Multimodality and measurement: A discourse analysis of English learners' interactions with traditional and multisemiotic test tasks

Laura J. Wright
University of Wisconsin-Madison

The inclusion of English Learners (ELs) in standardized testing in the United States has been of concern since the reauthorization of Elementary and Secondary Education Act under No Child Left Behind. This concern has grown as computerized tests have begun to be used in the majority of U.S. schools. The modalities computerized tests afford pose new potentialities and constraints for supporting ELs' understanding. This paper examines ELs' interactions with traditional and innovative test tasks developed as part of the ONPAR Science project. Using a discourse analytic perspective, video data of students interacting with test materials is investigated to examine how ELs understood assessment items. We propose that meaningfully including ELs in standardized testing requires understanding the semiotic modalities of the test and their respective affordances, and how modalities create potentialities and constraints for ELs' interactions.

Keywords: assessment; English learners; cognitive interviewing, discourse analysis, multimodality

Word count: 8,370 words

Introduction

In the United States, the reauthorization of the Elementary and Secondary Education Act (ESEA) under No Child Left Behind (NCLB) had serious ramifications on assessment policies and practices for English Learners (ELs) (Abedi, & Gándara,2006; Kopriva, 2008). Most notably, this policy meant that ELs were assessed rigorously in English language proficiency as well as content areas. Prior to NCLB, many ELs did not participate in annual content testing because of the language barriers that these assessments posed (Rivera & Collum, 2006). It was widely accepted that if ELs could not understand the language of a test, they could not be validly and reliably tested.

Testing of ELs remains a concern today because U.S. schools have begun to use new computerized assessments. Two consortia, Partnership for Assessment of Readiness for College and Career (PARCC) and Smarter Balanced intend to measure students' understanding of the new college and career readiness standards. Given how challenging the new standards are, as well as the medium of the test, these assessments pose even greater questions about the inclusion of ELs in standardized testing.

Because of the high stakes attached to these standardized tests and consequences for schools and districts when students' scores do not meet expectations, new ways to validly and reliably test ELs have been sought. The research presented here reports on students' interactions with innovative test tasks developed as part of the ONPAR Science project. ONPAR uses an access-based framework (Carr, Kopriva, & Rex, 2007; Kopriva, 2008) and replaces typically linguistically complex test items with a variety of multisemiotic features such as graphics, animations, multiple languages, and sound. These features provide affordances (Gibson, 1979; Kress, 2010) that have the potential to allow ELs greater access to the test items and meaningfully interact with them. Using a discourse analytic perspective, video data of students interacting with traditional items and

ONPAR test tasks during cognitive interviews is examined to investigate how students with varying degrees of language proficiency understood the assessment items. Specifically, the discourse analytic notions of *trouble* and *repair* (Schegloff, Jefferson & Sacks, 1977) are used to identify trouble sources during the interviews to investigate what aspects of traditional items students find confusing in comparison to the multisemiotic test tasks. We propose that meaningfully including ELs in standardized testing requires understanding and accounting for the multisemiotic resources offered by the tests, their affordances, and how ELs rely upon these resources when interacting with test items. Accounting for the potentialities and constraints of the modalities can better inform test design for this group of learners.

Background

Assessment of ELs and accommodations

Standardized test scores have long indicated that ELs' scores are lower than their English proficient peers in all subject areas, including science (Lee & Fradd, 1998; National Center for Educational Statistics, 2001; National Science Foundation, 1994). Some research suggests that achievement gaps can be partially attributed to traditional assessments confounding language proficiency and content knowledge (Abedi, Leon, & Mirocha, 2003; Francis, Rivera, Lesaux, Kieffer, & Rivera, 2006; Kieffer, Lesaux, Rivera, & Francis, 2009; Kopriva, 2008; author, 2012; Ruiz de Velasco & Fix, 2002; Wolf, Herman, & Dietal, 2010). In other words, content tests use a large amount of language to express the meaning of an item and many ELs have difficulty identifying what an item is about before even trying to answer it (author, 2012; Martiniello, 2008). This makes it difficult to know whether an EL's performance can be attributed to understanding content material or to what extent language proficiency also plays a role. Language is the primary modality of most standardized tests and because it is often inaccessible to ELs, it constrains their meaning making potential in the test environment.

Because of this, researchers have suggested a number of approaches to accommodating ELs in testing situations (Abedi, Hofstetter & Lord, 2004). An accommodation is defined as an alteration to either the assessment itself or its administration procedures (Rivera, Collum, Shafer Willner, & Sia, 2006). The purpose of the accommodations is to help ELs overcome the barrier posed by the modality of the test. Typical accommodations for ELs include the use of bilingual dictionaries and glossaries, English dictionaries, and allowing extra time. However, many more accommodations exist; Rivera and her colleagues report (2006) that there are over 75 different accommodations used in the United States, though many of these are not appropriately tailored to the needs of ELs. An appropriate accommodation must take into account what factor affects an EL's performance on a test, but yet is not construct relevant to the assessment (Francis, et al., 2006). Thus, because language is the non-construct relevant factor affecting the performance of ELs on content tests, an appropriate accommodation must provide either direct or indirect linguistic support.

While many different types of accommodations are currently used, the effectiveness of different approaches is questionable. In a meta-analysis on the effectiveness of different accommodations strategies for ELs, Kieffer et al. (2009) report that the only accommodation that has been shown to improve the performance of ELs is the use of English dictionaries and glossaries. Moreover, this

accommodation strategy is most effective when ELs are also offered extra time to complete the assessment and have practice with this strategy during instructional time. While other strategies are popular such as translating tests into students' native language (Bowles & Stansfield, 2008; Fortuny, et al., 2005), or simplifying the language of the assessment (Abedi, Courtney & Mirocha, 2003; Abedi & Lord, 2001), these methods have met with mixed results (Hudicort-Barnes, Noble, Roseberry, Suarez, & Warren, 2008). Because all communication is multimodal (Kress & van Leeuwen, 2001), understanding how other modalities may be used to support ELs' understanding in testing environments may be a promising approach. Cognitive interviewing provides a methodology for investigating how student understanding via different modalities can be accounted for in the design of multisemiotic test items intended for linguistically and culturally diverse students.

The access-based approach

An approach that has been used to design multisemiotic test tasks is the *access-based framework* (Kopriva, 2008). This approach advocates for changing structural and contextual factors of test items as relevant to make content more accessible (Carr, et al., 2007). Access-based test items provide alternative means to "access meaning, solve problems, and demonstrate solutions without lessening the rigor of the item or changing the construct being measured …" (Carr, et al., 2007; p.8). The overall aim is to design a test form that is parallel to the content of a traditional test so that scores are comparable for all students. That is to say, the constructs measured across test forms remain the same, but construct-irrelevant characteristics of an item may vary across the different forms.

The access-based framework draws upon principles of universal design techniques such as using plain language, plain formatting, a reduced reading load, and visuals, yet it does not adhere to any of them strictly. In other words, developing access-based items does not require following a checklist (e.g., only using simple present tense), but rather is a process in which all design principles are taken into account. The adaptations made to test items using the access-based framework address the linguistic support needed by ELs. Recent research in the area of test accommodations suggests that drawing upon these design principles in developing assessment items is effective for increasing students' access to test content (Kopriva, 2008; author, 2010).

ONPAR

The ONPAR-Science project took place from 2007-2011 in the United States. As part of the projects more than 50 science test tasks for Grades 4 and 8 were developed, as well as high school biology and chemistry. Tasks were designed based on conceptually difficult released test items from state and national tests. ONPAR tasks measured the same science constructs as the traditional items so that they could target the same challenging science content. In order to investigate students' understanding of the constructs across test formats as well as elements of visual design and computerized testing, a series of cognitive interviews was planned as part of the research and design phase (author, 2009). Throughout this phase, cognitive interviews were conducted to determine if/when ONPAR techniques were effective with ELs. In

addition, randomized experimental trials were conducted to compare outcomes of traditional items and ONPAR tasks.

