Paper #3
Abstract Title Page.

Title: Designing Dynamic and Interactive Assessments for English Learners That Directly Measure Targeted Science Constructs
Author(s): Dr. Rebecca Kopriva[1], David Gabel[2], Catherine Cameron[2]
Affiliations of authors: 1: University of Wisconsin, Wisconsin Center for Education Research, School of Education, University of Wisconsin-Madison,1025 West Johnson Street, Suite 785, Madison, Wisconsin 53706; 2: Center for Applied Linguistics, 4646 40th Street NW, Washington DC 20016
Author's emails: rkopriva@wisc.edu, dgabel@cal.org, ccameron@cal.org
Contact email for paper: rkopriva@wisc.edu

Abstract Body

Background / Context:
Recent trends are towards computerizing large scale assessments, and this movement has been accelerating with the RTTT assessment grants and the establishment of SMARTER Balanced and PARCC consortia. As we move towards widespread computer delivery of assessments, there is an opportunity not to be missed of not only improving what assessments can measure, and increasing the speed of feedback, but also improving the accessibility and validity of these assessments for ELLs and others.
This presentation will explain and illustrate how computer-based innovative test tasks are designed in a multi-semiotic environment to effectively and comprehensibly convey meaning to all students, especially English language learners (ELLs), students with learning disabilities (LDs), selected other students with disabilities, and non-identified native English speaking poor readers, giving them the tools they need to express complex thinking in mathematics and science. With text playing a peripheral rather than primary role, multi-semiotic representations of context, content, and task demand are combined with innovative student response mechanisms to engage cognitive schemas. These response mechanisms allow students to show what they know and can do by manipulating various stimuli for the purposes of modeling, assembling, classifying, drawing and using relational templates to demonstrate complex causal inferences. ONPAR uses this methodology to create tasks that leverage the dynamic and interactive capacities of computers to convey meaning to and from the test taker not only during the delivery of the task demands but also in how students are equipped to give evidence of their learning. Being able to legitimately express knowledge and skills is one of the most difficult barriers for students with language or attention challenges when language and language structures typically associated with more complex skill levels are not accessible. Several ONPAR projects in elementary, middle school and high school math and science have focused on allowing these student groups to demonstrate their understanding of more complex concepts, skills and reasoning at a level commensurate with how students proficient in reading and writing English explain their understandings. While the ONPAR projects are using these methods in computer-interactive environments, teachers can and should routinely deploy similar methods in their content classrooms to better measure what all students, including these focal groups, can do. These kinds of applications will be discussed.

Purpose / Objective / Research Question / Focus of Study:
This presentation will explain and illustrate how computer-based innovative test tasks are designed in a multi-semiotic environment to effectively and comprehensibly convey meaning to students, especially English language learners (ELLs), students with learning disabilities (LDs), selected other students with disabilities, and non-identified native English speaking poor readers, and give them the tools they need to express complex thinking in mathematics and science.

Setting:
Research was conducted in elementary and middle school students.

Population / Participants / Subjects:
Research was conducted for math and science ONPAR projects with elementary and middle school students. A series of cognitive labs conducted with individual students early in the projects to refine the ONPAR approach.

For the science project, focal groups included all levels of English language learners (ELLs) as well as the control group of non-identified native English speaking students. A total of 513 fourth-grade and 468 eighth-grade students from 26 schools in five states took part in the ONPAR science study.

For the math project, the focal groups were extended to also include students with learning disabilities (LDs), other students with disabilities that the teachers thought might benefit, and non-identified native English speaking poor readers, as well as the groups included in the science project. A total of 1048 fifth-grade and 1258 eighth-grade students from four states took part in the ONPAR math pilot.

Intervention / Program / Practice:
Interactive and Dynamic items are built within four environments which form an organizational and conceptual framework within ONPAR's approach. These environments – 1) introduction 2) problem environment 3) task question and 4) response environment – are made explicit through multi-semiotic representations of content and context that focus the test taker on the item target and minimize construct irrelevant variance.
Four techniques are used within these environments: 1) Multisemiotic representations used to replace large amounts of text, and build meaning and complexity across screens 2) Multisemiotic representations used to focus students on item target 3) Varied item types used to directly measure the construct 4) Intuitive user-interface design used to minimize the construct irrelevant variance associated with figuring out how to interact with the item.

