

Ensuring Assessment Integrity

Increasing Assessment Validity and Student Engagement

Adisack Nhouyvanisvong, Ph.D.

May 2020

Naiku is a next generation assessment platform, providing teachers with comprehensive assessment tools to help teachers collect data about their students to make informed instruction.



12100 Singletree Lane, Suite 185
Eden Prairie, MN 55344
Phone: 612-356-2458
E-Mail: info@naiku.net
Web: www.naiku.net

Introduction

Educational assessments should provide important data to teachers and administrators to serve classroom, school, and district improvement needs. Teacher's ongoing use of assessment to guide and inform instruction has been shown to be one of the most effective instructional strategies to improve student learning (Black & Wiliam, 1998). Data from common formative assessments is a principal component of effective PLC teams (DuFour, et al). Results from school/district assessments such as benchmark assessments can inform teacher instruction and planning, decision-making at the classroom, school, and district levels, and inform educational policy.

Whether the data is being used by individual teachers in the classroom, within PLC teams, or by administrative leaders, it is important that the assessment data is reliable and valid to support sound educational decisions. Invalid or inaccurate data caused by student cheating can render data-driven instructional decision making ineffective or worse.

Mitigating the effects of student cheating is important anytime, but particularly so if students are taking assessments outside of the classroom without teachers present. This paper outlines strategies to help ensure the integrity of assessment data in remote learning environments.

Purpose of Assessments

Educational assessments serve many important purposes as part of a comprehensive assessment system. Four important purposes from Data Use (n.d) are highlighted below:

1. **Communicate expectations for learning.** It is important for students, teachers and parents to know what knowledge and skills are important to learn and how learning will be measured. Assessments communicate these expectations frequently to the various stakeholders. It is important to vary the assessment tasks and item types (i.e., include essays, performance tasks, constructed response items, not just multiple-choice items) to communicate that various knowledge and approaches to learning are valued.

2. **Plan instruction.** Assessments provide teachers with immediate and actionable data about student learning and performance. Teachers need information to develop and adjust curriculum and instruction to best serve their students. To do so, assessment results must be aligned to content and learning targets and provide accurate and valid feedback on students' strengths and weaknesses relative to those learning targets.
3. **Monitor and evaluate learning.** Assessments can be used to monitor and evaluate learning. They provide educators with information on the effectiveness of educational programs and curricula. Data from assessments can help administrators make mid-course corrections or modification where student performance is lagging or highlight an educational program or curriculum that helps students excel.
4. **Predict future performance.** With annual statewide accountability tests given at the end of the school year, assessment results can provide data to predict whether students, classes, schools, or districts are on course to meet those year-end goals. Results can be provided at individual and group levels to identify those who need more help and those who excel and may benefit from more advanced instructional programs.

Quality Criteria for Assessments

Regardless of whether assessments are purchased or newly developed, assessments should meet certain criteria for quality. Below, four criteria from Data Use (n.d.) for the selection and use of assessments are summarized:

1. **Validity.** Validity is often thought of as a criterion of whether an assessment measures what it purports to measure. More accurately, it is the extent to which the inferences made from the assessment results support the purpose for which it is used. A benchmark assessment may be valid for one purpose but have little validity for another. When evaluating assessments for validity, it is important to consider whether the assessment is assigned with district and school

learning goals, provides reliable information for the intended use and score interpretation, is instructionally sensitive, provides useful reporting for intended users and purposes.

2. **Reliability.** Reliability refers to the expected consistency of test scores. Assessments with high measures of reliability provide consistent results across time, across forms, and across settings. When evaluating assessments for reliability, users should examine the item difficulty (appropriate difficulty is in the range of 0.2 to 0.8), item discrimination (desired values are greater than +0.3), and reliability coefficient (test-retest, parallel forms, or internal reliability coefficients should be 0.8 or greater.)
3. **Alignment.** Benchmark assessments should be aligned to the standards or learning targets that the school or district is teaching. Alignment of the assessment to the learning goals is paramount. When evaluating assessments for alignment, also consider the framework used to develop the assessment and items, the distribution of items by cognitive level demands, and the range of items for diagnosing specific learning strengths and weaknesses.
4. **Utility.** Ultimately, assessments must have utility. The assessment results must be useful and help teachers and administrators make effective instructional decisions. When considering assessments for utility, look for assessments that are user-friendly, easy to administer, scored in a timely fashion, and include robust reports useful for different stakeholders.

Strategies to Maximize Assessment Integrity

The following strategies can be used in combination to aid in minimizing the effects of student cheating.

Open Book Assessments

Open book assessments mitigate cheating as, by definition, looking up information through additional sources is allowed. To be effective, assessment questions should challenge students, even if they have

Ensuring Assessment Integrity

access to outside informational resources. Questions should principally be at the Application or Analysis level (or higher) of Bloom's Taxonomy, rather than Understanding or Remembering.

Lock-down Browsers

Student use of lock-down browsers prevents students from opening other applications or tabs to look up answers. In online assessment within the classroom this can be an effective way to minimize students "Googling" answers. Lockdown browsers for use with online assessment solutions are freely available, such as the university consortium developed [SafeExam Browser](#). Assessment software providers may also provide custom solutions for specific platforms, such as Naiku's Chromebook app.

