

Assessment Analysis and Reflection

Inferences

The assessment was changed between the time I handed it in and the time I taught the matter unit. My cooperating teacher preferred me to use the standards that the first grade team had set already. I was free to incorporate activities and lessons that I developed but the points that I had to hit were clearly written out. My cooperating teacher also advised me to make the assessment simpler due to time restrictions. As a guest in the classroom I complied. Due to this, I had to change the assessment by deleting certain items that we did not cover during the unit, edit some questions to make them fit what we specifically covered in the unit (ex: changing milk to food coloring in question #2), and reduce the number of multiple choice options from four to three. The Table of Specifications for this assessment was modified to reflect the changes.

As a whole, the class did very well and performed at the level I have come to expect from them from my observations and experience checking their work. My cooperating teacher and I expect all of the students to do well on assessments because the content is approached in so many ways, with so much support during the lesson and outside of the lesson if need be. The bell curve does not seem to apply to this first grade class. We always have the outliers who, for some reason, have not performed to the best of their abilities, but always have a high performance result.

42% of the students scored perfectly, 74% missed one or fewer points, 89% of students missed two or fewer points. One student missed 4 points and one student missed 6 points. The mean score was 98%. The majority of the class was present for the entire unit, which is why this average does not surprise me. The student who missed 4 points missed the first two days of the unit. Although I made sure to verbally explain the concepts he missed on the days he was absent, he missed the hands-on experiences we did on those days and the foundational content that would have remained in his brain more effectively if he had been able to experience those activities.

The high performance by the students could be interpreted as a too-easy assessment, but since I know these students and how the unit was taught I would disagree. The students all scored so high because the Intended Learning Outcomes were completely analyzed, dissected, and fit into the unit very purposefully. The only question that gave the students a lot of trouble (#6), with six out of the nineteen students incorrectly responding, was classified a higher level, “application” question. I decided that the students did not fully understand because we had not actually done the experiment discussed in the question, which is how the rest of the questions were formed. Throughout the unit I had had the students predict what would happen to numerous types of matter. Many of the students made many wrong predictions. This could have led them to believe I was trying to trick them again with a prediction about what would happen to the fake sugar in question #3. I decided to throw this question out due to its confusing, unreliable nature.

Two Individual Students

The student who missed 4 points missed the first two days of the unit. Although I made sure to verbally explain the concepts he missed on the days he was absent when he returned, he missed out the hands-on experiences we did on those days and thus missed the foundation-building activities that would have benefitted him during the entire unit. He, in a way, missed out on the entire unit because of this. He missed questions:

- “3. Think back to our experiment when we mixed sugar with water. Based on this, what can we predict will happen when we mix fake sugar with water? The fake sugar will _____.
12. Do you think hot chocolate mix will dissolve faster in hot water or in cold water?
13. Place each item in the correct column: Will dissolve in water; Will NOT dissolve in water.”

All of these questions involved dissolving specific materials that we discussed on the first two days and elaborated on throughout the unit. I can infer that by missing those essential hands-on experiences that enhance memory, he was unable to recall that information on the assessment.

One student who I expected to easily make a perfect score missed one question. He missed the question that many of the students missed that was discussed above. Question #3 was a higher level question that, based on previous observations and assessments, the student was completely capable of answering correctly. I believe he was still in the mind-set that I was trying to “trick” him. Some of the predictions I had the students make during their experiments were meant to be wrong so the students would think about why their prediction was wrong. I do not believe that this method was unhelpful in the learning process, but I would change the phrasing of the question next time so external factors would not affect reliability and validity.

Instructional Decisions

If, in the near future, I were to go back and fill in any missing holes on this unit I would bring back the concept that matter dissolved faster in hot water than in cold. Between them, three different student missed questions #5, 6, and 12, which all relate to heat and dissolving. Instead of throwing this concept into other lessons during the matter unit I would make an entire lesson just about temperature and the way it affects dissolving. While I was teaching this concept it seemed like I was presenting it well and memorably with a verbal explanation, relatable examples from the students’ lives (hot chocolate), a demonstration (making Jello) and experiments (a race to dissolve sugar in cold and hot water). These were all presented on different days and perhaps would have been more effective if they had been presented together with more explicit instruction.

As noted in my Table of Specifications, there were two assessment items that covered this concept. In the future I would rate these items as either medium or high priority instead of low, because it is an important part of the Virginia Standards of Learning for first grade and I do not feel like the entire class fully understood this concept.

If I ever teach this unit in the future, I will make sure to spend more time on dissolving matter in hot water as mentioned above. I will also spend more time incorporating units of measure into this unit because it is an important part of the scientific method and the students should have at least a basic understanding of units of measure to be successful in upper grades. Judging from the results on the assessment I gave to my class, the majority of the students learned the correct units of measuring temperature and mass, but one student confused temperature and mass and one student chose 'temperature' as the answer for measuring how hot something is and how much mass something has.

Validity and Reliability

The fact that the mean score in this class was 98% is a good indicator that the validity and reliability for this assessment were strong. The only question that I had trouble with and which could have been more valid and reliable was question #3, which was discussed earlier. Six out of seventeen students answered incorrectly. Three of those students are among the regular top performers and higher level thinkers in the class. The other three students who answered incorrectly are not the highest level thinkers but still should have been able to make a good prediction based on the information provided in the question.

As discussed above, the previous, tricky work completed during the unit that involved me giving the students purposefully tricky experiments to encourage deeper learning probably influenced their thought process for question #3. They had come to expect a trick and this was not what I was trying to measure. This question does not have good construct validity because it was not measuring what I wanted it to measure. Instead of measuring the students higher level thinking skills in application and what they remember about how sugar dissolves in water, it was measuring how the students had come to expect a tricky question from me.

To measure the content in this question more accurately I could have created a second question that uses the higher level thinking skills and related to the same content as question #3. By doing so I would be enhancing parallel reliability. Having responses from two of these types of questions from each student would allow me to determine if the question was reliable or if the student were just making good or random predictions.

Internal consistency reliability was strong. All of the students who missed two or more questions missed the same type of questions. As discussed above regarding the student who missed the first two days of the unit, he missed three questions that all related to dissolving. This directly correlates to the content that he missed on those first two days. By missing all three of them he showed me that it was not a mistake, but he is actually not very strong with that concept.

If I had the opportunity I would love to re-test the students to measure the content that they were able to remember over a longer period of time. Perhaps an assessment before the end of the year to judge what material needs to be reviewed would be beneficial. The test was multiple-choice and would benefit from a second go around to determine test-retest reliability to make sure the students were not just excellent guessers.

The three students who incorrectly answered the questions about temperature and how matter dissolves in heat (discussed above) shows me that the formative validity was not strong in that content area. In the future, I could take those questions and make sure to cover that content more in-depth.