ONPAR test tasks utilize multisemiotic features to convey information to test takers. Semiotic modalities are typically considered to be visuals or graphics, action, sound, and language (Kress, Jewitt, Ogborn & Tsatsarelis, 2001; Kress & van Leeuwen, 2001). Each semiotic modality offers different affordances for meaning making. The term affordance refers to the potential uses of an entity (e.g., an object, picture) that arise from its perceivable properties or its design (Gibson, 1979; Kress, 2010). The goal of using different modalities is to minimize the challenges that the modality of language may pose and to provide alternative modalities that afford opportunities to "access meaning, solve problems, and demonstrate solutions without lessening the rigor of the item or changing the construct being measured" (Carr, et al., 2007; p. 8). Because meaning making in science is multisemtiotic (Kress, et al., 2001; Lemke, 1990, 1998; author, 2008), it is appropriate to assess meaning in science assessment contexts in a multisemiotic fashion.

As shown in the image below, ONPAR tasks include: (1) graphics and animations, (2) text prompts with hyperlinked word and phrase support in the form of graphics and symbols, (3) a speaker button that provides an oral English reading or native language translation of the prompt, and (4) an animated help button to demonstrate how to physically respond to an item (see Kopriva, 2009 for further description of ONPAR features).

[Insert figure 1 here]

Graphics and Animations

Graphics and animations may be considered a primary means of conveying information in ONPAR tasks because of their size and location on the screen; graphics and animations are at the center of the screen and occupy the greatest amount of screen space. The ONPAR test directions also prompt students to focus first on imagery when looking at an item. As such, imagery is intended to activate a student's schema (Bransford, 1984). When students initially see the graphics in an ONPAR task, they should think about the topic that is being tested. That is not to say that imagery is the only means of providing information to students; the relationships among the modalities convey overall meaning to students. However, graphics and animations are integral to the meaning of the task.

Written Language

Written language is another important multisemiotic feature. Many ONPAR screens contain a text prompt at the top of the screen. The prompts function to maintain precision in what an item is asking or support other stimuli. In other words, text prompts help specify the meaning of visuals and ask students to answer a question. In research on visual literacy, textual information has been found to help students in understanding the meaning of imagery, especially if it is difficult to interpret (Ametller & Pinto, 2002). Thus, text prompts can help make ONPAR's visuals clearer to students, while visuals may help support written language.

ONPAR test prompts are developed with several guiding principles. Prompts are typically only one sentence, non-construct relevant terminology contains hyperlinks to provide additional support to students in the form of graphics,

animations, or symbols, and complex grammatical constructions such as various types of clauses, and shifts in tense are avoided. Written language is reviewed by a linguist to ensure that it is worded as simply as possible, and a content specialist to ensure it still conveys accurate scientific information.

Spoken language (Sound)

In addition to imagery and written language, ONPAR provides optional sound through a speaker button. The speaker button either provides the text prompt read aloud in English, or in one of several other languages such as Spanish or Korean. At the beginning of the test, students choose the language they would like to use during the test. The sound button is an optional feature that students activate by choice with each question. For example, if a student sees a prompt and wants to hear it, he/ she simply presses the speaker button. Providing an oral version of the text prompt in English or a student's native language is intended to support the meaning of the prompt, which is relevant to the construct tested. For some students, the oral version may be redundant, but for others, a native language translation may provide necessary information. For example, a student may know some words in English while other words in their native language. The translation allows them to draw upon both languages to make meaning.

All of ONPAR's mulitsemiotic features are intended to work together in complementary ways to enhance ELs' understanding of the tasks. While the tasks are designed to provide access to ELs, the project needed evidence of student understanding of tasks vis a vis these features. The following section discusses how the project undertook this.

Methodology

As part of the research and design phase of the ONPAR project, 14 rounds of cognitive interviews were conducted to examine how students understood the multisemiotic test tasks. Analyzing ELs' interactions with test tasks at this level was fundamental to designing appropriate test tasks for them. A total of 161 students participated in cognitive interviews.

In order to conduct the cognitive interviews, schools were contacted about the ONPAR project and asked to participate voluntarily. Schools then assisted in recruiting current and former ELs for the study. Students received informational letters and permission slips to take home to their parents. Students were also asked to provide assent to the interviews prior to all sessions. Spanish-speaking students were asked if they preferred be interviewed in Spanish by a Spanish speaking interviewer.

Using a researcher-developed cognitive interview protocol, interviewers interacted with individual students as they worked through items. The protocol contained the same basic questions across all items in order to see trends in interpretive strategies. Interviewers asked students about each item, focusing on interpretations and problem solving strategies. Students worked through items individually and were asked to explain their responses retrospectively after each screen so that the cognitive load posed by an item was not increased by language processing (Bowles, 2010). Interviewers asked students open-ended questions to clarify answers when needed. Each interviewee interacted with approximately five of ten ONPAR tasks. At many of the interviews, students also responded to one traditional constructed response item.

An observer was present during each interview to take notes and video record the interview. A sample of approximately 50 interviews from the larger corpus was transcribed using Atlas.ti software. This allowed for analysis of student discourse, coding across interviews, and analysis of visual data. After data were transcribed, researchers reviewed the videos and transcripts and discussed trends in all interviews, focusing specifically on times in which communication appeared to break down. Using the notions of *trouble* and *repair* (Schegloff, et al., 1997), instances were identified across interviews where communication appeared to break down. These instances were further examined to determine potential sources of trouble. Repair organization describes how parties in conversation deal with problems in speaking, hearing, or understanding. In reviewing interviews, instances in which trouble with *understanding* was deemed important. Trouble sources related to understanding may provide evidence that the ELs struggled with some aspect of a test item. Transcripts with trouble sources were analyzed to investigate the kinds of trouble sources across students, and to generalize about the challenges students face when interacting with different test tasks.

While there are many examples from the data that could illustrate students' interpretations of multisemiotic items, this analysis focuses on interviews with eight Grade 4 students from similar language backgrounds to compare how interpretive strategies differ across two item formats even when the construct is intended to be the same. This illustrates the importance of considering the affordances of a test item's semiotic representation and how it affects ELs' interpretations. Four students interacted with a traditional item intended to test the concept of buoyancy and four students interacted with an ONPAR multisemiotic task intended to test the concept of buoyancy. The ONPAR task was designed with the released traditional test item in mind. The table below shows the pseudonym of the students, the version of the item with which they interacted, and the language in which the student was interviewed.

[Insert table 1 here]

Traditional Item

The traditional item used for the study is a released Grade 4 science item from 2005 The National Assessment of Educational Progress (NAEP). It is an item with graphic support that asks students to compare the buoyancy of objects made from different materials (see images below). The NAEP description states that, "students must predict whether objects can float based on their perceived density levels and size and as a result, when objects are submerged, to determine how much liquid is displaced." The item is comprised of two parts; both parts of the item ask for similar kinds of information—comparing the buoyancy of two different balls—in similar formats—constructed response. The NAEP website indicates that a correct response would entail a statement that the water level goes up more in Cup 1 with correct explanation, a partially correct response would entail a statement that the water level goes up more in Cup 1, with explanation or an incomplete explanation, and an unsatisfactory or incorrect response would entail a statement that the water level goes up more in Cup 2, or that Ball 2 pushes the water level higher in Cup 2. Correct responses for this question would reflect that the different balls have different densities and therefore have a different effect on the water level.

According to the NAEP website, the first part of this item received an "easy" rating, meaning that the majority of students who took this item were able to correctly

respond to it. For the first part of the item, the NAEP website indicates that 62% of students answered the item correctly. However, the second part of the item received was more problematic for students; 60% of students answered this part of the item incorrectly. This may be due to the grammatical complexity of the item. Even though the questions in the part I and II of the item are about similar ideas, buoyancy and water displacement, the grammatical forms of the questions are different.