Unfortunately, student use of lock-down browsers may not be an option when students are outside of the classroom. In addition, the use of a lock-down browser on the student's test-taking device does not prevent a student from using an additional, separate device – such as their phone or tablet – to look up answers. Teacher observation/proctoring within the classroom can mitigate this, but outside of the classroom other measures outlined in this paper should be used.

Time-limited Assessments

Offering time-limited assessments can help mitigate the ability for students to cheat as they won't have enough time to look up the answers. Educators at Penn State University suggest that 1 minute per multiple choice item is sufficient for students who are prepared (Online Assessment, n.d).

Assessment management software may facilitate this process. For example, teachers in Naiku can set a time limit. Students will see a countdown timer when they start the test, and can only submit answers, not edit, once the time is expired. Teachers can grant additional time as needed.

Online Proctoring

Proctoring is a time-honored method for discouraging cheating, though it may be challenging to accomplish online. Some university online learning systems utilize the student laptop camera to view the student while they are taking the test. This may not be feasible for K-

12 schools, but teachers may have access to other methods of online proctoring. For example, Naiku has a Watch Progress report for teachers to view student progress in real-time. Teachers can see if students are “unfocused”, i.e. they have switched to another tab or application and are no longer focused on the test. Teachers can also view student and class progress in real-time to identify students that are struggling or questions that are difficult for the class.

Additionally, post-test reports provide teachers detailed information such as how long students spent on the assessment, as well as on each question. This can be helpful to see if they were rushing their work or perhaps simply copying answers.

Item Complexity and Type

Utilizing higher complexity level questions and a variety of items types beyond multiple choice can help mitigate cheating. Asking students to infer, interpret, or summarize information presented (Understanding level, Bloom’s Taxonomy), or correlate, calculate, or estimate (Analyzing, Bloom’s) are examples that require more than simply looking up or remembering information.

While higher complexity level questions can be designed in multiple choice format, utilizing other format types can often be a more effective way to assess student knowledge. Though essay or narrative responses serve this purpose, there are many other item types that teachers can use within assessment software that can effectively assess higher order thinking while also have the benefit of automatic scoring of student responses. As an example, students could be asked to arrange sentences within a paragraph, or categorize information provided in a science experiment.

Assessment management software can have item search features so teachers can quickly find question items that meet specified criteria – not just academic standard, but also other attributes such as item type and DOK or Bloom’s complexity level. These, particularly in combination with included professional item banks and/or shared community item banks, help facilitate assessment creation.

Randomize Assessment Items/Options

Randomizing the order of assessment items and/or answer choice options (for multiple choice and multiple response item types) minimizes cheating from sharing answers between students.

Some online assessment software provides for random question selection from a bank of questions, for example, randomly pulling 10 questions from a bank of 50. This method is best used for formative assessment or as a practice test. Because the questions are different and not guaranteed to be the same complexity level, it is not suitable for common assessments for teacher teams or school/district assessments.

The Algorithmic item type is a specific question type that provides randomization while maintaining complexity level. Particularly well suited for math problems, with this item type the numbers within a question can be randomized for each student within set boundaries. As an example, the question “ $4 + 6 = ?$ ” could be randomized such that one student sees “ $2 + 8 = ?$ ” and another sees “ $3 + 6 = ?$ ”. In this example the first number is randomly generated between 1-5 and the second from 6-9 for each student.

Show Student Work

Asking students to show their work or explain their answer, particularly when using higher complexity questions, can help mitigate cheating.

Essay type and other narrative questions naturally lend themselves to this. Some online assessment software may additionally allow for uploading of additional student work as attachments such as project photos, presentations, or multimedia clips, for example.

Showing student work can be particularly relevant with fill-in-the-blank questions. This question type can be automatically scored with the answer justification as an additional measure to help validate answer integrity. As an example, Naiku’s “Tell-me-more” feature provides a free-field journaling box for the student to provide additional information to justify their answer. This information is correlated with their answer as it is presented to their teacher, as well as for post-test student reflection.

Summary

In summary, educators can increase assessment validity when using online assessments, even in a remote learning environment, by using multiple strategies to engage students and minimize the chance of student cheating. Using test questions that require students to employ higher thinking skills and/or test questions that request students to explain their answer can be used in combination with assessment administration options such as randomization and time limits to minimize the likelihood of cheating.

References

Black, P. & Wiliam, D. (1998). Inside the Black Box: Raising standards through classroom assessment, School of Education, King's College, London, United Kingdom.

Data Use (n.d). Assessment Systems: Benchmark Role. Retrieved from http://datause.cse.ucla.edu/ba_comprehensive.php

Ensuring Assessment Integrity

Data Use (n.d). Purpose of Benchmark Assessments: Retrieved from http://datause.cse.ucla.edu/ba_purpose.php

Data Use (n.d). Selecting and Using Benchmark Assessments. Retrieved from http://datause.cse.ucla.edu/ba_criteria.php

Online Assessment (n.d). Academic Integrity. Retrieved from <https://sites.psu.edu/onlineassessment/analyze-evidence/>

About the Author

Dr. Adisack Nhouyvanisvong is an expert in educational assessment, including computer-based and adaptive testing. He has created and ensured the psychometric integrity of large-scale educational assessments for states and national organizations, taught at the University of Minnesota, and is an adjunct professor at St. Mary's University where he has taught educators and graduate students in educational assessment practice and instructional strategies. He has been published in peer-reviewed journals, regularly speaks at education conferences and is currently President of Naiku.