**Assignment #2 Research into Theory and Classroom Praxis**

Madison Jerome, Siddhi Nirajbhai Shah and Avery Kotylak

Department of Education, University of Regina

ECS 401: Critical Perspectives of Assessment

Dr. Cristyne Herbert

**Brainstorming and Final Five Questions**

**Brainstorming Questions Developed as a Group**

1. What are some techniques you use in your classroom for assessment? Summative vs formative?
2. Does this change subject to subject?
3. What does triangulation look like within your school district?
4. What evidence do you need/how do you properly assess students to ensure that they are learning? I.e a student in grade 2 who is reading at a grade 1 level? (how do ensure your assessments are valid and reliable)
5. What creative assessing strategy have you used to engage student learning and peak interests?
6. What is your assessment philosophy?
   1. Within your philosophy do you believe that grading is an effective way to assess students' learning? Why or why not? If not, how do you provide evidence of your students' learning?
7. How do you make your assessments equitable/culturally responsive for all students?
8. Since we are shifting away from grade averaging and testing, how do you tell the students' story of achievement?
9. Thoughts on late marks/bonus marks?
10. What are the challenges you have faced with assessment in your classroom?
11. How does assessment affect students' self-esteem and confidence?
12. What advice would you give a new teacher about assessment in the classroom?
13. What is your strategy to design differentiated assessments based on students' different funds of knowledge?

**Final Five Interview Questions**

1. How do you engage in continuous assessment in your classroom? What do you feel is the relationship between validity, reliability, and your own professional judgment?
2. If your assessment of students reveals that students are below grade level, how do you differentiate your instruction and assessment to support these students?
3. How do you make your assessments equitable/culturally responsive for all students?
4. Can you identify one or two challenges you have faced with assessment as an elementary educator and explain them? How have you overcome these challenges?
5. What would you say is the most important thing for a brand new teacher to keep in mind when it comes to assessment?

**Literature Review of Reliable and Valid Assessment within the Classroom**

Classroom assessments are deemed credible based on two types of criteria, validity and reliability (Cizek, 2009). When reviewing reliable and valid assessments, it is important to understand the meaning of both terms within a classroom context. Assessment validity is evident when an assessment evaluates what it is supposed to assess whereas assessment reliability is the ability of an assessment to repeatedly and consistently obtain the same results (Cizek, 2009; Marzano, 2017). While both are highly important to evaluate the quality of an assessment, there is one that takes precedence (Cizek, 2009; Marzano, 2017).

When designing classroom assessments, Cizek (2009) and Marazano (2017) explain that validity is considered to be a fundamental standard. This is because it measures the extent to which a students' knowledge and skills are justified (Cizek, 2009). Messick (1993) explains that there are three types of validity which include criterion-related validity, construct validity and content validity (as cited in Marzano, 2017). Hathcoat (2013) also explains that the perspective, whether instrumental or argument based, used to examine these three types of validity varies their meaning (as cited in Marzano, 2017). Marzano (2017) continues to expand on the instrumental perspective of classroom assessments and states that all three types of validity must be found in assessments in order to be valid. For criterion-related validity to be evident, from the instrumental perspective, one must compute and compare the new test to one already deemed valid (Marzano, 2013). Within the instrumental perspective construct validity is the ability to examine and justify qualitative data found in classrooms where content validity examines what it is meant to examine (Marazano, 2017). Marazano (2017) argues that by using the instrumental perspective, teachers are unable to control criterion-related and construct validity and that teachers have more control using an argument-based perspective to validity, therefore, it is more suitable for classroom assessments.

The argument-based perspective on validity is all-encompassing (Morazan, 2017). Rather than focusing on the assessment alone, in the argument-based perspective, teachers can use multiple resources to gather data to ensure that students are understanding what is being taught (Morazan, 2017). For example, criterion-related validity, with an argument-based perspective, uses formative assessment data to predict summative assessment data (Marzano, 2017). Marzano (2017) explains that using perspectives are critical for validity in assessment, where Bonnor (2013) suggests there are principles that contribute to it.

Bonnor (2013) provides five principles which should contribute to the validity of a classroom assessment. These principles include assessment highlighting instruction, limiting any and all bias, emphasizing process over the product through observation and feedback, evaluating the effects of one’s own assessments and incorporating all stakeholders to ensure validation (Bonnor, 2013). Bonnor (2013) also explains the importance of systemic validity, which states that assessments are “components of a dynamic system, and thus decisions about assessment methods provide feedback that determines the future of the system” (p. 101). Another critical concept when it comes to validity is to consider cultural validity.

