Re-Imagining K-12 Schools Student Assessment

CJ Greer, Joey Levin, and Steffenie Williams

Professor Anjalé Welton

TA Abby Koberstein

Spring 2021

Introduction

The American K-12 education environment has seen a tremendous amount of change over the past few decades, and the push to utilize standardized assessment methods has drastically increased. For example, standardized testing is one of the most commonly used and well known testing methods in the United States. When addressing this topic, American policymakers claim standardized testing allows for comparisons to be made amongst schools in regards to their student achievement, establishes accountability for teachers, and has the ability to inform instruction for educators. However, it is clear that standardized testing and other assessment methods in America do not even come close to accurately portraying the full picture of student educational outcomes. In fact, standardized assessment practices have enforced a strict system of segregation used to separate by intelligence, socioeconomic status, opportunity, and privilege. Thus, this policy brief will discuss the ways in which American K-12 students are being marginalized through the current standardized assessment methods. It will conclude with recommendations for alternative grading methods that provide authentically beneficial outcomes for students and their development.

Brief History of Standardized Assessment in America

The use of standardized assessment dates back to the mid nineteenth century. An academic named Horace Mann introduced the concept of using exams in Boston schools to gain objective information about the quality of teaching and learning in urban schools, monitor the quality of instruction, and compare schools and teachers within each school (Gallagher, 2003). Mann's exams were incredibly successful and were subsequently adopted by school systems across the country. Achievement tests, considered objective and comprehensive, became a popular method for assessing learning. Almost a century later, World War I helped kickstart renewed interest in standardized testing as the U.S. Army required a method for quickly identifying potential officers among large numbers of recruits. To do so, Arthur Otis and Robert Yerkes developed the Alpha Army Test, which gauged a soldier's mental capabilities (Edwards, 2008). The Alpha Army Test, which had an efficient and effective scoring method, became a model for many future standardized tests. This test changed the narrative around standardized testing and shortly after patent and copyright requests for tests soared (Edward, 2008).

Years later in 1965, the first set of federal laws were passed requiring the use of standardized tests (Edwards, 2008). With the Elementary and Secondary Education Act, the federal government aimed to eradicate poverty through the establishment of Title I funds and other grants directed to support educational resources and opportunities for low-income students and to help build the capacity of state educational support (Diem et al., 2021). However, through this act, the Johnson Administration administered standardized student assessment and influenced education with combative funding tactics that involved withholding federal money if their expectations were not met (Diem et al., 2021). This was not met with apathy; in 1966, the National Center for Education Statistics authorized research studies examining the issue of inequitable testing amongst diverse populations. This research, now referred to as the Coleman Report, found that a student's home environment was the most important factor affecting school achievement (Edwards, 2008). During this time period, supporters of standardized testing used the findings of the Coleman Report to assert students' home environments, and not standardized

testing biases, were responsible for the disparities in test scores amongst diverse populations (Edwards, 2008). As a response, the Johnson administration and many of his supporters rebutted this viewpoint and instead stated the system created these inequities and therefore needed to be reformed to make the American dream a reality for everyone (Diem et al., 2021).

Twenty years after ESEA was passed, the original efforts to eradicate poverty utilizing an equity-based approach to American education flipped entirely during the Reagan administration. In the now infamous 1983 report known as *A Nation at Risk,* Reagan and US Secretary of Education Terrel H. Bell wanted to study the nation's education problems and garner support for change through the usage of a presidential commission (Horsford et al., 2019). In short, the summary of this 18-month study could be synthesized into the following statement from Reagan's administration, "Downwardly spiraling pupil performance rendered the US education system dysfunctional, thereby threatening the nation's technological, military, and economic preeminence," (Horsford et al., 2019). Through this fear mongering, Reagan's Administration recommended adopting academic and performance standards, as well as state and local achievement tests that aligned with those standards. The focus was no longer about the needs of students, but instead on how the US fared in academic performance and how to marginalize historically underrepresented students in the process (Diem et al., 2021). By the end of the 1980s over 200 million standardized tests were administered annually to determine student academic readiness (Edwards, 2008). Fast forward to 2001, George W. Bush's Administration reauthorized the ESEA, called No Child Left Behind. Through this legislation, NCLB took a very different approach compared to the original intentions of ESEA. In the George W. Bush version, his administration used market-based one-size-fits-all instructional reforms that put enormous emphasis on standardized tests and punitive sanctions when standards and requirements were not met (Diem et al., 2021). This led to reinforced racial inequities and stereotypes that blamed students from low-income households and students of color for school failure, and protected the system that was supposed to serve them (Diem et al., 2021). In the two decades since NCLB was passed, the usage of standardized assessment has become even more prevalent in modern education and is now worth almost $2 billion per year. In fact, a 2015 investigation revealed companies such as Pearson, ETS, Houghton Mifflin, and McGraw-Hill spend approximately $20 million per year lobbying American politicians to promote pro-standardized assessment policies (DeMatthews, 2021).

