE FOR THIS CONFERENCE?**
East-Central Ballroom

Heidi Andrade*, University at Albany; **Derek Briggs**, University of Colorado Boulder; **Sue Brookhart***, Duquesne University; **Kristen Huff***, Curriculum Associates; **Mark Wilson***, University of California, Berkeley; **Caroline Wylie***, Educational Testing Service; facilitated by **Lorrie Shepard**, University of Colorado Boulder

*Members of the NCME Special Conference on Classroom Assessment Task Force. Other members are Neal Kingston, University of Kansas; Alison Bailey, University of California, Los Angeles; Jade Caines Lee, University of New Hampshire; Dale Whittington, John Carroll University