**Best Practices in Developing a Universally Designed Assessment on Scientific Reasoning**

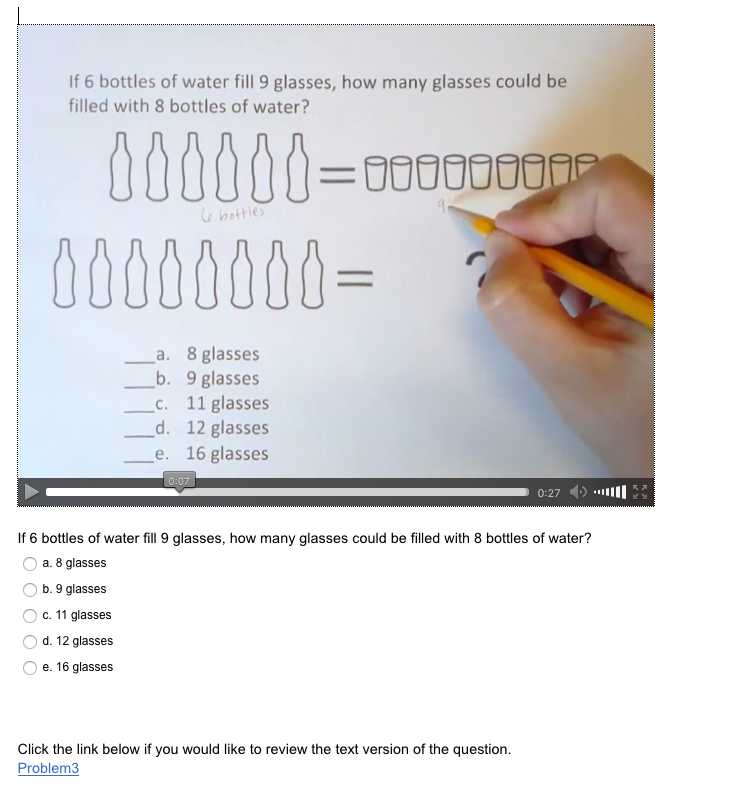
(authors deleted for blind review)

**Abstract**: The purpose of this session is to showcase an example of a universally designed assessment on scientific reasoning. This technology-based assessment incorporates universal design for assessment principles so that it meets a wide range of needs of learners who will take the assessment. The problem with many assessments is that they impose barriers for many students who may struggle with learning issues unrelated to the content of the test (e.g., issues with decoding text, limited vocabularies, shortened attention spans, etc.). These traditional tests end up not accurately measuring the construct that they intend to measure because it is testing a host of skills irrelevant to the content of the test itself. Our goal is to share best practices in developing a universally designed assessment as well as our lessons learned in the process and future plans for testing.

“If assessments that purport to *measure* student learning are not ... universally designed, those assessments can pose barriers or obstacles for students with learning disabilities, obstacles that interfere with their ability to demonstrate what they have learned” (Christensen, Shyyan, & Johnstone, 2014, p. 23). To address this issue, we set out to convert a traditional paper-based test on scientific reasoning developed by Lawson (1978) to a computer-based test that incorporated universal design for assessment features. During initial rounds of testing with the paper-based test, we found that many students struggled with demonstrating their knowledge because of learning issues unrelated to the test content. To figure out the specific issues that students were facing, we observed two students with disabilities take the assessment with the help of a special education professor who provided several supports. These supports included reading the question aloud, pointing to the words that were read, gesturing to items in the problem, and labeling information on the problem. Although a computer-based test could provide certain supports, such as read aloud and vocabulary definitions, we felt that this was not sufficient for supporting students. Thus, we set out to create video-based items that would provide the same types of supports that were provided from the specialist.

To guide our design of these new test items (See Figure 1), we incorporated many of the universal design for assessment principles proposed by Thompson, Johnstone, and Thurlow (2002):

1. Measuring What the Test Intends to Measure: The design of our test eliminated irrelevant constructs related to reading and other attention issues because the problem is read aloud to students and important information is gestured to.
2. Respecting the Diversity of the Population: The language of the test was reviewed by experts in special education, educational technology, and science to ensure that the test did not disadvantage students based individual characteristics, such as age, ethnicity, gender, disability, etc.
3. Providing Clear Text: The contrast, fonts, and blank spaces in the videos and the user interface of the test were reviewed by experts to ensure maximum readability.
4. Offering Clear Visuals: Diagrams as well as hand gesturing within the video provided clear visuals for students. In the process of developing the videos, we observed that some of the diagrams needed additional clarity. For example, arrows were added to an image of a ruler measuring liquid in a beaker to indicate the precise measurement for students to use in the problem.
5. Including Concise and Readable Text: The text was read aloud to students by a narrator. In developing the narrations, we came across text that was awkwardly phrased and the language was revised accordingly.
6. Allowing for Format Change: The test itself embedded a variety of test formats: students have the option to watch the videos or read the text to themselves. They can also pause, rewind, and replay the video as many times as needed.



**Figure 1:** Test item in a Universal Design for Assessment

In summary, in the process of developing the videos for the test, we observed issues with the actual test items themselves which we iteratively corrected. We will share our lessons learned from this iterative process of test development along with tips and tricks for how to create these types of assessments for your own classrooms. Finally, we will share our future plans for testing this assessment with the target population.

**References:**

Christensen, L. L., Shyyan, V., & Johnstone, C. (2014). Universal design considerations for technology-based, large-scale, next-generation assessments. *Perspectives on Language and Literacy*, 40(1), 23.

Johnstone, C., Thurlow, M., Moore, M., & Altman, J. (2006). *Using systematic item selection methods to improve universal design of assessments. Policy directions*. Number 18. National Center on Educational Outcomes, University of Minnesota.

Lawson, A. E. (1978). Development and validation of the classroom test of formal reasoning. *Journal of Research in Science Teaching, 15*(1), 11-24. Retrieved from http://www.public.asu.edu/~anton1/LawsonAssessments.htm