

Caily Bridgeland

Test Analysis and Critique

Assessment Conditions and Purpose:

I teach a fourth grade classroom at Matoaka Elementary in Williamsburg-James City County. My homeroom class is composed of twenty-four students, with eleven boys and thirteen girls. I am part of an inclusion classroom and eight of my students have special needs: four students have Individualized Education Plans, one has a Student Assistant Plan, four of these students see a reading specialist, one works with a speech therapist, and one is an ESL student. My classroom is also diverse in its socioeconomic status; two students who live in a trailer home. There are eighteen Caucasian students in my class, four African American and two Hispanic students.

In order to provide a controlled environment that avoided systematic error, I set up the classroom in a way that student desks were spread out to prevent cheating or copying work. I also ensured that the open area outside of the classroom was not being utilized for an hour so that students would not be distracted by noise. I avoided culturally biased language and expressions in the reading passages I chose, and to accommodate students with special needs, I read the test aloud to two students whose IEP required a test read aloud to them. Also, the special education teacher requested that one of the students pulled out for reading have modifications to the test, which I allowed for a particular student. She was allowed to skip two of the short answer responses.

The purpose of the assessment was to prepare fourth grade students for the Reading and Language Arts SOL 4.5. This assessment helped assess students' present level of performance before the third quarter Developmental Reading Assessment. SOL 4.5 targets students'

comprehension of fictional texts, narrative nonfiction texts, and poetry. Students were tested on their ability to use a variety of reading strategies to comprehend short passages. They were also assessed on their ability to derive the author’s purpose, extract the main idea of a reading passage, identify problem and solution, isolate sensory words, summarize supporting details, recognize similarities between various texts, label cause and effect, and make predictions and inferences. The students were assessed on all parts of SOL 4.5, with the exception of reading with fluency and accuracy, which requires individual oral assessments.

This assessment evaluated the requirements for SOL 4.5 through the lens of the Mystery

Content	Knowledge	Comprehension	Application	Analysis	Synthesis
The student will explain the author’s purpose (e.g., to entertain, inform, or persuade).		X – M 12, 13 (63%) (100%)			
The student will describe how the choice of language, setting, characters, details, and other information contribute to the author’s purpose.		X—L 17c (88%)			

and poetry units that I introduced to the class during my student teaching experience.

<p>The student will describe in depth a character, setting, or event <u>drawing</u> on specific details from the text (e.g., words, actions, or a character's thoughts).</p>			<p>X—M 15a, 15b (83%) (75%)</p>		
<p>The student will understand that narrative nonfiction is a story based on facts.</p>		<p>X—L 1 (71%)</p>			
<p>The student will identify the facts contained in a piece of narrative nonfiction.</p>		<p>X—L 10 (50%)</p>			
<p>The student will identify the main idea or theme of a text and summarize using supporting details.</p>		<p>X—M 14a, 18a (92%) (67%)</p>			
<p>The student will identify the problem (conflict) and solution.</p>		<p>X—S 3, 14b (96%) (71%)</p>			
<p>The student will discuss the similarities and differences between text and</p>		<p>X—L 18b (50%)</p>			

previously read materials (e.g., similar themes and topics, patterns of events).					
The student will make connections between the text of a story or drama and a visual or oral presentation of the text, identifying where each version reflects specific descriptions and directions in the text. *		X			
The student will identify sensory words that describe sights, sounds, smells, and tastes, and describe how they make the reader feel.		X—M 7, 8 (96%) (88%)			
The student will refer to details and examples in a text when explaining what the text says, drawing conclusions/making inferences from text.			X—S 9, 17a, (96%) (71%) 17b, (75%)		
The student will identify cause and effect relationships		X—S 2, 4, (92%) (71%)			

		5, (58%) 6 (75%)			
The student will make, confirm, or revise predictions.					X—S 11, (71%) 14c, (92%) 17b (75%)
The student will read familiar text with fluency, accuracy, and prosody. *			X		
The student will read with sufficient accuracy and fluency to support comprehension *			X		
The student will become aware of when they do not understand, (e.g., by reflecting upon and learning to articulate what exactly is causing difficulty). *		X			

Students in the Aggregate:

Student scores out of 46 pts: 23, 25, 26.5, 30.5, 31, 32.5, 35.5, 37, 37.5, 38, 39, 39.5, 40, 40, 40, 40, 41.5, 42.5, 43.5, 43.5, 44, 44.5, 48.5
--

Overall Median	Overall Mean	Overall Mode
$39.75/46 = (86\%) = B$	$37.625/46 = (78\%) = C+$	$40/46 = (87\%) = B+$

Gender:	Girls (13)	Boys (11)
Mean	$43.5/46 = (95\%)$	$36.5/46 = (79\%)$

Classroom Type:	General Education (16)	Inclusion (8)
Mean	$40.65/46 = (88\%)$	$31.56/46 = (69\%)$

Inferences of Data:

While analyzing my Table of Specifications, I was able to make inferences about which ILOs the majority of my students mastered and others that I needed to review in the future. For example, I felt confident that the majority of my students had a high level of mastery of identifying sensory words that describe sights, sounds, smells, and tastes, and describe how they make the reader feel. The results on the table of specifications showed that 88% and 96% of students identified the correct sensory word, which I was pleased with because we have emphasized this skill during our poetry unit. I was also pleased with my students' overall results on the inferencing ILO, because this was an area where they had achieved poor scores on previous assessments and is considered to be the most challenging reading strategy. I spent a lot of time teaching my students this strategy in class, and of the three questions, 92%, 71% and 75% of students answered this question correctly. I was happy with the 71% and 75% as well because they were assessed as the second to last part of a short response question, and a few students got the wrong answer because they forgot to answer one part of the three-part question.

Overall, students were comfortable with referring to details and examples in a text when explaining demonstrated that they were able to draw conclusions and make inferences from text; most of their penalization was forgetting to include that part of their writing in the short answer response section.

I recognized that certain students mastered a part of one specific ILO, but not the entirety of the ILO. For example, for the ILO that assesses an author's purpose—to entertain, inform, or persuade—every student in the class received 100% accuracy on question number thirteen, which assessed an informative passage, while only 63% of students got question number twelve correct, which assessed the same intended learning outcome but was assessing a persuasive passage. This makes sense because fourth graders have had more experience in my cooperating teachers' classroom with informative passages and have only recently been introduced to distinguishing fact from opinion.

I also thought it was worth noting that the two de-emphasized ILOs that I spent very little time teaching were the two that both accrued a fifty percent. I think that for number ten, which dealt with identifying the facts that appear in narrative nonfiction, students had not had practice with this skill as much as they had with sensory words in the poetry and mystery unit. While our class did read a nonfiction novel earlier this year, they were rarely asked to address the specific number of facts in a small passage, while the sensory words from the poetry and mystery unit were emphasized constantly and were fresh in their minds. For question eighteen b, I accredited part of the low score to it being a three-part essay question and the final question on the test. I actually changed the point value of question eighteen from being worth a total of eight points to being worth a total of four because so many students thought that they were supposed to choose to answer one of parts a, b, or c in the question instead of all of the parts of the question. This led

certain students to fail to even attempt to answer question eighteen part b, which would make sense for a lower percentile to answer it correctly overall. I also think that there could be systematic error with question number five. For the cause and effect questions, I had three underlined portions that were causes in a row, and observed students circling all three as “cause” and then changing their answers after. Number five is the most difficult and least straightforward of the three, so I think that students assumed that one of the answers had to be an effect and second-guessed themselves even though they knew the correct answer.

There was obviously a large gap in performance on the test between general education students and students who have IEP/SAPs and/or are pulled out for special education instruction. While the general education classroom scored a mean of 88% on the assessment, the special education students scored a 69%. When looking at the specific questions that special education students scored incorrectly on, I realized that they were mostly the short answer response questions. Looking back, this makes sense because all but one of the special education students are reading below grade level. I realized that these reading passages may have been too difficult for them and that they may have misunderstood vocabulary words, which would be a threat to construct validity because the question was not assessing intended skills but instead knowledge of vocabulary.

This would make sense as to why there was also a large gap between genders in regard to performance on this test. Girls scored a mean of 95% while boys scored a 79%. In my homeroom class, there are only two boys out of the eight students in the accelerated reading group. Also, out of the eight special education students who are below fourth grade reading level, five are boys.