When transforming the NAEP item for the ONPAR study, the same exact wording of the item was used. In the first question of the item, the task demand is written in a syntactically simple form in present tense (*In which cup will the water level rise the most?*). Then, students are prompted to provide an explanation in the form of a constructed response (*tell why you think so*:).

[Insert Figure 2 here]

In the second question, the task demand is grammatically constructed as a conditional (*If she put this hollow ball in one of the cups, do you think the water level would rise more or less than it would if ball 2 were put in the cup?)* with a subjunctive verb (*put*). The subjunctive is used to indicate that something is an unreal or hypothetical situation. Furthermore, the second part of the question requires a comparison of the water level based on the two different balls being put into a cup– *would the water rise more or less than it would if ball 2 were put in the cup*. A third person singular pronoun (*it*) is used as a referring term in this sentence to reference the water level. While this is grammatically appropriate, the construction makes the question difficult to understand.

[Insert Figure 3 here]

ONPAR task

As noted above, the ONPAR task aims to test the same scientific concept of buoyancy as the traditional NAEP item. However, rather than presenting the steel balls and wooden ball as separate questions, the ONPAR task represents all three balls at the same time. The question asks about what will happen to the balls in the first part of the question, and the water level in the second part of the question. Other differences between the items include ONPAR's use of animation, colour, and pop up features. The pictures below depict what students see in the ONPAR task and are described more fully below (this task is also viewable on the ONPAR website).

The first screen shows three glasses of water. Students must click the "Go" button to begin the task. Once the "Go" button is clicked, an animation begins. Students see a wooden board appear over the three glasses of water and then a hand places three different balls on top of the board. Two of the balls look like metal (gray) and the third ball looks like wood (brown). As seen below, the first metal ball is bigger than the other two. One metal ball and the wooden ball are the same size.

[Insert Figures 4 and 5 here]

Once the balls are placed, the hand disappears, the board is removed, and directional arrows appear above the glasses to suggest that the balls fall into the water. Then, a question appears at the top of the screen, *what will happen to the balls?* The figure

below shows the first question of the task. Students are to manipulate the balls to show whether they will sink or float.

[Insert Figure 6 here]

On the first screen, there are several help features that students can use. Each of the three balls features a hyperlinked picture that depicts the interior of the ball; the steel balls are solid and the wooden ball is hollow. In addition, the question features hyperlinks. *What will happen* is hyperlinked to the arrows, indicating that this part of the question references the action of the directional arrows. The word *balls* is also hyperlinked to the steel and wooden balls.

[Insert figure 7 here]

Once students respond to the first screen, they click the check button and the second interactive screen appears. This screen is a slightly different colour to visually indicate that the question is different. A new question appears at the top of the screen, *what will happen to the water level*, and directional arrows appear next to the water level in the glasses. The placement of the balls from the first screen carries over to the second screen and students have to manipulate the dashed line on the cup to indicate whether and how high the water level rises for each ball. On this screen, the question also features hyperlinks. *What will happen*, again is linked to the directional arrows at the water level, and *water level* is linked to the water in the cup.

[Insert figure 8 here]

Analysis

In this section, examples from cognitive interviews are provided to show students' interpretations. First, the difficulty that ELs face with the linguistic demands of traditional paper and pencil items are illustrated by examining how four ELs interacted with the traditional science item; construct relevant vocabulary, paraphrasing the question, and producing answers were all sources of difficulty. In addition, through the use of repeated clarification strategies, the interviewer is able to gain insight into ELs' conceptual understanding of science constructs. Thus, the analysis shows the constraints that the primary modality of the traditional test items, language, poses. Further, it shows that alternative modalities such as static visual support may not be enough to compensate for the barriers posed by the language of the item. The result is that this item does not validly and reliably measure ELs' knowledge. In the second section, examples of how four ELs interacted with the ONPAR task are provided. The examples show that ONPAR's multisemiotic features afford ELs the opportunity to demonstrate content knowledge without language barriers. However, there are other trouble sources for ELs interacting with the ONPAR item. Design and use considerations are discussed in light of these findings.

Traditional item

Vocabulary trouble

All four of the ELs in this analysis struggled with vocabulary in the traditional item. Two of the most salient terms *identical* and *hollow*, were central to the meaning of the item. The term *identical* is visually supported; the graphic shows two glasses of water of the same size, filled to the same level. In a scientific context, it is important to control variables that may affect the outcome of an experiment and therefore it is important for students to understand that the variable of size is being controlled. The word identical is central to the meaning of this item. When asked to read the question aloud, students regularly faltered at the term *identical*, indicating a trouble source. Further, when asked to define or explain it, there was also evidence of trouble insofar as the students could not provide a definition or translation in Spanish.

The first example shows a trouble source in Jose's interview with the word *identical*. Jose was interviewed in English and, according to his educational records, his English proficiency was higher than other interviewees. While he spoke with little hesitation, when reading the question aloud, Jose used the word *indicated* rather than *identical*.

Jose: uh shown in the picture below- shown in the picture below Christina has two **indicated** cups that are filled to the same level with water. She also has two solid ball-ss steel balls. Christina puts ball one in a cup-- cup one and ball two in cup two. In which cup will the water level rise most?

Jose's trouble with *identical* suggests that he may not have recognized this word, or that he misinterpreted it. Both *identical* (the word in the question) and *indicate* (the word Jose reads) are cognates in Spanish and a bilingual student could confuse the words with each other. However, if the cognates have led him to misunderstand the semantic meaning of *identical*, it may make the question more difficult for him.

The other term that prompted trouble was *hollow*. This term was important for students to understand because it was an indicator of the density of the objects. Review of all four interviewees' transcripts indicated that this word was a trouble source and instances of repair were often initiated around it. One student, Maria, who was interviewed in Spanish, stated that she did not know what it meant—and that it was a confusing aspect of the item for her. In the example below, *hollow* is recast and by the interviewer.

Maria: Um, Christina has another ball that is the same size-as ball two, but this ball is made of wood and is .. hal- hol- hollow?
Interviewer: Hollow, yeah.
Interviewer: ¿Qué significa *hollow* - tú sabes qué significa *hollow*?
(What does *hollow* mean- you know what *hollow* means?)
Maria: Mm hmm.
Interviewer: ¿Eh, algo que fue confuso de esa pregunta, o entendiste?
(*Eh, [is there] anything that was confusing to you in this question, or did you understand?*)
Maria: Uh, the word hollow.
Interviewer: Que es la palabra hollow.
(What is the word hollow.)
Um, hollow significa eh … hueco, que hay un hueco adentro.
(*Um, hollow means ... hole, that there's a hole inside.)*

Vocabulary knowledge is considered to be one of the most important aspects of language development for ELs (Graves, August, & Mancilla-Martinez, 2012). Understanding the meaning of key vocabulary in test items is crucial because terms convey important semantic meaning. For example, if an EL understood the concept of *hollowness,* he or she may be able to answer the question correctly. Without understanding this key term, the meaning of test questions may not have been clear, even if this concept was clear to a student. Misunderstanding key vocabulary ultimately hinders ELs' opportunities to answer questions reflecting their content knowledge.

Trouble with the task demand

A second trouble source was explaining or paraphrasing the question. Most students had no difficulty identifying the task demands; they could read the relevant portion of the item to the interviewer. However, they had difficulty putting it in their own words or translating it. This aspect was most noticeable with the second part of the traditional buoyancy item when it was grammatically more complex (*If she put this hollow ball in one of the cups do you think the water level would rise more or less than it would if she put ball 2 in the cup*).