Standardized Assessments

Standardized assessments have been a controversial tool used in educational settings since their introduction. The Iowa Test of Basic Skills created by Everett Franklin Lindquist, education professor at the University of Iowa, was used throughout the country as a way to test the 'skills' or 'achievements' of students in each grade. Although the names of standardized assessments have changed (STAAR, MAPP, WKCE etc.), the goals and problems with them remained. Standardized assessments are inequitable. Sometimes these assessments are linked to teacher pay, school evaluations and in some cases, even school closures. From a theoretical standpoint, "the model used to assess performance and competence in this society is monocultural (Ratteray, 1974; Mercer & Lewis, 1978; Savage & Adair, 1977). Its underlying assumption is that individuals must attain and demonstrate certain competencies deemed essential for effective functioning as members of society. A main criticism is that the model ignores the relevance of

culturally different experiences that foster other equally important competencies essential to the survival of the group or individual. A Black child, for example, may learn Black English instead of Standard American English (Labov, 1972). The home environment of the black child may nurture a socially interactive style that is in contrast to the individualistic and competitive style fostered in the classroom (Gordon, 1982). The black child may therefore be at a distinct disadvantage in an educational institution that utilizes the "favored" language and style" (Williams, p. 192-193). Languages, lived experiences, culture and more should be considered assets of our students, not deficits.

Standardized assessments create a deficit by default, giving advantage to the dominant culture while all other students are at a disadvantage. Negative test scores can be used against the student, placing them in 'low' classes, tracks or special education. "Standards-based reform, including testing, has gained increasing support as a strategy to improve schooling among legislators, educators and the general public (Hanushek & Raymond, 2005; McNeil, 2000; Neal & Schanzenbach, 2010). The Center for Public Education (2006) noted that a high stakes test has consequences attached to the results. The test results often determine whether a student will be promoted to the next grade, graduates from high school or whether a student is admitted to college" (Smith & Szymanski, p. 16-17). Students with experiences that are similar to those who create or write the questions for these assessments are at an advantage based on race, culture, socioeconomic status and lived experiences.

Another criticism of standardized assessments is creativity. How are students able to show that they can think creatively or critically when they are being asked to choose from a pre-ordained list of answers? They can't. This format of testing was created because they are quick to grade and inexpensive to give (Carroll, 2013). What about higher-level thinking? These assessments are graded by computers and there is no place for this kind of answer on a standardized assessment (Carroll, 2013). When educators and students spend hours and hours preparing for high stakes standardized testing little time is left for focusing higher order thinking and learning. As McComas and Abraham (2004) noted: Numerous studies reveal that educators know that higher-order questions - encompassing critical thinking, creativity, problem solving - engage learners and ensure the transfer of knowledge (learning). "Higher order thinking occurs when a person takes new information and information stored in memory and interrelates and/or rearranges and extends this information to achieve a purpose or find possible answers in perplexing situations" (Lewis & Smith, 1993, p.136; Smith & Szymanski, p. 17). When educators are required to teach to the test, they are forced to teach students using lower-order questioning. Educators must spend valuable time teaching students how to answer multiple choice questions administering practice tests - up to 113 in a student's lifetime (Kamenetz, 2015). Other questions to consider: Who writes the questions for these assessments? Who is deeming these questions and answers acceptable ways for students to prove their academic worth? Standardized assessments do not measure the change or growth of students from year to year. This year's test-taker's scores will be compared to last year's test-taker's scores. Comparing these scores show that the students in these classes were different, not that individual students, classes or schools are improving or failing. The Network for Public Education asserts that, "assessments should be used as diagnostic tools, to help teachers figure out where students are in their learning. But in most states, teachers are forbidden to see the actual test questions or provide feedback to students. Teachers do not see how their students answered specific test