Analyzing the Results of Two Students:

The first child that I decided to study in depth was a new student because I wanted to spend extra time analyzing his assessment to see what his strengths were from previous instruction and where his educational gaps appeared. This student was homeschooled previous to attending Matoaka, sees a speech therapist, and also works with the guidance counselor once per week due to social-emotional issues in the classroom with his peers. Overall, he received a 39.5/46, yielding an 86%, which was eight percent higher than the classroom mean and one percent lower than the mode. I chose to analyze his assessment because he surpassed many of the students in the classroom on certain skills and ILOs (such as assessing the main idea of the passage,) which 67% of students got wrong during the short response section. I also chose to analyze this student because he lacked certain skills that many of my students have already mastered in our classroom this year and I wanted to target specific ILOs to go over again with him. While 92% of the class achieved a correct answer for number two, which was a very basic question distinguishing cause and effect, this student had difficulty with this ILO and thought that the effect came before the cause. It was clear that his unfamiliarity with cause and effect was a pattern because he also got another cause and effect question wrong, circling the answer where an effect came before the cause. Two out of his three wrong answers on the multiple choice section were about cause and effect. The other question that he answered incorrectly was number 11, which assessed the ability to make predictions, which 71% of the class answered correctly. In this question, he answered that the candidate who was more qualified would win instead of paying attention to the fact that stated that the winner was a “popular” candidate who was less qualified. I noticed a trend in his short answer responses; he enjoys writing his opinion in his work. I believe that the reason that he got this multiple choice question incorrect was because he was inserting his own opinion into the answer.

Something else that I found interesting about this student was his strength in stating the main idea in his short answer response questions. Last year my homeroom class achieved poor SOL scores overall in being able to state the main idea of a passage. This student's greatest strength is his ability to summarize the main idea very effectively. While his responses lacked detail and sophistication, he was one of the only students who answered all necessary components of each of the short response questions. While other students accrued points for their ability to use specific detail and description from the text in their passages, this student's writing was very terse, yet his analysis of the main idea itself was clear and consistently correct.

Analyzing this student's assessment helped me realize his strengths in stating a main idea and providing legitimate predictions and inferences; however, it also helped me to pinpoint his specific need to learn cause and effect and develop the skill of incorporating specific description and detail from a passage to strengthen an argument. These specific, targeted skills were helpful to his parents who were worried about the gaps in his education from being homeschooled and wished for an analysis on what he needed to work on.

The second student's assessment that I focused on was a boy in my class who received 25 out of 46 answers, which is a 54%. This student has a high socio-economic status and his parents are highly involved in the school and actively help him with individual instruction and tutoring with his homework assignments. This student has good grades in all other subjects and has always been on grade level with reading until the middle of this year. Recently there has been a slide in his reading grade, and he was the only student in the general education classroom (that was not pulled out for special education,) who did not pass his reading benchmark. I analyzed his assessment and realized his difficulty with reading passages, especially those that used sensory language or figures of speech. This student answered most concise, literal questions correctly

that other students who got A's on the test did not answer correctly. Where this student was having difficulty was predicting, inferring, and analyzing the meaning of sensory words from a reading passage. He missed all three of the questions measuring the inferencing ILO, showing an inability to understand meaning in a passage that is not stated literally. In accordance with my prediction that he understands literal statements but has difficulty with inferencing, this student answered question thirteen correctly, which showed his understanding of informative passages, but answered question number twelve incorrectly, which showed a lack of understanding of persuasive passages. I think that in the future, this student needs practice with hearing reading passages presented orally. I think that hearing the prosody and expression in the passages will help him understand them better.

Inferences about Instructional Decisions:

Administering this assessment gave me many ideas about instructional decisions that I will make in the future in regard to testing. I have several short-term goals that are specific to my current class of students: I realized that their use of description, figurative language, and written expression are sophisticated for their age overall, but that they lack basic organization of a main idea and mechanics that are essential to effective learning and writing. While they were able to utilize very specific detail from the text to describe passages and novels that they had read, they were unable to understand the relationship between the main ideas of texts to one another: exactly 50% of students did this properly. I think that I should have emphasized this more strongly in my instruction and assessment. I also am concerned that they are having difficulty with fact and opinion. While 100% of students were able to recognize an informative (factual) passage, only 63% could properly identify a persuasive (or opinion) paragraph.