The following example shows how Isabel struggled with the task demand. Isabel correctly identifies the task demand, and then the interviewer prompts her to explain the question in her own words. After the interviewer asks, Isabel begins to try to explain and then trails off, followed by a prolonged silence. Silence is often an indicator of a trouble source, suggesting that the interviewee needs time to think about something. In this case, because the prolonged silence is followed by a re-start and another pause, it seems that Isabel needs time to formulate her ideas in words. Isabel finally responds that ball 2 would be more or less than the wood ball, which is correct, but when the interviewer prompts her to identify *what* is more or less, Isabel identifies *water*. From a scientific perspective, the question is asking about the *water level*, not merely water, and so Isabel's answer is not as precise as it could be.

Interviewer: OK, and now what's it asking you?
Isabel: If it's less...(3) If Ball 2...(3) would be--more or less than the wood ball.
Interviewer: More or less what?
Isabel: Water.

A second example of the task demand as a source of trouble for ELs comes from Pepe. Pepe shows a clear misunderstanding of the task demand, stating that both balls were put together into the same cup.

Interviewer: Okay, so what is the question?
Pepe: (silence) Do you think the water level will .. reez (rise) more or less than it will ball two were put in the cup.
Interviewer: What do you think that means?
Pepe: If they put .. the ball--the both balls together .. on the same (points to screen) ... on the cup. To see if it will have more water than the other one. Or the same, or less.

While Pepe is clearly able to identify the task demand when prompted by the interviewer, moving beyond identification to a correct interpretation is difficult for him.

These examples show that explaining a question is difficult, suggesting that there may be difficulty with comprehension of the task demand. If students are not able to comprehend the information a question requires, they will surely not be able to answer it.

Productive trouble

A final area of trouble was expression of answers to the second part of the item. Whether trying to write (type) or orally respond to the question, all students struggled to do so. Despite this challenge, all students were able to express some content understanding when the interviewer attempted to clarify what they knew. The interviewers' persistence in trying to understand ELs afforded them the opportunity to express their knowledge.

The first example is from Isabel's written response to the item. This example shows that Isabel's written answer would have been difficult to interpret because she uses a non-specific referent, *it*. It can refer to either the wooden ball or the steel ball, so her answer is unclear. The interviewer works with her to clarify her answer.

Interviewer: So you wrote...*Because it weighs more*. So, which one weighs more?
Isabel: The steel ball.
Interviewer: OK, but they're asking you if (points to a position onscreen) they put the--if you put-put the wood ball, it would cause more water to go up. So, you said *more*. You said *because-it weighs more*.
Isabel: Uh, I think it's less.
Interviewer: Oh, OK.
Isabel: [So I change--]
Interviewer: [So,] it has to do with the weight.
Isabel: Um-hmm.
Interviewer: It's fine. If you want to change it, you can change it.
Isabel: (begins changing her answer)
Interviewer: So, now you put, *Because the wood ball weighs less, it's going to go--up less.* That's what you said?
Isabel: Yeah.

Isabel identifies *weight* as a key factor to determining water displacement. This is a first start to understanding the relationship between density and buoyancy. However, her first written answer lacks specificity needed to understand which of the two objects weigh more and what effect that would have on the ball. The interviewer prompts her to explain her answer and then gives her the opportunity to clarify and elaborate her thinking. Without the ability to specify which material she thinks weighs less, it would be impossible to judge her understanding of this idea.

A second example of productive difficulty as a trouble source comes from Pepe. Pepe's written answer is incoherent at first. However, with further questioning, Pepe is able to demonstrate understanding of the idea of water displacement.

Interviewer: Okay, so, (reading Pepe's answer) *because- because the ball is made of wood and the other is made of steel. So if I put the wood ball it will have the same amount of water will rise*. Or more? ... So your answer is telling me that it's because of the material that it's made out of? But what about the material?
Pepe: The wood does not it will stay- it doesn't not have pressure.

Interviewer: Okay.
Pepe: And if you put the wood ball in the water, it will stay floating.
Interviewer: Oh, okay, the wood will stay floating. How about the steel ball?
Pepe: It will go down and the water will go up.

Pepe's written answer states that the wood ball would cause the water to rise the same amount. The interviewer prompts Pepe to clarify his answer and Pepe says that the wood doesn't not have pressure. The answer is unclear and the interviewer seeks clarification again. Finally, Pepe is able to state that the wood ball will stay floating and the steel ball will go down (sink) and the water will rise. While it is not clear that Pepe has a full understanding of the phenomena, he is able to express more clearly what he understands orally than in writing.

These examples show how difficult it is for ELs to express their understanding of the phenomena in writing. When interviewers are able to prompt them to explain and clarify what they mean, the students are able to provide greater evidence of what they do/ do not know. Thus, constructed response items which rely on evidence of student understanding through writing may be particularly difficult for making judgements about ELs' conceptual understanding.

This section has shown three areas of trouble for ELs with traditional test items: vocabulary, comprehension of the task demand, and difficulty with producing comprehensible written responses. All three of these areas deal with linguistic aspects of traditional items that need to be considered for creating valid and reliable tests for ELs. In the section that follows, ELs' responses to the ONPAR test task are examined.

ONPAR task

The analysis of the student responses to the introductory and first response screen, focus on several aspects: (1) whether students are able to describe the pictures, (2) their interpretation of the animation, and (3) how students understand the three balls of differing materials and sizes. In the analysis of responses to the second screen, students' interpretation of the water level is examined. Comprehension of these visual aspects would demonstrate whether ELs understand the visual depiction of the construct in such a way as to gain insight into their understanding. The primary trouble sources identified are related to students' ability to interpret the visual depiction, suggesting that the multisemiotic tasks, too, present difficulties for student comprehension. However, these trouble sources are different from the linguistic trouble sources present in the traditional items. This suggests that the modalities presented by the test offer different meaning making potentialities and therefore differential opportunities for students. In this section, two examples of misinterpretation are presented. Conversely, two cases in which meaning is accurately interpreted are also presented.

Trouble with the grammar of visual design

The primary area of misunderstanding is due to the multiple ways in which visuals can be interpreted. While students may be able to correctly interpret an object within an image, correct interpretation of visuals relies on students understanding the visual as it is used in context, or the grammar of visual design (Kress & van Leeuwen, 2006). The first example in which the grammar of visual design affects meaning

making comes from students' interpretation of the size of the balls. While the balls in the ONPAR task are depicted in two sizes, some students interpret the size to be small, medium, and large. The example below from Ines illustrates this. In this example, after seeing the animation, the interviewer asks Ines what she saw. Ines describes that she sees three balls of three different sizes.

Ines: (clicks on the 'Go' button animation plays on the screen)
Interviewer: OK, can you explain what happened?
Ines: Um, they put a piece of wood, and then they put (gestures with three fingers) three balls.
Interviewer: OK, and then what happened?
Ines: The put a b--a (points toward the screen) big one, (points toward the screen) a medium one, and (points toward the screen) a small one.
Interviewer: OK, can you point out which one's the big one?
Ines: Uh, (points at the ball furthest to the left of the three) this one.
Interviewer: The medium?
Ines: (points at the middle ball) This one.
Interviewer: And the last one, (waves her hand) the small one?
Ines: (points at the ball furthest to the right of the three) This one.

In this case, Ines is mistaken in her interpretation of the size of the different balls. It may have been the fact that three balls are used, or the placement of the balls that interfere with her interpretation of the sizes. The misinterpretation of the size of the balls may not have led her to a wrong answer (i.e., she may have determined that the middle ball still sank), however, the example shows that the meaning of visual imagery is not necessarily transparent and that similar kinds of meaning making that takes place with linguistic signs, takes place around visuals. Thus, it is the grammar of the visual that led her to make this mistake.