items and learn nothing about how their students are doing, other than a single score, which may arrive long after the student has left their classrooms (Strauss, 2014). To be effectively used as educational tools teachers should be given access to the questions and scores for each standardized assessment. This would allow educators to use the information gathered to inform their practice, alter their lessons and curriculum by using the content diagnostically for each individual student. These assessments can be considered extremely high stakes for some students because Title I schools (schools with 35% of students coming from low-income families) lose federal funding when students do not meet basic proficiency requirements on standardized assessments. This means that schools in areas with the highest levels of poverty are expected to perform well on standardized assessments, despite the glaring inequalities they must overcome, because funding is linked to these students' scores. These schools can be closed, students displaced, educators relocated, and already struggling communities would be strained even further. If schools succeed in acquiring federal funding from Title I, which is to be used for programs that will 'close the gap,' they may find that the district chooses to disseminate the funds into other schools and programs throughout the district. It is required that the district uses the funds for certain content, but does not require that it all be used in the schools that need it the most. Instead of spending 1.7 billion dollars on standardized assessments (Chingos, 2012, p. 1), districts should be using that money to hire qualified, diverse, anti-racist educators.

Standardized assessments "were imposed to have some kind of system of equity and objective measures of how students were doing" (Kamenetz, 2015). Instead they are inequitable and marginalize students. Standardized assessments are used in ways that do not reflect the abilities of students of color, ELL, students with disabilities and students from low socioeconomic statuses. Time at school should be used for teaching and learning, not testing and test preparation. Standardized assessments do not inform instruction. These assessments do not measure important skills - creativity, critical thinking, cooperpation, collaboration, leadership and kindness. They are inequitable and unnecessary. Standardized assessments are great determiners of one thing - they determine if students are good at taking standardized assessments.

Standardized Grading

Similarly to standardized assessments, grading and reporting are pivotal elements of nearly every education system in the world. Grading represents teachers' evaluations, formative or summative, of students' performance. Reporting is how the results of those evaluations are communicated to students, parents, or others (Guskey et al., 2015). Because of their fundamental value, almost everything students do within the standardized system is given a letter grade to represent their performance. These results are then translated to a numerical system called a grade point average. A student's Grade Point Average (GPA) is how schools measure academic performance. A GPA is calculated by dividing the total amount of grade points by the total amount of classes attempted. Student GPAs may range from 0.0 to a 4.0. The overarching issue that has arisen is that this system is heavily flawed when it comes to accurately conveying the student's knowledge and abilities. For example, a student who received an A (Excellent) for one paper, C (Fair) for a second, and B (Good) for a third got ordinal grades, 4, 2, and 3. When the instructor succumbed to the arithmetic temptation, this student got a GPA of (4+2+3)/3 = 3.00. Imagine that instructors code the primary colors as Red = 4, Yellow = 3, and Blue = 2 and there

are three apples with these three different colors. Following the same rationale and equation, these apples have a cumulative GPA of 3.00, meaning that they are on average yellow (Soh, 2011). Even though all three apples are a different color, the current system in place makes it impossible to discern that information given just the GPA. This example highlights just how broken letter-based grading is, and how it does not provide the full picture of these students as individuals. Unfortunately, to judge the qualities of an individual student using all pertinent information is highly inconvenient from an administrative perspective. It is much simpler to use a single number to summarize each student's intellectual quality and then make important decisions based on that number (Soh, 2011).

