

The Longer-Term Impact of COVID-19 on K–12 Student Learning and Assessment

Kyndra V. Middleton, *Howard University*

Abstract: *Due to the precipitous onset of the coronavirus disease, teachers and students across the nation were thrust into a new environment, and the impact of this new experience will be felt both shorter and longer term. This academic year saw “test pollution” with the switch to online instruction, and student learning was significantly impacted by stress, anxiety, illness, being forced to learn in a vastly different method than previously experienced, and the increased potential to fall behind due to lack of access to materials. Classroom assessment, teaching and learning, and measurement and interpretation of student growth are among the numerous areas that have been affected by the sudden switch of schools to online instruction that will require much thought in order to examine the impact of the significant deviation from the classroom norms on which much of previous research has been based. Educators, educational researchers, and policymakers have been presented with a challenge that does not have a definitive answer. There are many unknowns that remain as schools plan to move forward with instruction. However, through collaboration, the knowledge that each of these professionals can contribute ensures that adequate decisions will be made that will benefit all students equitably.*

Keywords: classroom assessment, COVID-19, student growth, teaching and learning

The sudden onset and extensive spread of the coronavirus disease (COVID-19) has undoubtedly created a more immediate significant impact in the lives of many across the nation. Teachers and students across the elementary and secondary levels were thrust into a new teaching and learning environment that most had not experienced before. In the area of education and the field of educational measurement, the longer-term impact will be felt and realized well after parents return to work and children return to school.

How COVID-19 Has Impacted Classroom Assessment

According to Ebel and Frisbie (1991), evaluation in education asks questions such as “How good is the performance?” “Have they learned enough?” and “Is their work good enough?” (p. 23). Each of these evaluative questions require judgment, and in fact, evaluation of student performance has been defined as the process of making a value judgment about the merit of the student’s academic accomplishment to aid in the improvement of future student learning (Brookhart & Nitko, 2019; Olin & Sullivan, 2002; Zughoul et al., 2018). Assessments have been one such method to inform student evaluation. Both measurement specialists and researchers have posited that assessments and procedures originally intended for other purposes are inadequate and inappropriate for the current situation and that the resulting test scores are being “polluted.” Messick (1984) described the term “test pollution” broadly as a systematic increase or decrease in test scores unrelated to the content domain represented on

a test. There are three main sources of test score pollution: (a) test preparation, (b) situational factors (e.g., stress, anxiety), and (c) external factors (e.g., socioeconomic status) according to researchers (Chalak & Tavakoli, 2010; Haladyna, Nolen, & Haas, 1991). This academic year saw “test pollution” across each of these sources with the switch to online instruction. Student learning (and hence test preparation) may have been impacted more than typical by stress, anxiety, illness, being forced to learn in a method that was vastly different from what they were used to, and the potential to fall behind due to lack of access to the materials to learn (Internet, quiet space, etc.).

The purpose of classroom-based assessment is to inform instruction, whether this be through the decision-making process of identifying strengths and weaknesses, for grading decisions, placement of students, to motivate students to study and learn, or for increasing/decreasing the pace of instruction (AERA, APA, & NCME, 2014; Brookhart & Nitko, 2019; Cizek, 2009; Goertz, Oláh, & Riggan, 2009). If the assessment scores (inclusive of informal and formal assessment through testing or observation) produce too much construct-irrelevant variance, the interpretation of the assessment results become invalid. While it is factual that scores on tests taken during class time while monitored by a teacher may be invalid for various reasons (e.g., cheating, illness, mental state, lack of focus or concentration, etc.), this may have become more of a problem during the last section of the academic year since students were not monitored when completing tests or assessments. For example, did the

student use notes or the Internet when he or she would not have been allowed to do so during class? Did someone else complete the test for the student?