A second example where students misinterpret visuals because of the grammar of visual design occurred around a support feature that was intended to convey that two of the balls are solid and one is hollow. As shown above, when students move their mouse on one of the balls, a pop up picture appears indicating the interior of the ball. Interviewers asked students about what they saw when this occurred. Two of the students stated that the balls broke in half. In the first example, Luisa states that she thinks this is the case, but does not know why this would occur. Thus, there is no real-world experience that she is drawing upon to conjecture this, but rather it is the visual itself that leads her to this conclusion.

Interviewer: Now those pictures that come up with your mouse what does that tell you?
Luisa: Hey (sounds like a noise—almost like the student is saying I don't know).
Interviewer: It's telling you hey?
Luisa: I don't know (not clearly enunciated).
Interviewer: That's what it's saying?
Luisa: I say huh? (student seems irritated).
Interviewer: No those pictures. What do you think that it's saying?
Luisa: That's it's going to break apart in the water.
Interviewer: Hmm? (leans in to hear better)
Luisa: That it's going to break apart (iconic gesture). In the water. (enunciating more clearly this time).

Interviewer: Oh really, why is it going to break?
Luisa: Ahhhhh… I don't know.

These examples show that the meaning of visual signs is not necessarily transparent to students. In the first case, Ines misunderstands the size of the balls based on a schema of size: small, medium, and large. The fact that there are three different balls triggers that they are of three different sizes. In fact, Kress & van Leeuwen (2006) suggest that diptych and triptych images convey a sense of comparison, so the student correctly interprets the comparative visual grammar here. She understands that she is to compare the properties of the ball against each other, but she does not grasp that there is more than one variable to compare. The saliency of size in the visual depiction may have been the most noticeable aspect, leading her to misinterpret the image.

In the second example, the student does not understand the meaning of the pop-up feature, and how it is intended to assist her in seeing the interior of the balls. She understands the pop up to show action rather than to specify a property of the ball. While the word *hollow* was confusing to students in the traditional item, the visual depiction of it is also confusing.

ONPAR successes

While the first screen presents some challenges for students' interpretation of visual meaning, the second screen of the item is quite clear to all students. This is in contrast to the second part of the traditional question that posed so many difficulties for students because of the constraint that the language of the item posed. In the case of the ONPAR task, the dashed line represents "water level." All four students correctly interpret the visual element. Furthermore, the response mechanism (dragging a dashed line up or down) allows ELs to express that the water level would rise when an object sinks. The following example shows how one student, Beatriz, responds to the second part of the ONPAR task. According to her educational records, Beatriz is a low proficiency EL. In the example, the interviewer and Beatriz look at the ONPAR screen and discuss the visual. Beatriz is able to identify the question and, even though she struggles with paraphrasing the question, she is able to do so and she is able to relate it to the graphic depiction.

Interviewer: Okay, y sigue a la próxima. ¿Okay, y qué cambiaba?
(*Okay, and continue to the next. Okay, and what changed?*)
Beatriz: Se enseña unas líneas.
(*Some lines appeared*.)
Interviewer: Okay, unas líneas. ¿Y qué crees que te enseñan las líneas?
(*Okay, some lines. And what do you think the lines show you?*)
Beatriz: Que suben las bolas.
(*That the balls are going up.)*
Interviewer: ¿Okay, y dónde está la pregunta? (student points to question on the screen) ¿Me la puedes leer?
(*Okay and where is the question?* (student points to question on the screen) *Can you read it to me?*)
Beatriz: "What will happen to the water level?"
Interviewer: ¿Y qué significa eso?
(*And what does that mean?*)
Beatriz: ¿Qué va pasar … uh … con el agua … "level"? Es como nivel.

(*What will happen … uh … with the water … level? It's like* level.)

Interviewer: Mm hmm. ¿El nivel del agua, verdad? Entonces cómo responderías tú a esta pregunta?

(*Mm hmm. The level of the water, right? So, how would you respond to this question?*)

Beatriz: Mmm … Subiendo el agua y bajándola.

(*Mmm … Raising and lowering the wáter*.)

This example shows the complementary role that language and visuals play with one another in the ONPAR environment. First, Beatriz does not present a concise representation of what the dashed lines mean (the balls are going up). Rather than stating that the lines indicate the water level, as other students do, Beatriz relates the lines to the balls. In addition, even though she clearly identifies what the question is, she struggles to explain what it means. The word *level* in particular presents some difficulty, though she is able to translate it into Spanish after thinking a moment. Beatriz is able to determine what the task demand is, to raise or lower the water on the screen. The language of the item and the visual seem to play complementary roles and support this student's understanding of the task demand, such that she responds to it in a meaningful way.

While the second part of the NAEP item proved to be difficult for all students, the second part of the ONPAR task proved the opposite. The different affordances of the modalities and the way they related to each other played a major role in allowing ELs access.

Despite some difficulties with comprehension of the visual depictions, there were success stories of student interaction with ONPAR's multisemiotic tasks. A complete transcript in the appendix shows a complete interview with one EL, Cecelia, about the ONPAR buoyancy task. The transcript is provided in its entirety to show the speed and ease of interaction with the entire task. In this case, Cecelia is able to move through the entire task quickly and easily, something that even the highest proficiency student is not able to do with the traditional item. From the interview, it is clear that Cecelia is able to display her content knowledge with the ONPAR task. She is able to successfully interact with the ONPAR interface to complete the task, successfully showing through the multisemiotic format what she has conveyed verbally.

Discussion

Given the importance of standardized assessments in the United States and the rapidly changing testing landscape, careful consideration of the affordances of testing materials and ways in which students understand test items is warranted. Trouble sources with traditional items were consistent across ELs' interviews and suggested that ELs struggled with both the receptive and productive language load posed by the dominant modality of traditional test items, language. This is not surprising given what researchers have shown in quantitative studies of test outcomes and the discussion about accommodations. However, there is power in examining ELs' interactions with test items to focus on the affordances of the semiotic resources used and how they introduce non-construct relevant barriers. From this analysis, it is clear that what has been deemed as an achievement gap is exacerbated by the language of content tests; language and content are indeed confounded. Until content tests do no conflate language and content, they will remain invalid for ELs. It is imperative to consider this in light of the millions of dollars being spent in the United States to

develop innovative tests to measure college and career readiness standards. With the use of computers come greater affordances to support ELs' understanding and unless these resources are utilized skilfully, this money will not be invested in ways that it should be.

Overall, the ONPAR tasks show promise in assessing ELs' content knowledge in an innovative way. While there are still trouble sources in interviews about the ONPAR task, students understood many aspects of the task that had been confusing in the traditional test environment. Many students were able to understand the intended meaning of visual depictions, and the language load of the multisemiotic task did not present the same barriers as traditional items. The multisemiotic features functioned in complementary ways to support ELs' understanding of the items and afforded ELs new opportunities for understanding and expressing their knowledge. However, the multisemiotic task was not without issues and can be further improved, taking into account ELs' understanding. Most notably, the affordances of visuals in relationship to language need to be carefully considered.

The following recommendations are suggested to improve upon the design and use of multisemiotic test tasks:

1. Design items with an understanding of visual grammar as a foundation (e.g., Kress & van Leeuwen, 2006). As seen in the interviews about the first ONPAR screen, visual meaning is made in systematic ways, just as language. The visual design of items must be grounded in an understanding of visual grammar so that the grammar can be relied upon as a meaning making resource.
2. Use language in targeted and complementary ways with visuals. The meaning of visuals alone is not always transparent to students. When students do not understand a visual, they appear to rely on language to clarify its meaning. As shown in this analysis, the second screen of the ONPAR task was quite successful because the visuals were linked to language. Further, previous research shows that the indeterminacy of visuals can be clarified and specific through targeted use of language (Amteller & Pinto, 2002). Thus, a limited amount of language is supportive to students in this environment.
3. Conduct fine-grained qualitative interviews with students during the research and design phase to examine how students interpret the meaning of items vis a vis the semiotic resources used. Employing cognitive interviews has provided insight into students' interactions with traditional and multisemiotic test items and has shown the value of gaining first-hand information from students when developing test items. Given the role that testing plays in annual accountability, test development has become an increasingly large industry within the United States. Yet, still relatively little is known about how students interact and understand test items vis a vis the modalities employed. Student interaction with different items has shown a variety of sources of trouble with both traditional and multisemiotic items. These insights have the potential to lead to improved item and test design for ELs, in particular.
4. Integrate training in visual literacy in classroom instruction. As visuals become a more dominant form of communication in testing situations, students need to become accustomed to reading images. This training needs to become part of instruction so that students have access to this a meaning making resource. This is especially true if test items become more reliant upon visuals for meaning making.