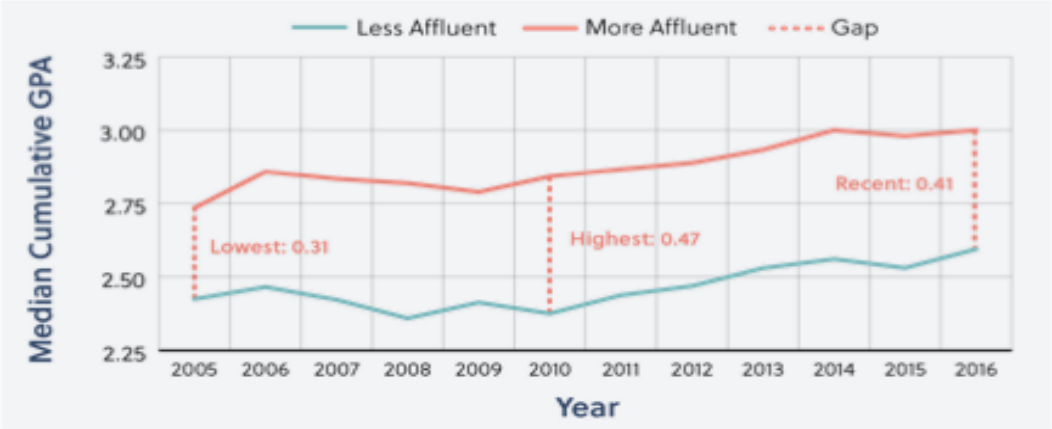
In addition to the fact that letter-based grading does not provide an entirely accurate assessment of student performance, it also has adverse effects on the students themselves. First, grades tend to diminish students' interest in whatever they're learning. In almost every study done that examined the impact of intrinsic motivation of receiving grades, or instructions that emphasize the importance of getting good grades, has found a negative effect on student learning (Kohn, 2013). Second, grades create a preference for the easiest possible task. When students have to decide what classes or coursework will count towards their grade, their response overwhelmingly avoids taking on any unnecessary intellectual risks. This means students will opt for shorter books and easier topics not because they are unmotivated, but rather they are rational (Kohn, 2013). They are responding to a system that tells them the goal is to get good grades and in return their academic success is more important than taking on intellectually challenging content that could lead to lower marks. Finally, grades tend to reduce the quality of students' thinking. This manifests differently for every student but common examples include: skimming books for "need to know" information, asking what is essential to know for the test, and being unable to respond to open ended questions where creative thinking is required (Kohn, 2013). Other noted findings of a grade-focused educational environment demonstrates there are increased rates of cheating, grades promote fear of failure even in high-achieving students, and the elimination of grades produces substantial benefits with no apparent disadvantages in higher levels of education (Kohn, 2013). Thus, the efficacy of standardized grading is highly questionable, unless these are the outcomes the current system intends to create.

The Consequences of Standardized Grading and Grade Inflation

Since grades have become such a pivotal component of modern education, it is not surprising that there has been an increase in grading inflation within the standardized system. Grade inflation is the awarding of higher grades than students deserve either to maintain a school's academic reputation or as a result of diminished teacher expectations. It occurs when the course grades subjectively assigned by teachers do not comport with objective measures of student performance. Researchers have been documenting the mismatch between school grades and external measures of student learning for decades. In fact, rising high school grade point averages (GPAs) have been accompanied by stagnant SAT, ACT, and NAEP scores, strongly suggesting lowered classroom standards. As for higher education, A's are now the most common grade awarded, despite constituting just 15 percent of grades in the early 1960s (Hurwitz et al., 2018). Grade inflation clouds measures of students' true knowledge and skills. Grade inflation can also exacerbate socioeconomic inequities in educational outcomes when it varies systematically by student or school background. On the one hand, teachers may be more apt to

inflate the grades of higher-performing, higher-income students in an effort to appease their pushy parents (Gerhenson, 2018). On the other, teachers may be more apt to inflate the grades of low-performing, low-income students to avoid both the difficulty and associated costs of remediating them. In a study done in the southeast region of the United States, researchers set out to examine the pervasiveness of grade inflation and how it correlates with other academic outcomes. Although this study was done in a specific part of the country, the researchers noted that this data can be nationalized with relative ease and should provide caution to educational administrators across the US. Their first finding demonstrated that while many students are awarded good grades, few earn top marks on their end of year exams (Gerhenson, 2018). Therefore, while students may find success in achieving a specific letter grade throughout the term, they struggle to show the levels of mastery on their end of year exams that are supposed to be affiliated with an "A" grade. The second finding from this study indicated during the years examined, grade inflation occurred faster in schools attended by more affluent students comparatively to schools attended by less affluent students. As demonstrated by the chart below, all GPAs have risen, but not in an equitable fashion (Gershenson, 2018).

