

Developing Systems of Assessment

Joan Herman

NRC Workshop

Washington, DC– September 13, 2012

National Center for Research
on Evaluation, Standards, & Student Testing

UCLA | Graduate School of Education & Information Studies | cresst.org



Overview

- Assessment systems: two views
- Examples of each
- Validity issues
- Implementation challenges



Systems Serving Multiple Purposes

1. Formative: Integrate instruction and assessment to facilitate significant science learning
2. Summative: multiple measures to provide comprehensive, standards-based view
3. Multi-level: feedback to serve accountability and improvement for multiple stakeholders



Coherent Systems: Formative and Summative



End of the Year Assessment
Builds K-12 to College Readiness

Examples of Integrated Systems

- SimScientists
- IQWST
- 3-Dimensional Fused Knowledge
- Display-based reasoning
- AP Exam

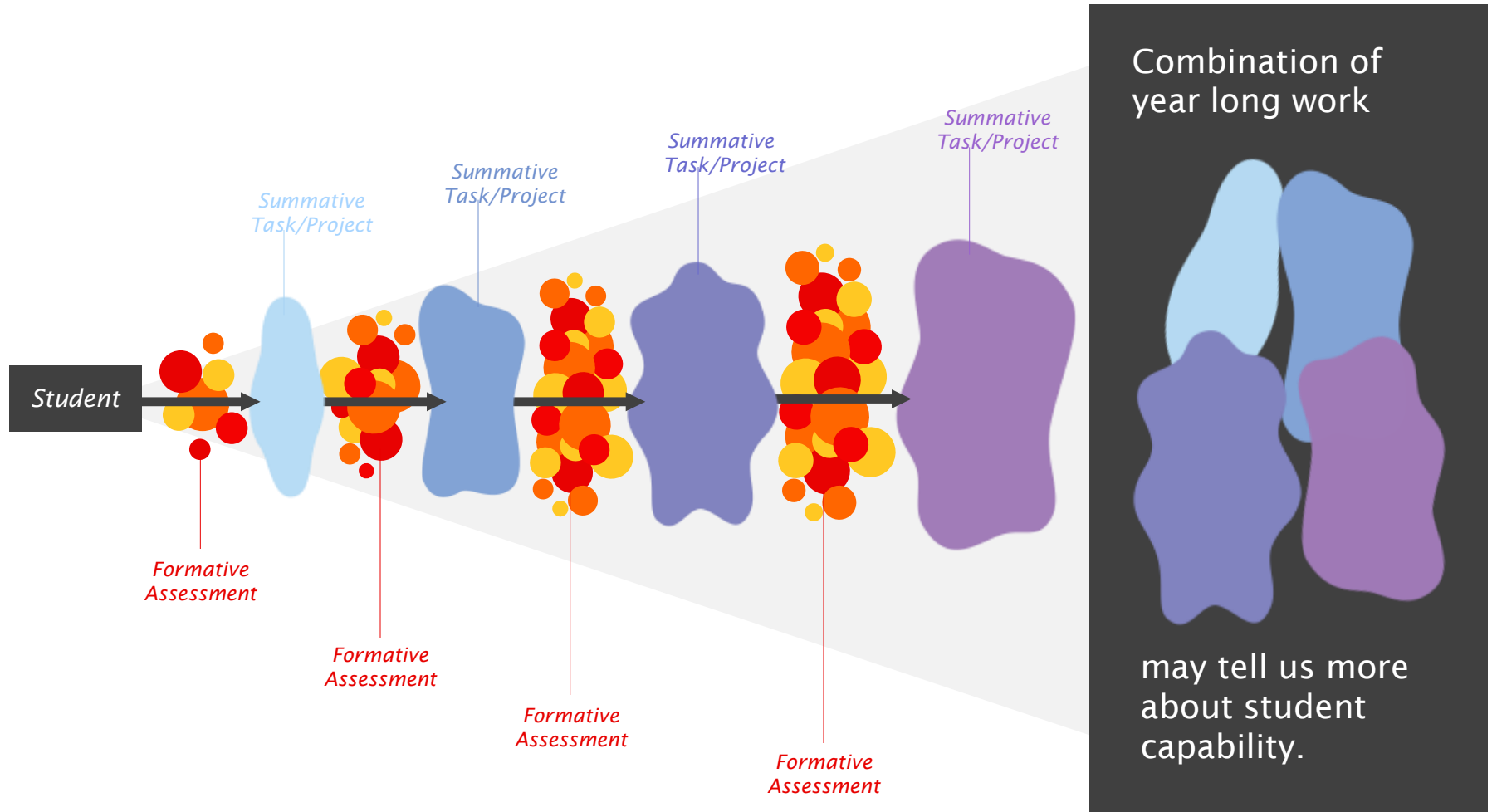


Validity Criteria for Coherent Formative Systems

- Focused on rich goals
- Learning-based
- Aligned/coherent
- Reliable
- Valid
- Accessible/fair
- Diagnostic
- Actionable
- Credible/meaningful
- Consequences



Coherence: Accountability and Classroom Evidence



System Examples Combining Classroom and On-Demand Tasks

- RI Diploma System
- Queensland
- Both rely largely on educator judgment



Rhode Island High School Diploma Requirement (2014)

- District designed, standards aligned
- Proficiency requirements, based on two of three:
 - ✓ *Exhibition*
 - ✓ *Portfolio*
 - ✓ *Comprehensive end of course exams*
- State assessment score at least partially proficient
- Rigorous criteria and quality control



Portfolio Tasks

- Variety of task types
 - ✓ *On-demand task*
 - ✓ *Extended task*
 - ✓ *Mini-exhibition*
 - ✓ *Exhibition*
- Developmental progression/rigor
 - ✓ *Classroom*
 - ✓ *School*
 - ✓ *PBGR*



Exhibition: Senior Project

- Applied learning goals:
 - ✓ *Problem solving*
 - ✓ *Communication tools and techniques*
 - ✓ *ICT tools and techniques*