Conclusion

In conclusion, listening to students' voices plays an important role in making design decisions for assessments and determining appropriate ways to test ELs. Students' voices from this study indicate that a one-size-fits-all approach is not necessarily an appropriate approach to accommodating ELs. The multisemiotic approach employed by ONPAR allows students multiple, complementary paths that support their understanding and shows the promise of designing tests that allow students a variety of approaches to meaning making. The multisemiotic approach affords ELs new opportunities for making meaning, which, in turn, can provide new opportunities for them to show their content understanding. As the United States moves to national computerized testing for content tests, these insights are particularly relevant.

References

Abedi, J., Courtney, M., & Leon, S. (2003). *Effectiveness and validity of accommodations for English language learners in large-scale assessments* (CSE Technical Report Number 608). Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing.

Abedi, J., Courtney, M., Mirocha, J., Leon, S., & Goldberg, J. (2005). *Language accommodations for English language learners in large-scale assessments: Bilingual dictionaries and linguistic modification* (CSE Tech. Rep. No. 666). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.

Abedi, J. & Gándara, P. (2006). Performance of English language learners as a subgroup in large-scale assessment: Interaction of research and policy. *Educational Measurement: Issues and Practices, 26*(5), 36-46.

Abedi, J., Hofstetter, C., & Lord, C. (2004). Assessment accommodations for English language learners: Implications for policy based research. *Review of Educational Research, 74*(1), 1-28.

Abedi, J., Leon, S., & Mirocha, J. (2003). *Impact of students' language background on content-based assessment: Analyses of extant data* (CSE Tech. Rep. No. 603). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.

Abedi, J., & Lord, C. (2001). The language factor in mathematics tests. *Applied Measurement in Education*, *14*(3), 219– 234.

Almond, P. J., Cameto, R., Johnstone, C. J., Laitusis, C., Lazarus, S., Nagle, K., Parker, C. E., Roach, A. T., & Sato, E. (2009). *White paper: Cognitive interview methods in reading test design and development for alternate assessments based on modified academic achievement standards (AA-MAS).* Dover, NH: Measured Progress and Menlo Park, CA: SRI International.

Ametller, J., & Pinto, R. (2002). Students' reading of innovative images of energy at secondary school level. *International Journal of Science Education*, *24*(3), 285-312.

Bowles, M. A. (2010). *The think aloud controversy in second language research*. New York: Routledge.

Bowles, M., & Stansfield, C. (2008). A practical guide to standards-based assessment in the native language. Retrieved 1/23/09, from the World Wide Web: http://www.ncela.gwu.edu/spotlight/LEP/2008/bowles_stansfield.pdf

Bransford, J.D. (1984). Schema activation versus schema acquisition. In R. C. Anderson, J. Osborn & R. Tierney (Eds*.), Learning to read in American schools: Basal readers and content texts* (pp. 259-272). Hillsdale, NJ: Lawrence Erlbaum Associates. –Reprinted in H. Singer & R. Ruddell (Eds.),*Theoretical models and processes of reading*. Newark, DE: International Reading Association.

Carr, T., Kopriva, R., & Rex, J. (2007). *Handbook on developing accessible assessment items for English language learners and students with learning disabilities*. Columbia, SC: South Carolina Department of Education.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data* (Rev. ed.). Cambridge, MA: The MIT Press.

Ercikan, K., Arim, R., Law, D., Domene, J., Gagnon, F., & Lacroix, S. (2010). Application of think aloud protocols for examining and confirming sources of differential item functioning identified by expert reviews. *Educational Measurement: Issues and Practice, 29*(2), 24-35.

Fortuny, L.A., Garolera, M., Hermosillo, R., Feldman, E., Fernandez Barillas, H., & Keefe, R. (2005). Research with Spanish-speaking populations in the United States: Lost in translation a commentary and a plea. *Journal of Clinical and Experimental Neuropsychology 27*, 555-564.

Francis, D., Rivera, M., Lesaux, N., Kieffer, M., & Rivera, H. (2006). *Practical guidelines for the education of English language learners: Research-based recommendations for instruction and academic interventions*. Portsmouth, NH: RMC Research Corporation, Center on Instruction. Available:www.centeroninstruction.org/files/ELL1-Interventions.pdf;

Garcia, O. (2009). *Bilingual education in the 21st century: A global perspective.* New York: Blackwell/Wiley.

Gibson, J. J. (1979/1986). *The ecological approach to perception.* Hillsdale, NJ: Erlbaum.

Graves, M., August, D. & Mancilla-Martinez, J. (2012). *Teaching vocabulary to English language learners.* New York: Teachers College Press.

Hudicourt-Barnes, J., Noble, T., Roseberry, A., Suarez, C., & Warren, B.. (2008, March). *How Do Children Make Sense of High-Stakes Science Tests? The Case of MCAS Participants*. Paper presented at the annual American Education Research Association. New York, NY.

Kieffer, M.J., Lesaux, N.K., Rivera, M., & Francis, D.J. (2009). Accommodations for English language learners taking large-scale assessments: A meta-analysis on effectiveness and validity. *Review of Educational Research, 79*(3), 1168-1201.

Kopriva, R. (2008). *Improving testing for English language learners*. New York: Routledge.

Kopriva, R. (2009). Assessing the skills and abilities in math and science of ELLs with low English proficiency: A promising new method. *AccELLerate!, 2*(1), 7-10.

Kress, G. (2010). *Multimodality: A social semiotic approach to contemporary communication.* London: Routledge.

Kress, G., Jewitt, C., Ogborn, J., & Tsatsarelis, C. (2001). *Multimodal teaching and learning: The rhetorics of the science classroom*. London: Continuum.

Kress, G. & van Leeuwen, T. (2001). *Multimodal discourse: The modes and media of contemporary communication*. New York: Oxford University Press.

Kress, G. & van Leeuwen, T. (2006). *Reading Images: The grammar of visual design*. London: Routledge.

Lee, O., & Fradd, S.H. (1998). Science for all, including students from non-English language backgrounds. *Educational Researcher 27*(3), 12–21.

Lemke, J. (1990). *Talking science: Language, learning, and values*. Norwood, NJ: Ablex.

Lemke, J. (1998). Multiplying meaning: Visual and verbal semiotics in scientific text. In J. R. Martin & R. Veel (Eds.), *Reading science: Critical and functional perspectives on discourses of science*. London: Routledge.

author

Martiniello, M. (2008). Language and the performance of English language learners in math word problems. *Harvard Educational Review, 78*(2), 333-368.

National Center for Education Statistics. (2008). NAEP data explorer. Washington, DC. Retrieved July 22, 2008, from htttp://nces.ed.gov/nationsreportcard/nde/viewresults.asp: U.S.Department of Education Institute of Education Sciences.

National Science Foundation. (1994). *Women, minorities, and persons with disabilities in science and engineering:* 1994 (NSF-94-333). Arlington, VA: Author.

NAEP item: http://nces.ed.gov/nationsreportcard/itmrlsx/detail.aspx?subject=science

ONPAR task: http://onpar.us/assessmentTasks/scienceTestlet.html

Rivera, C. & Collum, E. (Eds.). (2006). *State assessment policy and practice for English language learners: A national perspective.* Mahwah, NJ: Lawrence Erlbaum and Associates.