Cumulative GPA Trends and Type of Schools



The final finding from this study is that cumulative student GPAs are increasing faster than ACT scores in affluent schools (Gershenson, 2018). As a result, students in more affluent schools systematically receive more optimistic evaluations of their current and future performance than their more disadvantaged peers, they will act on this misleading information (Papageorge et al. 2018). Thus, the implications of this section highlight the importance of critically analyzing grade inflation as a consequence of standardized grading and the role it plays in perceived student academic achievement. College admissions officers and future employers should have heightened awareness about these consequences triggered by standardized grading and in return place less emphasis on student GPAs.

COVID-19 K-12 Education Experience

COVID-19 has created incredible challenges for young people and educators traversing the K-12 education space in the past year (Dibner et al. 2020). We've seen their resiliency, humanity, and passion on full display throughout this pandemic. Each of these actors holds a unique perspective

in navigating their journey during this unprecedented experience. Engaging with the rigors of the preparation and standardized assessment adds to young people and educators' challenges. Standardized assessment during Covid-19 spotlights several shortcomings present before the pandemic such as educator/youth stress, lack of reliability in assessing the educational needs of a young person, addressing the educational debt present in our K-12 system[1], and the dearth of uniformed assessment for youth across states (DeMatthews, 2021; Johnson, 2021; Ujifusa, 2020; Starr, 2020). Each of these problems uniquely places additional stress on young people, educators, and state governance navigating the standardized testing during this timeframe. Additionally, several states voiced the position of pushing back against testing as the data gathered during this time is an inaccurate depiction of students' abilities with the harsh realities of Covid (Toppo, 2021; Johnson, 2021; Ujifusa, 2020); Starr, 2020. As we grapple with this new reality, we must approach this issue holistically. Knowledge about the pandemic's impact on stressed educators and youth outputs in schools is sparse (DeMatthews, 2021). It's critical to take time to understand how COVID-19 impacted young people and educators throughout the K-12 education system. A lack of acknowledgment around their lived experiences is a disservice to their humanity.

Young people's resilience comes to light through numerous aspects of their schooling experience. The literature highlights their viewpoint in navigating an educational world without physical group learning. School-centric spaces provided a buffer for numerous societal ills faced by students. The pause of in-person services impacts the social development, mental health support, and schooling experience youth need (Dibner et al. 2020; Powers et al., 2020; Antoni, 2020). Another issue that intensified through the pandemic is student insecurity. Student's insecurities around food, parent employment (financial stability), and internet access have all influenced the academic experience of young people during this virtual era (Kaden, 2020; Simpson, 2020). Internet usage specifically poses distinct complications for youth in engaging with school. For some, finding suitable spaces for virtual schooling requires leaving their homes to enter restaurant and coffee spaces (Strauss, 2020; Simpson). This access problem is only the beginning of the challenge for many students. Young people voiced their displeasure in virtual schooling around the length of videos, large groups with multiple speakers, screen time, the lack of structural consistency between classes, and lack of instructional support opportunities (Black et al., 2020; Simpson, 2020). This insight illustrates the online tension our youth navigate during this era of Covid-19. The abrupt lack of child supervision stemming from school closure thrust some older siblings into the additional role of caregiver for siblings during the day (Antoni, 2020; Kaden, 2020). Balancing caregiving and learning is a steep responsibility for any individual. The intersection of these various tensions highlights the determination so many students have in striving for their education in our current context. Adding traditional standardized assessment does not authentically evaluate young people during this time.