The Impact of COVID-19 on Learning

In the educational world, students have also experienced deficits in learning as a result of new approaches to teaching that were implemented. In numerous online surveys conducted since March, teachers have consistently indicated they were not prepared to teach online and that a significant percentage of their students did not even log in to complete assignments. Teachers also taught less new material, especially in high-poverty schools (DeWitt, 2020; Herold & Yettick, 2020). This lack of preparation unquestionably impacted teaching ability which in turn impacted student learning since teaching leads to learning and learning can inform teaching. The current discussion will focus on student learning and assessment issues that have arisen or have been intensified by the coronavirus pandemic.

The necessary response to the pandemic, which significantly impacted instruction for almost 3 months of the 2019–2020 academic year and has the potential to continue its impact into the 2020–2021 academic year, included quickly moving schools and universities to online instruction without much time for training and implementation of standardized procedures. While classroom instruction is typically not fully standardized, there are some standardized components within schools. Teachers within schools have autonomy in instruction; however, this autonomy comes with common expectations for students. For example, the teacher teaches a lesson for a class period with certain procedures built in to assess student understanding and performance. Traditional classroom-based tests are even more standardized than classroom instruction with instructions provided on remaining quiet, students not using any aids to help with the test, etc. With the switch to online learning, the standardization of both teacher and student performance was not easily attainable.

The Impact on COVID-19 on Teaching

Teachers within the same schools responsible for teaching the same subject approached teaching differently. This difference may be due to a lack of knowledge of evidence-based pedagogical approaches to teaching online, lack of knowledge of technology, family/personal issues, illness, or many additional reasons. What is known from surveys and interviews conducted is that students within the same school had varying levels of instruction ranging from minimal at best to innovative (DeWitt, 2020; Read, 2020; Schwartz, 2020). For example, some teachers provided assignments and requested they be submitted. There was no teaching of new material involved so students essentially reviewed what they were previously taught for the last 3 months of the academic year. Other teachers uploaded short asynchronous videos for student learning while others held synchronous lessons with students. This lack of standardization of online instruction can and will pose problems when students return to classes in the fall. The students who were in a class together during the spring will no longer be in the same classes. There has already been a concern that students of varying abilities are sometimes placed in the same class. This issue will be intensified this fall with the potential for an even wider variety of learn-

ing abilities that teachers will have to address which means instructional time for new material will be delayed.

The Impact of COVID-19 on the Interpretation of Student Growth

Classroom instruction and assessment help both teachers and students recognize strengths and weaknesses as well as prepare students for standardized tests. The necessity to review older material longer and the delay in learning new material can have lasting effects on student growth measures. The divergence in classroom instruction during the last quarter of the academic year will impact standardized statewide and national test scores for at best the next 2 years but at worst even longer. Classroom instructional divergence creates additional variance in test scores so the ability to compare the same student's test performance relative to others across last year and this year will be problematic. To illustrate, the calculation of student growth which compares a student to his or her peers who received similar previous scores on a prior assessment, if utilizing student growth percentiles (SGP), has been impeded since there will no test score for the 2019–2020 academic year. Thus the 2018–2019 to 2019–2020 growth as well as the 2019–2020 to 2020–2021 growth will be unable to be calculated. Even if schools are able to test in the fall to make up for the missing spring testing, there will be 1.5 years' worth of growth for the first comparison and only 0.5 year worth of growth for the second comparison. Further, the variability in instruction will likely create larger variance in the scores due to the varying degrees of instruction and learning as a result of online schooling (e.g., instruction type, lack of Internet access, mental state changes due to being impacted by the virus, etc.). Research will be needed to compare past 2-year SGPs (2017–2019) and current (2019–2021) SGPs once established instead of 1-year SGPs to determine whether there is a differential impact of instruction, and hence an impact on SGPs, as a result of COVID-19 (Betebenner & Van Iwaarden, 2020).

