Using Cognitive Labs to Refine Technology-Enhanced Assessment Tasks and Ensure their Accessibility: Insights from Data Collected to Inform ONPAR Elementary and Middle-School Science Task Development

Laura Wright, Center for Applied Linguistics
Rebecca Kopriva, University of Wisconsin

In 2007 to early 2008 ONPAR project staff conducted cognitive labs in three different school districts. The purpose of the labs was to inform item development and determine if the multi-semiotic methods being considered were viable for assessment purposes. Specifically, the labs were designed to evaluate if and how particular item methods were successfully measuring the expected content knowledge, skills and abilities, and if and how students understood what was being asked of them. Because the ONPAR items focused primarily on more challenging content, the first objective focused on if and how non-text semiotic representations were successfully conveying the intended meaning, how relevant and irrelevant cognitive density was either facilitating or distracting students from the intent of the items, and how patterns in effective methodology related to different kinds of item targets. The latter objective addressed four principal issues: overall item accessibility, success of on-screen supports, item transiting and navigation, and students' computer skills.

Method

Fifty-eight students in grades 4-9 completed the cognitive labs. A majority of participating students were native Spanish speaking ELs with proficiency levels 1 and 2 of 6 (in other words, students with low level of English proficiency), but native English speaking students were also interviewed. Science ability estimates were also collected from the teacher. Participating students ranged from mid to high in science ability, with

over-sampling of low English proficient students who teachers rated as demonstrating high science ability. This over-sampling was deliberate in order to distinguish between content related ambiguities in the tasks versus language issues. Using a cognitive interview protocol developed by the ONPAR research team, interviewers interacted with individual students as they worked through the ONPAR items. English speaking students were interviewed in English; Spanish-speaking students were asked if they would like to be interviewed in English or in Spanish by a fluent Spanish speaking interviewer. Students worked through each screen individually. Interviewers asked the students questions at the end of each screen, focusing on how the students interpreted the meaning of multi-semiotic information, and, when a response was required on a particular screen, asked students to explain their problem solving process and why they answered as they did. Besides the protocol questions, interviewers asked students open-ended questions to clarify answers when needed.

While the cognitive labs focused on evaluating the ONPAR items, each student interview period also included one or two traditional items (selected from state or federal publically available released science items) interspersed with the ONPAR items. Traditional items were included to compare accessibility and processing strategies across traditional and ONPAR item versions.

Besides focusing on particular item methods, and as part of addressing overall access to the intended meaning of the tasks, the cognitive labs were also used to evaluate multiple variations of item design. In some instances, we tested two variations of the same item to find how the design, pacing, and multiple vs. single screen layout impacted the way that students approached and understood the items. The labs also evaluated how students'

computer faculty impacted their comfort level with the screen layout and how well versions of the pre-test tutorials oriented students to the ONPAR methodology.

Results

The results of the cognitive labs were used during development of the ONPAR tasks in order to help refine their content and design. Several key findings emerged. First, results suggest non-text stimuli could be often used as *the* primary communication vehicle, rather than only or largely acting in a supporting role to text. This included conveying meaning about the problem to the students and students reporting that many of the non or low-text response opportunities they interacted with could adequately convey what they know and can do. The non-text stimuli included static visuals, animations and simulations, continual interactive opportunities, and response spaces where students could assemble, drag and drop, draw, or create using isolated words, phrases, or visuals. Second, again and again the lab results showed that reducing language in test items, did not, by itself, render the items accessible. Rather, a coherent approach to designing items was required if much of the meaning load was to be conveyed using non-text approaches. Such designs, from introduction to the completion of the last response, had to be considered a priori, and multi-semiotic representations were purposefully placed to maximize the communication of meaning while minimizing the language load.

Third, assuming language support as explained below, text *was* found to be the best way to communicate the precise question or statement of the problem that was requested of the student. While context, problem building, and response environments could convey meaning successfully in ways that were largely text free, the staff found that the specific statements or questions that students responded to were best conveyed

via supported English text. Additionally, by systematically varying the amount of English text used in select ONPAR tasks while at the same time holding all other parts of the tasks the same, the researchers found that students on the whole performed better with complete, succinct sentences where selected context-relevant, target irrelevant, words were supported via rollovers and were accompanied with an audio recording. Fourth, the incorporation of the speaker icon into the screen directly to the left of each question or problem statement was found to be effective, which, when pressed, allowed students to hear the question or statement in English or their native language. However, language elsewhere on the screen could be supported by rollovers as necessary, but audio support was typically not needed.

Fifth, the project staff found that incorporating animated, interactive 'help' icons on the response screens of the tasks increased student accessibility. The purpose of these icons was to show students how to functionally interact with the screen to provide their response. Students activated those icons more than any other support on the screen. Sixth, a standardized layout of all ONPAR tasks and each type of screen were found to be more helpful to students than tasks where order of information changed. In the effective layout there was some type of introduction, some type of problem-building 'event', and one or more question and response screens. Seventh, the labs found that animation length was a factor. Brief animations appeared to improve communication of particular information while long animations taxed student interest. For more lengthy animated demonstrations of problems, various techniques have been devised to maintain student focus. Over tasks, researchers also noted that students generally proceeded from the top left of the screen to the bottom right and that layouts that reflected this action facilitated the successful

communication of meaning. Finally, the cognitive labs revealed the importance of a 10-minute interactive tutorial comprised of multiple practice items.