The surrounding literature on K-12 educators' experiences in COVID-19's presents several conversations around the perception of teaching, the shift in the teaching profession, and the response to the pandemic. Educators' dynamic and multifaceted roles became on full display throughout the pandemic (Kaden, 2020). Parent's awareness of educators' challenging roles in developing young people became more apparent (Garbe et al., 2020a; Garbe et al., 2020b). This

[1] The term “Educational Debt” makes reference to Gloria Ladson-Billings reframing of the term “Achievement Gap” into an asset driven assessment (Ladson-Billings, 2006).

illumination around educators assists in showing the "unnoticed work" for students often involved in the teaching profession (Kokkinos, 2007; Pogere et al., 2019). Unnoticed work by educators crafts the foundation throughout good pedagogy and youth development for K-12 students. In searching for equitable responses that championed good pedagogy and youth development, educators found that shifting to an online-based teaching platform with a "traditional classroom instruction" skillset didn't translate well (Toth, 2020). Educators' realization around the necessity of shifting their practice manifested itself in the form of enhanced workloads (in and outside the classroom), intentionally engaging with students more frequently (calling home, delivering supplies, etc.), and acquiring new teaching skills necessary for a virtual learning space (Kaden, 2020; Sullivan et al., 2020; Lepp et al., 2021). The teaching profession's response to these numerous issues stemming from COVID-19 largely shaped the experiences of students and parents alike (Grooms & Childs, 2021). K-12 educators' commitment to providing learning for young people proved foundational in navigating the pandemic. Adding the additional stress of standardized testing preparation further burdens the teaching profession that many educators do not necessarily need during these trying times. The mission of standardized testing becomes arbitrary when viewing the experiences of educators and youth in addition to the historically problematic nature of standardized testing.

Overview of Standardized Assessment in Wisconsin

Students in Wisconsin are required to take **46** standardized assessments during their K-12 schooling. This does not include the assessments given by classroom teachers in various subjects.

- The Wisconsin Forward Exam in grades 3-8 in English Language Arts (ELA) and Mathematics, at grades 4 and 8 in Science, and 4, 8, and 10 in Social Studies. There are 17 total assessments and the estimated time for test administration in each grade is approximately:
    - 2 hours and 20 minutes for ELA (English Language Arts)
    - 1 hour and 45 minutes for Mathematics
    - 1 hour and 40 minutes for Science
    - 1 hour and 30 minutes for Social Studies.
- Dynamic Learning Maps (DLM or Maps) at grades 3-11 in ELA and Mathematics, at grades 4 and 8-11 in Science, and at grades 4, 8, and 10 in Social Studies. There are a total of 26 assessments. The DLM Assessment is not a timed test. Each student is administered the test on an individual basis. A student may take the test over several days, as long as it is completed within the testing window. The total time required to administer the test varies, however, it is estimated the entire test will take about 2.5 hours to complete.
- ACT Aspire for grades 9 and 10. There are two assessments with five sections and the estimated time for test administration in each grade is approximately:
    - 45 minutes for English
    - 75 minutes for Math
    - 65 minutes for Reading
    - 60 minutes for Science
    - 40 minutes for Writing

- ACT with Writing for grade 11. There is one assessment with five sections and the estimated time for test administration is approximately:
    - 45 minutes for English
    - 60 minutes for Math
    - 35 minutes for Reading
    - 35 minutes for Science
    - 40 minutes for Writing

Best Practices & Recommendations

Holding onto the lessons learned throughout the COVID-19 pandemic is critical as we begin transitioning into a new K-12 learning space by reintroducing in-person learning and assessment. The tension between virtual and in-person learning revealed that our education system upholds inequality (Kaden, 2020; Black et al., 2021). Educators, parents, young people, and policymakers should use this time to reimagine the potential of school-centric spaces (Poletti, 2020; Richmond et al., 2020). The equity issues present in schools before the pandemic can be eradicated by working collaboratively with one another, centering the needs of all youth people, and intentionally addressing inequity. Below we've highlighted two feasible policy recommendations for addressing the equity issue found in standardized testing assessments for K-12 students. These practices allow for the overarching goals of creating a more equitable and holistic evaluation of young people and combating score-driven teaching practices in the state of Wisconsin.