Acknowledging that the loss of instructional time has been disadvantageous to student learning, some school districts have encouraged students to attend summer school to make up for lost learning. Due to the ongoing pandemic, only a small percentage of these course offerings are face-to-face, and the most vulnerable populations are falling farther behind with online instruction. Some of these students have unstable housing or are homeless while others lack reliable Internet access or computer accessibility (Tornay, 2020). Despite the challenges that have resulted from the sudden necessity to switch to online instruction, there is some promising news as well. School closures has allowed some districts to use money saved to offer summer instruction for elementary students and smaller in-person classes for migrant students which allows for more hands-on experiences. Others are offering robust hands-on, smaller classes with local partnerships while others have turned summer online classes into a video game format, complete with the ability to earn prizes (Castelo, 2020; Hammond, 2020; Tornay, 2020). Just as with the range of instructional methods utilized during the last quarter of the academic year, summer course offerings and opportunities range from districts offering no summer learning opportunities to districts offering expanded opportunities.

The inequities that exist in learning will widen as a result of this non-uniform implementation of summer school. While

many school districts are doing the best they can under the circumstances, the already-troubling test score gap will most likely widen since studies have shown under normal circumstances, summer learning loss exacerbates the inequities that exist between certain subgroups such as those from lower socioeconomic class and students with disabilities, and this loss equates to about 1 month of learning (Alexander, Pitcock, & Boulay, 2016; Cooper, Nye, Charlton, Lindsay, & Greathouse, 1996; Kerry & Davies, 1998). The student who chooses or is able to take advantage of the summer school offering may be quite different than the type of student who chooses not to or is unable to take advantage of summer school. Ideally, the already low-performing student would be the student who educators hope would enroll in summer school. However, this may not be the case which could further widen the educational learning gap which leads to a widening of the test score gap.

Adding the limited instructional time students have had over a 3-month period to the summer learning loss over an additional 2- to 3-month period creates an additive effect of learning deficit that teachers will be forced to attempt to make up for once schools resume. This additional review time will hinder new classroom instruction, creating a ripple effect for the next academic year which can and will lead into the following academic year unless deliberate steps are taken throughout the next year to both review the old knowledge that was lost and teach the new material quickly and in-depth enough in a way that is not too accelerated and/or cursory so some students do not get left behind. It will indeed be a very delicate balance that must be maintained.

Conclusion

Each of the topics mentioned previously are serious issues that can significantly impact the measurement of student learning in K–12 education and that will require thoughtfulness and additional research to examine the impact of the significant deviation from the prior classroom norms which much of data and research have been based upon. While some in educational research have previously utilized hierarchical linear modeling (HLM) as a method to address certain research questions, one might argue that now more than ever is the time to utilize HLM during the current unprecedented educational environment since this statistical method will allow for within-school variance to be better accounted for. Without utilizing a method such as HLM, the large variability within school will be ignored which will in turn affect the reliability and validity of any study done examining academic achievement during this time. Whereas a previous typical research question may have included the different methods teachers used for instruction (the how), the current research question would need to include whether teachers taught or not. This error variance can be accounted for in HLM, but the sample size requirement may pose an issue since the recommendation for minimum sample size at any level is 30 groups, and even then, the estimated standard errors are too small (Maas & Hox, 2005).

Educators, educational researchers, and policymakers have been presented with perhaps the biggest challenge of their careers. There are many unknowns that remain as schools plan for moving forward but create contingency plans in case that is not possible. However, through working together, the knowledge that each of these professionals

contributes to the educational measurement field ensures that adequate decisions will be made that will benefit all students equitably.