Ruiz de Velasco, J. & Fix, M. (2002). Limited English proficient students and high stakes accountability systems. In D. Piche, W. Taylor, & R. Reed (Eds.), *Rights at risk: Equality in an age of terrorism.* (pp. 245-261). Washington, DC: Citizens' Commission on Civil Rights.

Schegloff, E. A., Jefferson, G., & Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. *Language*, 361-382.

Solano-Flores, G. (2010). *Vignette Illustrations as a Form of Testing Accommodation for English Language Learners: A Design Methodology for Use in Large-Scale Science Assessment.* Paper presented at the Annual Conference of the National Council of Measurement in Education, Denver, Colorado, April 29-May 3, 2010.

Willis, G.B. (2005). *Cognitive interviewing: A tool for questionnaire design.* Thousand Oaks, CA: Sage.

Wolf, M. K., Herman, J. L., & Dietel, R. (2010, Spring). *Improving the validity of English language learner assessment systems* (CRESST Policy Brief No. 10 - Full Report). Los Angeles, CA: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Tables

Student Pseudonym	Interview Language	Item
Isabel	English	Traditional
Maria	Spanish	Traditional
Pepe	Spanish/ English[1]	Traditional
Jose	English	Traditional
Cecelia	English	ONPAR
Luisa	English	ONPAR
Ines	English	ONPAR
Beatriz	Spanish	ONPAR

Table 1: Student participants

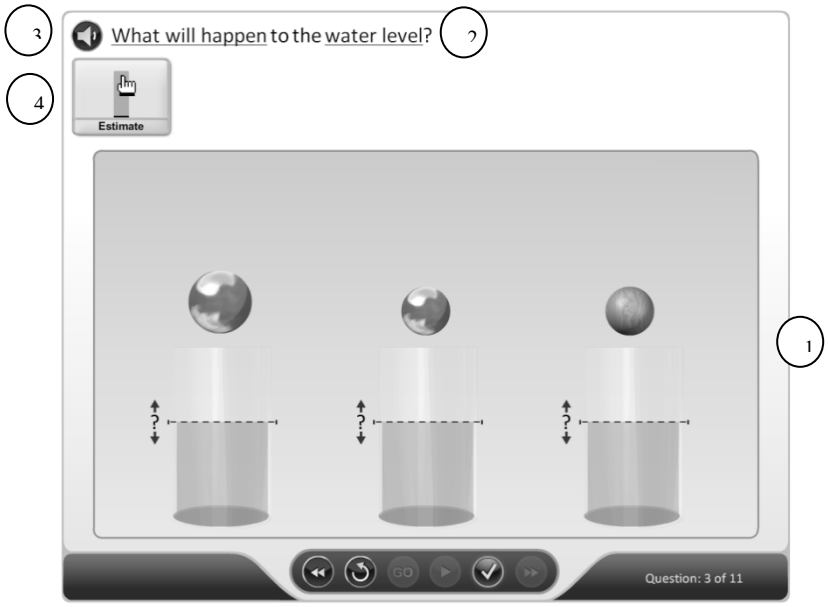

Figure 1. ONPAR layout and features

1. Graphic: conveys information of task
2. Text prompt: conveys task demand
3. Speaker button: provides text prompt translated into Spanish (optional)
4. Icon: demonstrates physical action needed to complete task (optional)

[1] This student began the interview in Spanish and then switched to English during the interview process.