Recommendation I: Reexamining the necessity of standardized testing for Primary Wisconsin students

Throughout the pandemic, a perspective shift on the role of standardized testing has happened in the realm of higher education. A reimagining around utilizing standardized testing for determining the trajectory of post-secondary students is taking place currently in our U.S. context (Vigdor & Diaz, 2020; Moody, 2020). The basis for eliminating these assessment measures centers on the historical evidence of racial bias, class bias, stress creation (during a pandemic), divergence from curricula goals, and the lack of actual assessment on content knowledge (Bhattacharyya et al., 2013; Middleton, 2020; Kaukab & Mehrunnisa, 2016). We argue that this conversation in the post-secondary section of our education system is transferable to the K-12 education sector. Our first practice recommendation is reexamining the necessity of standardized testing assessments for WI K-12 students. Primary and secondary students face the same challenges utilized by higher education spaces in turning away from admissions-based testing. A shift away from our standardized forms of evaluation would begin deconstructing the competition-based strategies that further disfranchise vulnerable students throughout the K-12 system. We fail our youth when we incentivize standardized testing instead of striving for their development as young people through schooling (Bhattacharyya et al., 2013). We must begin to have discourse around this issue. Additionally, we have shown the ability as a country to redetermine the parameters for education measures (which align with our federal funding structure). For instance, the Every Student Succeeds Act (ESSA) reauthorization gave authority

to states in determining their own goals for determining student success (U.S. Department of Education, 2015). The shift in federal oversight opens the possibility for exploring different funding metrics for schools while investigating holistic assessments for students.

Recommendation II: Removing Letter-Based/Standard-Based Assessment and Implementing Skills-based Assessment for Wisconsin Primary students

As previously outlined, standardized testing assessments do not perform the intention of their role (Bhattacharyya et al., 2013; Middleton, 2020; Kaukab & Mehrunnisa, 2016). Creating a conversation around these well-documented issues is only the beginning of addressing this issue. Our second recommendation focuses on tangible actions/alternatives for standardized assessment following the discourse around the subject. We argue that removing letter grade assessments for primary students opens the door for a skills-based assessment that's more holistic in evaluating a child's academic profile, trajectory, and needs. The standards-based assessment[2] found in testing is predicated only on students reaching certain "thresholds" or "distinction" (Hendry et al., 2012). Alternatively, skills-based assessment emphasizes qualitatively assessing students instead of focusing just on the score they produce. Removing letter grades is not a utopian fantasy, primary schools all across the country have successfully implemented different grading methods proven to be more beneficial for their students by enhancing young peoples' ability to think creatively and increases their level of engagement with academic material (Kohn, 2013). Some illustrations of skill-based assessments include student self-assessments (rubrics, checklists, self-reflections), teacher observations of students and their work (teachers documenting classroom work ethnographically), interviewing students on reading and math, skill portfolios and public defenses of student work, multiple measures of skills assessment, and external reviews by education consultants (Meier & Knoester, 2017). These assessments reduce anxiety (for students and educators) and the happiness associated with these practices are great "indicators of success, linked to grades, retention, and employment" (Kamenetz, 2015). An example of the utilization of skills-based assessment instead of standard-based comes from New York state. New York Performance Standards Consortium (NYPSC)[3] claims to utilize skills-based assessment in helping students grow in areas of critical thinking, problem-solving, and public speaking while providing resources to help young people, parents, and educators shift away from standard testing-based schooling (Meier & Knoester, 2017). Exploring these standards alternatives is well within the state and local government's purview (and interests) as the average US state spends over $1.7 billion on standardized assessment that could shift towards other education needs (Ujifusa, 2012). Additionally, the shifting away from standardized based assessment for matriculation at the higher education level mitigates fears individuals may hold in moving away from grade-based/standardized-based assessment (Kohn, 2013).

[2] Standards are written descriptions of the quality of work expected at different levels. These descriptions are usually constructed with key criteria in mind (Hendry et al., 2012).

[3] New York Performance Standards Consortium (NYPSC) is a group of 36 high schools in New York City, Rochester, and Ithaca (Meier & Knoester, 2017).