Cultural validity is the ability of an assessment to address the influence of socio-cultural factors on students (Basterra et al., 2010). Language is one thing to consider as it is how humans gain understanding (Basterra et al., 2010). In order to validate assessment for those with English as a second language, Bassterra et al. (2010) suggest differentiating the assessment to the students strengths which may be adjusting question types, oral journalling, portfolios or graphic organizers. When differentiating assessments, teachers should also consider reliability.

Morzan (2017) explains that in order to achieve assessment reliability in the classroom, one must follow a new paradigm based on precision. Through this lens, it is encouraged for teachers to use the scores of individual students from parallel assessments (Morzan, 2017). Assessment reliability in the classroom would typically follow a test and retest procedure (Cizek, 2009). Cizek (2009) explains that another type of reliability within the classroom would be scoring reliability. This is then divided into intrarater consistency, consistency within a scorer, and interrater agreement, agreement of many (Cizek, 2009). Youn and Chen (2021) present three rubrics and their relation to reliability in aid to increase reliability in classrooms.

An analytical rubric uses many categories to assess, these will then be averaged and a grade will be awarded (Youn & Chen, 2021). However, due to the time consuming nature of this rubric, intrarater consistency may be limited and therefore reliability is not as distinct (Youn & Chen, 2021). The second rubric type that Youn and Chen (2021) review is a checklist rubric. This rubric allows students to review what they have and have not done within the assessment (Youn & Chen, 2021). It is expressed that checklist rubrics are time consuming which can affect the intrarater consistency (Youn & Chen, 2021). A holistic rubric, Youn and Chen (2021) explain, is quite reliable as students are given only one mark based on their performance. Although, where it supports reliability it lacks validity (Youn & Chen, 2021).

As previously discussed, validity and reliability are two criteria that are a requirement for credible assessment within the classroom (Cizek, 2009). Because validity is one of the most important criteria within an assessment, there are multiple ways to establish a valid assessment (Cizek, 2009; Marzano, 2017). One must consider the perspective of validity, systemic validify and cultural validity (Basterra et al., 2010; Bonnor, 2013; Marzano, 2017). In order to create reliability within classroom assessments, Marzano (2017) suggests reviewing scores from parallel assessments. Ensuring one’s assessments have validity and reliability is essential for student learning.

**Interview Learning Summary**

During the interview, the teacher shared insights about their experience with continuous assessment, validity, reliability, and professional judgment in the classroom. They discussed the challenges faced while teaching in an online environment and the difficulties of enduring students’ accountability and authenticity in their work. The teacher highlighted the benefits of one on one assessments and individualized learning which encourages students to be more likely to engage and demonstrate their understanding. The interviewee emphasized ensuring that they align their teaching with the learning outcomes and indicators found in the curriculum.

The teacher also addressed the issue of accuracy and consistency in assessments. They acknowledged that different conversations and contexts might impact the consistency of assessment across different semesters or academic years. Additionally, the impact of test writing conditions on students’ performance was mentioned. The teacher had advocated for using professional judgment to assess students’ understanding rather than solely relying on students’ test performance and emphasized the need for alternative assessment methods to capture students' true abilities.

Regarding equity and cultural responsiveness, the teacher stressed the importance of using inclusive examples and questions encompassing various cultures and backgrounds. They emphasized the need for teachers to be aware of cultural differences and avoid using examples that might exclude or marginalize students. It was suggested to incorporate diverse perspectives and practices to create a more inclusive learning environment. This led us to inquire further into how assessment can be tailored to meet the needs of students with diverse cultural and religious backgrounds while respecting their beliefs. Montenegro and Jankowski (2017) elaborate on this topic by explaining it is important to differentiate assessment so that it is culturally relevant. This will allow students to be more engaged in their own learning (Montenegro & Jankowski, 2017).

In conclusion of the interview, the teacher discussed their own experiences with continuous assessment in different teaching contexts, specifically in an e-school and a visually impaired program. They mentioned the challenges of ensuring the authenticity of students' work in an online environment and the benefits of individualized assessments and one on one interactions.

This interview aligns with the research in terms of the importance of validity in assessments and the need to consider the purpose of the assessment (Cizek, 2009; Marzano, 2017). She emphasized the significance of validity and stated that assessments should align with the intended outcomes and indicators. She has also highlighted the need for assessments to be purposeful and meaningful.