Looking at this data, I will continue using free-write reading response journals, (which allows students the freedom of developing their written expression) while continuing to study sensory words with the poetry and mystery units. These forms of instruction clearly help students build written expression. However, I need to ensure that I increase the use of basic reading strategies, especially fact versus opinion, main idea, and compare and contrast.

I also have several long-term goals, especially when creating this assessment in the future. In order to utilize as many ILOs for my SOL as possible, I created a test that was simply too long. Even if broken up into two separate sessions, five short answer responses was too many for my fourth graders, especially because each had two or three requirements and different ILOs being assessed. What kept happening was that students were burnt out by the end of the test and only chose to answer one part of the question, even though I could infer that many of my students could have answered all parts of the question correctly based on their performance in class. In the future, I will be sure to only use a small number of short response items if they are going to be as lengthy as mine were. Students had a pattern of answering the first short response question correctly and writing an extremely detailed and lengthy response—92% of students got number 14 parts a and c correct, while only 75% of students got a correct answer for question number 17 b, which assessed the same ILO. I am not good at multiple choice questions myself, which makes me value and want to provide a diversity of questions to accommodate different learning styles and strengths. I do see value in short answer responses, but I believe that they have to be chosen carefully and sparingly so that students do not spend too much time on them and lose momentum and focus for other questions.

Conclusions based on Validity and Reliability:

I knew that SOL 4.5 would pose a challenge in promoting reliability when creating this

assignment, because this SOL assesses a students' ability to comprehend reading passages, and no two reading passages can be the same. Reading passages automatically give certain students advantages over others because they relay information about a topic which certain students may be more familiar with. Construct validity is difficult for this assessment, because a student could do well on a reading passage simple because he or she has more experience with a topic in the passage over another student. If this background knowledge helps a student answer a question correctly instead of his or her ILOs, this is a threat to construct validity. To avoid this, I did everything in my power to use straightforward language, and only assess students on facts that are given expressly in the passage.

This particular SOL is also difficult in establishing face validity and concurrent validity, because it is impossible to pick passages that are exactly the same. It would be extremely difficult to recreate a similar test because different reading passages are needed. Unfortunately, a certain amount of systematic error is present with reading passage questions because any passage can be culturally biased, especially if it is an opinion passage.

One way that I will attempt to ensure predictive validity in the future would be to assess fewer ILOs at once, but have more questions measuring each type of ILO. If a person is assessed on a greater number of questions, he or she is less likely to be impacted by one question that produces a negative or confused response based on the language or content of one specific question.

I also think it was a huge mistake to have my final question on the assessment be worth eight points, (which is twice as much as the other short response questions). Students had spent too much time writing short response essay questions and were tired so kept forgetting to

respond to certain parts of the question. My cooperating teacher suggested that I change the point values and have each multiple choice question be worth two points instead of one, and have question number eighteen be worth four points instead of eight due to student fatigue, carelessness, and an interruption of the test for some students when the allotted time period to take the test was over. Next time, I will be sure to have a shorter test with fewer short response questions, fewer ILOs at once, but a larger number of questions assessed for each ILO.

I will also increase reliability when assessing the short answer responses. Even though I had a rubric that was extremely helpful, I think that it would promote inter-rater reliability to have someone else grade my students' assessments to ensure that I am not unconsciously influenced by factors that are not the intended criteria of learning, such as taking into account how a student usually performs on tests, how hard a student works, or how much I like or relate to a student. I will also promote intra-rater reliability in that I will break up my grading into smaller periods of time so that I do not give a student a lower grade because I am tired after a long session of grading. I did attempt to do this when grading the assessments, but the short-answer responses took so long to grade that I kept insisting that I graded a certain number of tests per day so that I could budget my time effectively.

Overall, giving this assessment was an extremely educational experience and helped me brainstorm how I hope to assess students in the future. I will always use a table of specifications and unpack each of the intended learning outcomes when I am assessing a new SOL.