Text is minimal, mainly used for precision purposes, and supported multi-semiotically through static visuals, animations, focusing techniques and audio in English or the students' first academic language.

Research Design:
Cognitive labs were conducted with individual students to assess the efficacy of the ONPAR approach. A "think aloud" approach was used to assess the accessibility of the items and to study student cognitive processes as they responded to the items. Each cognitive lab session included an interviewer, and an observer, and followed a protocol created to elicit the desired information.

ONPAR Science:
An experimental study with equivalent groups randomly assigned fourth- and eighth grade students to one of two randomly assigned science test forms: (a) a traditional full-text form consisting of existing state items from which the ONPAR items had been built, or (b) an ONPAR form that used complete sentences, onscreen support, and audio translation of text in the student's native language. Test forms were delivered on laptops transported to the schools. The ONPAR items used dynamic simulated contexts and asked students to demonstrate what they knew by manipulating stimuli in various ways. At each grade, each the forms consisted of 11 items. The traditional fourth-grade form included 10 multiple-choice and 1 constructed-response items, while the traditional eighth-grade form included 9 multiple-choice and 2 constructed-response items. ONPAR forms each contained 11 item tasks. Students were assigned to one of four groups for testing based on their level of English language proficiency as measured by the

WIDA-developed ACCESS for ELLs® assessment: (a) *low proficiency ELs* (Level 1 or 2 out of 5), (b) *mid-proficiency ELs* (Level 3), (c) *high-proficiency ELs* (Level 4 or 5), and (d) *non-EL controls*.

ONPAR Math
The ONPAR math pilot was a similar design to the Science study. The study used a randomized design to investigate how students performed on 4th and 7th grade ONPAR tasks as compared to the traditional test measuring the same content. The focal groups included students who were classified as ELLs, learning disabled (LD), native English speaking poor readers with no IEPs (poor readers), and other students with disabilities who teachers felt may have been helped by the interactive items. The forms included 19-21 intermingled ONPAR and traditional items which were randomly assigned to participating students. The forms were delivered remotely through internet browsers.

Data Collection and Analysis:
Student responses from the two studies were collected and scaled across forms and results for the focal groups were analyzed and compared to the control groups within forms and within groups between forms. Additional measures, including teacher ratings of students' content ability on specific constructs measured by the test questions, and state math and reading data were used to validate the results.

Cognitive Lab results were used to inform item development.

Findings / Results:
Results from the randomized experimental study in science indicate that, when academic ability was controlled, low English proficient ELLs and non-ELLs in both grades 4 and 8 performed similarly on tasks which utilized ONPAR techniques, while there were significant differences in how the low ELL students performed on more traditional approaches. On the other hand, there were no significant differences between how non-ELL students performed on either form. Viewing these results as a whole, these findings argue that the ability of low ELL students can be measured by this method at a level that is on par with their non-ELL peers. Further, it argues that there might be flexibility in the ways ELs can be tested without advantaging them relative to scores obtained from traditional test formats used by non-ELLs. A large number of cognitive labs contributed substantively to the continued refinement of the ONPAR techniques.

The experimental study in mathematics was completed in winter 2010. Preliminary results show that when adjusted for student math ability, the interaction between test type (ONPAR versus traditional) and reading level (poor vs. good) is significant with a combination of LDs and poor readers but not for native English speaking good readers, and that when controlling for math ability the LD/poor readers and good readers performed similarly on the ONPAR tasks. Correlations from the experimental study and several cognitive labs have also suggested that several of the ONPAR techniques are actually better at capturing data about complex skills and concepts than traditional methods for all students.

Another randomized project with high school students in biology and chemistry is currently underway, with pilot testing scheduled for winter 2011/12.

Conclusions:
This presentation will explain and illustrate how computer-based innovative test tasks are designed in a multi-semiotic environment to effectively and comprehensibly convey meaning to students, especially English language learners (ELLs), students with learning disabilities (LDs), selected students with disabilities, and non-identified native English speaking poor readers, and give them the tools they need to express complex thinking in mathematics and science. The ONPAR projects in elementary and middle school math and science have focused on allowing these student groups to demonstrate their understanding of more complex concepts, skills and reasoning at a level commensurate with how students proficient in reading and writing English explain their understandings..