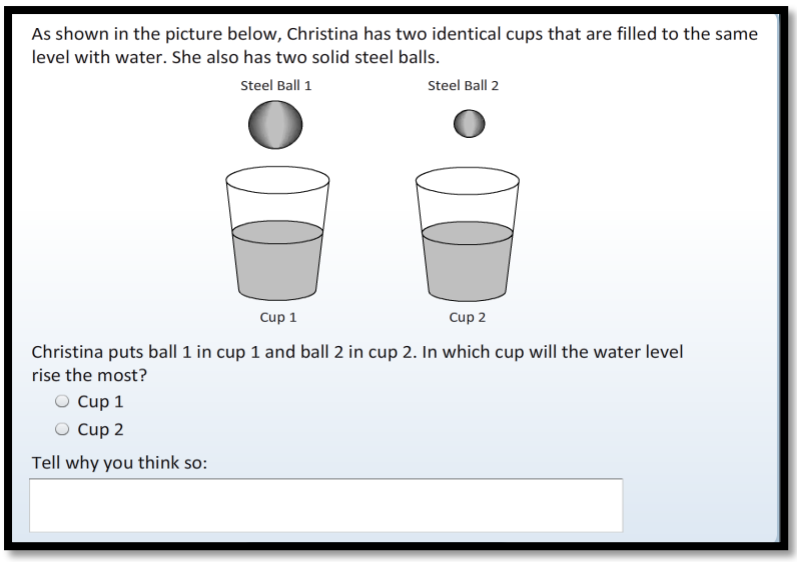

Figure 2: Traditional Item Screen 1 as represented in ONPAR study

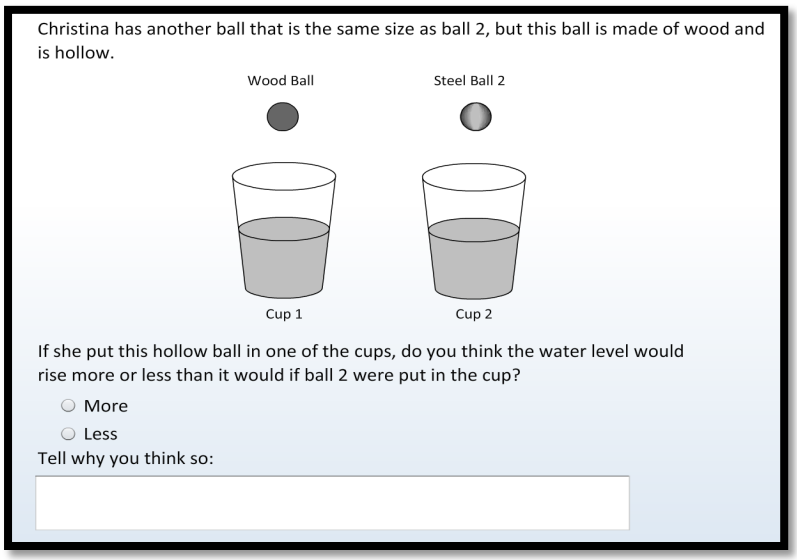

Figure 3: Traditional Item Screen 2 as represented in the ONPAR study

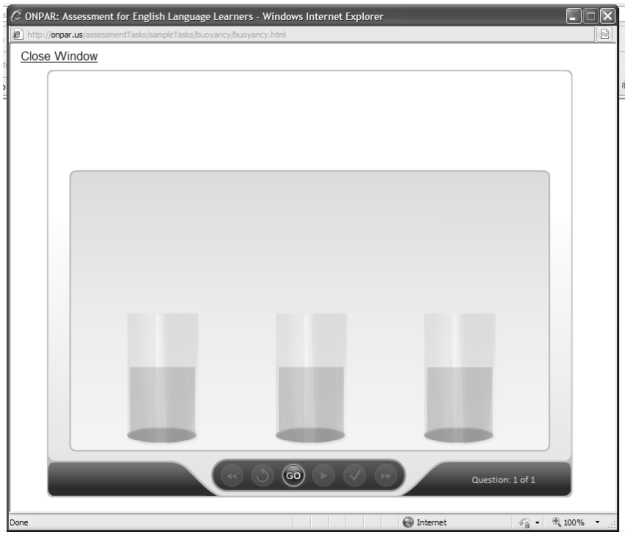

Figure 4: ONPAR orientation screen

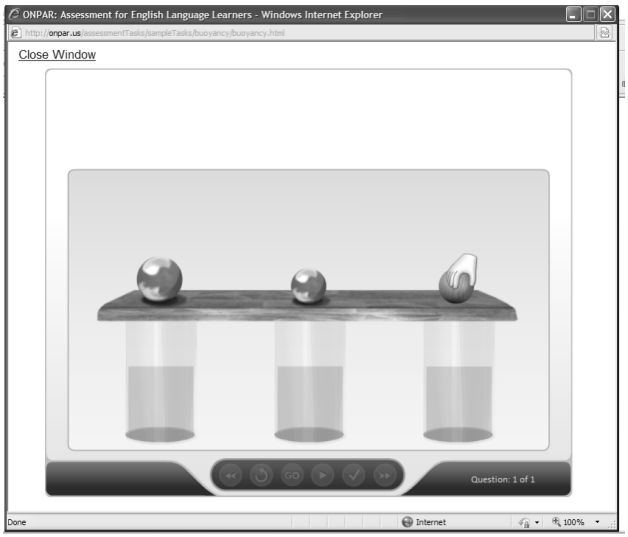

Figure 5: ONPAR animation

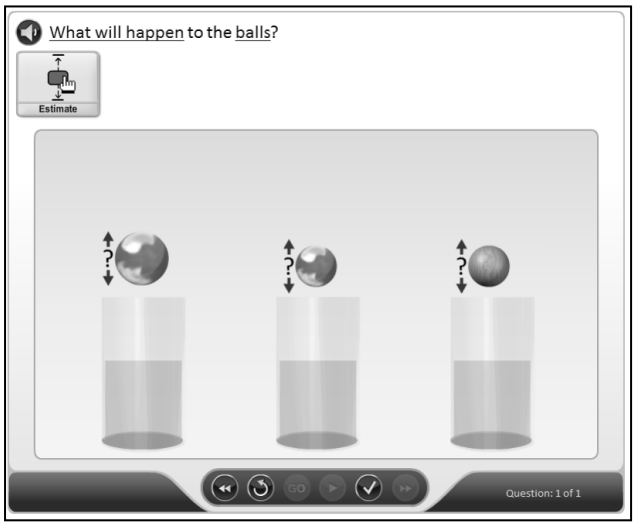

Figure 6: ONPAR question 1

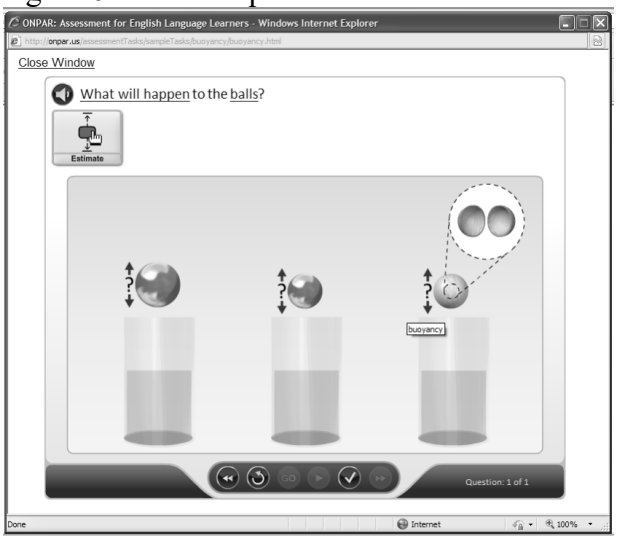

Figure 7: ONPAR visual support for “hollow”

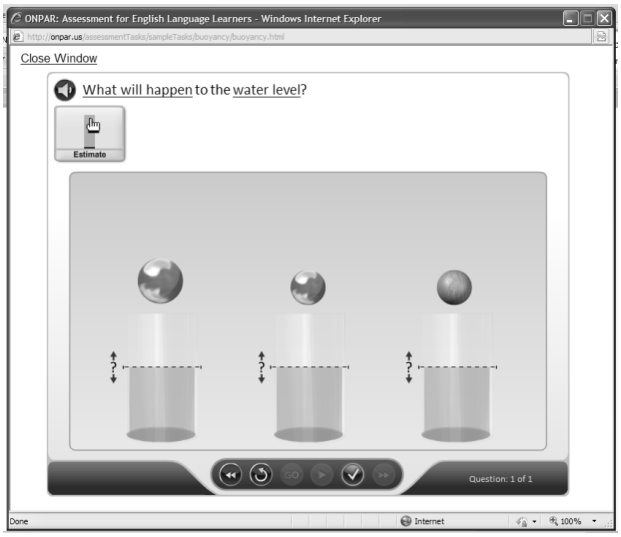

Figure 8: ONPAR question 2

Appendix: Cecelia transcript

Interviewer: OK, we can hit 'Submit' and go on to the next question. Go ahead and click on Number 3.
Cecelia: (clicks on Number 3)
Interviewer: What do you see in the screen?
Cecelia: Um...water in the--the--shade.
Interviewer: OK, you see water? What do you think these are (points to the three objects on the computer screen)?
Cecelia: Glass of wa--um, glasses.
Interviewer: Very good. You see three glasses of water. Go ahead and hit the 'Go' button.
Cecelia: (clicks on the 'Go' button)
Interviewer: Nice. Um, so, what it is you change on the--screen?
Cecelia: Um...they put..um..little balls on top of them, the glasses.
Interviewer: Um-hmm then, what's the question? Can you read it to me?
Cecelia: What would happen to the ball? (reading).
Interviewer: What do it's asking, in your own words?
Cecelia: Um if, maybe they will float.
Interviewer: Right, it's just--OK. So, can you try to answer this question?
Cecelia: (clicks on the balls in the cup icons, drags the mouse to adjust their positions, and releases the mouse button clicks the ball on the water surface in the remaining cup, adjusts the level, and releases the mouse button)
Interviewer: Very nice. So, how did you come up with this answer?
Cecelia: Um..the gray balls are too heavy, and the wood ball seems--is lighted.
Interviewer: so, because it's light, it's going to float?
Cecelia: Yes.
Interviewer: Awesome. Alright go..uh..ahead and click the 'Submit' button. We'll move ahead to the next screen.
Cecelia: (clicks the 'Submit' button).
Interviewer: Now what do you see different?
Cecelia: The water.
Interviewer: OK, and what's the question?
Cecelia: What will happen to the water level? (reading)
Interviewer: And what's it asking you, in your own words?
Cecelia: Um...the water will go down..or up.
Interviewer: Very good. You want to try to answer the question?
Cecelia: Um (takes the mouse and adjusts the water levels continues making adjustments to the water levels in the cups)...(7) Like that. Stay the same.
Interviewer: That one will stay the same line?
Cecelia: The water's not that heavy.
Interviewer: Right, and the other ones, one went up high and one went up...higher. I mean, (points at the screen) it went up, but it didn't go up as high. And w--why did you answer that?
Cecelia: Um, this one is little, so it goes a little up.
Interviewer: Um-hum.
Cecelia: And this one's bigger, and it goes a lot higher.
Interviewer: Very good. Alright, uh, anything on the screen that's kind of confusing, that you don't understand?
Cecelia: No.
Interviewer: OK, you want to hit the 'Submit' button?

Cecelia: (clicks the submit button)
Interviewer: Let's go to Number 5.

Transcription Conventions

[ ]	overlap of speech between two speakers
-	speaker has truncated a word
?	Rising intonation for question
!	Exclamation
.	Falling intonation indicating end of an utterance
..	indicates a brief break in speech (barely perceptible)
…	pause of medium length:
(2)	Long pauses are in numeric form
X	speech that is not audible
<X water X>	words enclosed are not clearly audible, but an estimate is provided
( )	indicate a transcriber's observation of relevant contextual information
@	laughter; one @ is used for each pulse of laughter
DID	word bears a primary accent
:	indicates the lengthening of a word or letter