 - ✓ *Self management*
- Personalized, choice
- Project examples: Core question -→ design a product/service, improve a system, trouble shoot, plan and organize an event/activity



Queensland Senior Certificate

- State establishes content standards/goals for 2-year syllabus in each content area
- Local schools develop specific, locally adapted course of work, specifying content, instruction and assessment (formative and summative) activities
- Local course accredited by external review board
- Teachers evaluate their students on a 5-point standards-based scale
- Teacher ratings reviewed through moderation cycles
- General, standardized test supports moderation and comparability



Queensland Assessment

- Varied task types:
 - ✓ *Extended experimental investigations: 4 weeks +*
 - ✓ *Extended response tasks involving research and secondary data: 2 weeks +*
 - ✓ *Supervised (on-demand) tasks*
- Curriculum-based, formative activities, with feedback, build to summative success



Queensland Assessment

- 11th grade “monitoring” tasks build to more demanding 12th grade “verification” tasks
- Final assessment based on work products distributed over time
- Exit standards assessed through principles of
 - ✓ *fullness (fully represent expected standards)*
 - ✓ *Latest (recent accomplishment supercedes)*



Validity Criteria for Multi-purpose Systems

- Meaningful, transferable scientific knowledge, practices and predispositions
- 3 C's: Comprehensive, Coherent, Continuous
- Instructionally sensitive
- Valid/fair for each intended purpose
- Useable/useful for intended purposes
- Feasible
- Consequences



Challenges in Moving Ahead

- Coherence of standards' tri-part progression
- Availability of instructional resources
- Teacher capacity
- Technical dilemmas
- Socio-political dispositions





National Center for Research on Evaluation, Standards, & Student Testing

UCLA | Graduate School of Education & Information Studies

VISIT US ON THE WEB

cresst.org

you@cse.ucla.edu