References

- Alexander, K., Pitcock, S., & Boulay, M. C. (Eds.). (2016). *The summer slide: What we know and can do about summer learning loss*. New York, NY: Teachers College Press.
- American Educational Research Association, American Psychological Association, National Council on Measurement in Education, Joint Committee on Standards for Educational, & Psychological Testing (US). (2014). *Standards for educational and psychological testing*. American Educational Research Association.
- Betebenner, D. & Van Iwaarden, A. (2020, June 2). Issues and considerations that the COVID-19 pandemic presents for measuring student growth [Blog post]. Retrieved from <https://www.ncea.org/blog/sgp/issues-and-considerations-covid-19-pandemic-presents-measuring-student-growth>
- Brookhart, S., & Nitko, A. (2019). *Educational assessment of students* (8th ed.). New York, NY: Pearson.
- Castelo, M. (2020, June 25). How K-12 schools are doing summer school virtually. *EdTech: Focus on K-12*. Retrieved from <https://edtechmagazine.com/k12/article/2020/06/how-k-12-schools-are-doing-summer-school-virtually-perfcon>
- Chalak, A., & Tavakoli, M. (2010). Sources of test score pollution: State of the art. *Curriculum Planning Knowledge & Research in Educational Sciences*, 26, 13–34.
- 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.
- Cooper, H., Nye, B., Charlton, K., Lindsay, J., & Greathouse, S. (1996). The effects of summer vacation on achievement test scores: A narrative and meta-analytic review. *Review of Educational Research*, 66(3), 227–268.
- DeWitt, P. (2020, May). Teachers work two hours less per day during COVID-19: 8 key EdWeek survey findings. *Education Week*. Retrieved from <https://www.edweek.org/ew/articles/2020/05/11/teachers-work-an-hour-less-per-day.html>
- Ebel, R. L., & Frisbie, D. A., (1991). *Essentials of educational measurement* (5th ed). Englewood Cliffs, NJ: Prentice Hall.
- Goertz, M. E., Oláh, L. N., & Riggan, M. (2009). *From testing to teaching: The use of interim assessments in classroom instruction*. Retrieved from https://repository.upenn.edu/cgi/viewcontent.cgi?article=1023&context=cpre_researchreports
- Haladyna, T. M., Nolen, S. B., & Haas, N. (1991). Raising standardized achievement test scores and the origins of test score pollution. *Educational Researcher*, 20(5), 2–7.
- Hammond, B. (2020, June 10). Risks, costs and goals of summer school in Oregon during coronavirus. *The Oregonian*. Retrieved from <https://www.opb.org/news/article/risks-costs-and-goals-summer-school-face-to-face-oregon-covid-coronavirus/>
- Herold, B., & Yettick, H. (2020, May). This is what teachers want us to now about pandemic earning. *Education Week*. Retrieved from <https://www.edweek.org/ew/articles/2020/05/11/teachers-work-an-hour-less-per-day.html>
- Kerry, T., & Davies, B. (1998). Summer learning loss: The evidence and a possible solution. *Support for Learning*, 13(3), 118–122.
- Maas, C. J., & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology*, 1(3), 86–92.
- Messick, S. (1984). The psychology of educational measurement. *Journal of Educational Measurement*, 21, 215–237.
- Olina, Z., & Sullivan, H. J. (2002). Effects of classroom evaluation strategies on student achievement and attitudes. *Educational Technology Research and Development*, 50(3), 61–75.
- Read, D. (2020, March 17). Embracing online teaching during the COVID-19 pandemic. *e Campus News*. Retrieved from <https://www.ecampusnews.com/2020/03/17/embracing-online-teaching-duringthe-covid-19-pandemic/>

Schwartz, S. (2020, May). Teacher appreciation week wrap-up: See how educators are innovating during COVID-19. *Education Week*. Retrieved from <https://www.edweek.org/ew/articles/2020/05/11/teachers-work-an-hour-less-per-day.html>

Tornay, K. (2020, June 25). Medford summer school expands after vulnerable students struggle with distance learning. *Mail Tribune*. Retrieved from [https://mailtribune.com/news/education/medford-](https://mailtribune.com/news/education/medford-summer-school-expands-after-vulnerable-students-struggle-with-distance-learning)

[summer-school-expands-after-vulnerable-students-struggle-with-distance-learning](https://mailtribune.com/news/education/medford-summer-school-expands-after-vulnerable-students-struggle-with-distance-learning)

Zughoul, O., Momani, F., Almasri, O., Zaidan, A., Zaidan, B., Alsalem, M., Albahri, O., Albahri, S. & Hashim, M. (2018). Comprehensive insights into the criteria of student performance in various educational domains. *IEEE Access*, 6, 73245–73264.