Regarding reliability, the interviewee touched upon the issue of consistency and accuracy in assessment practices. They mentioned that using the same test repeatedly may not account for the different conversations and content covered in different semesters or over the years. In the interview, the teacher also recognized the impact of external factors, such as test writing conditions and students’ mental health, on test performance. The observations correspond with the research on assessment reliability in which Marzano (2017) discussed factors like test-related procedures, scoring consistency, and integrated agreement. The interviewee acknowledged that professional judgment and alternative assessment methods, such as hands-on tasks and one on one conversations, could provide a more reliable indication of students' knowledge.

It was further discussed, in the interview, that individualized learning is important as it pushes students to reach high levels of achievement. They mentioned working with students at different proficiency levels and encouraging collaboration among students to support their learning and growth because students use language to aid in their understanding (Basterra et al., 2010). The interviewee used strategies throughout her career similar to what Bassterra et al. (2010) suggested which included oral journalling and graphics. While this aspect of the interviewee’s response did not directly align with the research on validity and reliability, it reflects a teaching approach that considers the diverse needs and abilities of students talked about in our research (Basterra et al., 2010).

Overall, the interviewee’s responses demonstrated an understanding of the concepts of validity and reliability in assessment, particularly the importance of validity in ensuring the purposefulness of assessments and the challenges of maintaining consistency and accuracy. Her emphasis on individualized learning and the use of professional judgment align with the research on alternative assessment methods and consideration of socio-emotional factors in the assessment process.

During this interview, the teacher shared that they create tests or assessments that focus on students’ ability to show learning rather than assessing the ability to recall. This made us realize that there is a need to create assessment tools that focus on the demonstration of learning and not rely on the cognitive capacity of memorization.

The teacher had expressed the need to find assessment tools and strategies to measure the learning through e-classes. The assessment tools should be fair and shield enough that it is eliminating the scope of manipulation while the students are giving the assessment using technology. This led to our research into the technology available for online assessment. Langenfeld et al. (2022) explains how technology can be leveraged to create more authentic and accountable online assessments through the use of digital-first learning and assessment systems. This system is set to create an individualized experience for the learner by allowing students to become autonomous within their digital classroom (Langenfeld et al., 2022). It engages students through decision making and self-reflection (Langenfeld et al., 2022).

**Correlation Between the Research Topic and Interview Summary**

One must consider the perspective of validity, systemic validity and cultural validity (Basterra et al., 2010; Bonnor, 2013; Marzano, 2017). The same outlook is also observed in the interviewee’s assessment philosophy who explained an incident that occurred during the e-class where some parents of grade seven students refused to give consent to discuss sexual education despite this topic being a part of the curriculum. This meant the interviewee needed to adapt and chose discussion to be a tool of assessment. This allowed students an opportunity to share their responses in an inclusive environment without the fear of being judged or questioned.

After the interview, a deeper contemplation led to the emergence of several new questions. This inquisition is paraphrased in the following question for more exploration. What are effective methods for ensuring consistency in assessments across different semesters or academic years? What alternative assessment methods can be used to accommodate students' individual circumstances and minimize the impact of test anxiety and mental health?

**Resources**

Basterra, M. D. R., Trumbull, E., & Solano-Flores, G., (2010). *Cultural validity in assessment: Addressing linguistic and cultural diversity.* Taylor & Francis Group.

Bonner, S., (2013). Validity in classroom assessment: Purposes, properties, and principles. *SAGE Publications, Inc.,* https://doi.org/10.4135/9781452218649

Cizek, G.J., (2009). Reliability and validity of information about student achievement: Comparing large-scale and classroom testing contexts. *Theory into Practice*, *48*(1), 63–71. https://doi.org/10.1080/00405840802577627

Langenfeld T., Burstein J. and von Davier A.A., (2022). Digital-first learning and assessment systems for the 21st Century. *Frontiers in Education, 7.* https://doi.org/10.3389/feduc.2022.857604

Marzano, R.J., (2017). *Making classroom assessments reliable and valid : How to assess student learning.* Solution Tree.

Montenegro, E. & Jankowski, N.A. (2017). Equity and assessment: Moving towards culturally responsive assessment. *National Institute for Learning Outcomes Assessment.* Retrieved May 29, 2023, https://files.eric.ed.gov/fulltext/ED574461.pdf

Youn, S. J., & Chen, S. (2021). Creating valid and reliable scoring rubrics for performance-based classroom assessment. *European Journal of Applied Linguistics and TEFL, 10*(1), 25-44. Retrieved May 26, 2023, https://login.libproxy.uregina.ca:8443/login?url=https://www.proquest.com/scholarly-journals/creating-valid-reliable-scoring-rubrics/docview/2516290287/se-2