References

Anderson, K. J., & Minke, K. M. (2007). Parent involvement in education: Toward an understanding of parents' decision making. The Journal of educational research, 100(5), 311-323. https://doi.org/10.3200/JOER.100.5.311-323

Antoni, J. (2020). Disengaged and Nearing Departure: Students at Risk for Dropping Out in the Age of COVID-19. Faculty/Researcher Works. http://dx.doi.org/10.34944/dspace/396

Bhattacharyya, S., Junot, M., & Clark, H. (2013). Can you hear us? Voices raised against standardized testing by novice teachers. *Creative Education*, *4*(10), 633. http://dx.doi.org/10.4236/ce.2013.410091

Black, E., Ferdig, R., & Thompson, L. A. (2021). K-12 virtual schooling, COVID-19, and student success. JAMA pediatrics, 175(2), 119-120. doi:10.1001/jamapediatrics.2020.3800

Carroll , J. (2013). The Brutal Reality of Bringing Kids up to Level: Are Critical Thinking and Creativity Lost in the World of Standardized Testing? *University of New Orleans , Educational Sociology Commons*. https://scholarworks.uno.edu/cgi/viewcontent.cgi?referer=http://scholar.google.com/&httpsredir=1&article=2688&context=td

Chingos, M. (2012). *Strength in Numbers: State Spending on K-12 Assessment Systems (Report)*. Brown Center on Education Policies at Brookings. https://www.brookings.edu/wp-content/uploads/2016/06/11_assessment_chingos_final_new.pdf

Dibner, K. A., Schweingruber, H. A., & Christakis, D. A. (2020). Reopening K-12 schools during the COVID-19 pandemic: A report from the National Academies of Sciences, Engineering, and Medicine. Jama, 324(9), 833-834.

Diem, S. & Welton, A. (2021). Anti-Racist Educational Leadership and Policy: Addressing Racism in Public Education. *Routledge, Taylor & Francis Group.*

DeMatthews, D. (2021) Standardized Testing Amid Pandemic Does Kids More Harm Than Good. *UT News.* https://news.utexas.edu/2021/03/15/standardized-testing-amid-pandemic-does-kids-more-harm-than-good/

Edwards, N. (2008). The historical and social foundations of standardized testing: In search of a balance between learning and evaluation." *JALT Testing & Evaluation SIG Newsletter*. https://hosted.jalt.org/test/edw_1.htm

Feldman, J. (2018). Inequity in Grading. *National Association of Secondary School Principals*. www.nassp.org/publication/principal-leadership/volume-19-2018-2019/principal-leadership-december-2018/inequity-in-grading/

Fontanesi, L., Marchetti, D., Mazza, C., Di Giandomenico, S., Roma, P., & Verrocchio, M. C. (2020). The effect of the COVID-19 lockdown on parents: A call to adopt urgent measures. Psychological Trauma: Theory, Research, Practice, and Policy. doi:10.1001/jama.2020.14745

Gallagher, C. J. (2003). Reconciling a tradition of testing with a new learning paradigm. Educational Psychology Review, 15 (1), 83-99.

Garbe, A., Ogurlu, U., Logan, N., & Cook, P. (2020a). Parents' experiences with remote education during COVID-19 school closures. American Journal of Qualitative Research, 4(3), 45-65. https://doi.org/10.29333/ajqr/8471

Garbe, A., Ogurlu, U., Logan, N., & Cook, P. (2020b). COVID-19 and remote learning: Experiences of parents with children during the pandemic. American Journal of Qualitative Research, 4(3), 45-65.https://doi.org/10.29333/ajqr/8471

García, E. (2020). The pandemic sparked more appreciation for teachers, but will it give them a voice in education and their working conditions. Erişim Adresi: https://www. epi. org/blog/the-pandemic-sparked-more-appreciation-for-teachers-but-will-it-give-them-a-voice-in-education-and-their-working-conditions.

Gershenson, S. (2018). Grade Inflation in High Schools (2005-2016). *Thomas B. Fordham Institute*. https://eric.ed.gov/?id=ED598893

Gordon, B. (1982). Towards a theory of knowledge acquisition. *Journal of Education*. 164, 90-105. Retrieved from: https://www.jstor.org/stable/42772833?casa_token=VkzKq-1l6SMAAAAA%3Av7hW_4Fw2GwL4EE8b8eTjn27ld7I2D-_H1FQL5GB-_wv3GZ5uJeBs4G422kAPTer-VvggulelPetU3I-a9jaRE18fIv62yrI-kyPp6bK5DoOKxYhaxg&seq=1#metadata_info_tab_contents

Guskey, T. & Munoz, M. (2015). Standards-Based Grading and Reporting Will Improve Education. *SAGE Journals.* https://doi.org/10.1177/0031721715579043

Hendry, G., Armstrong, S., & Bromberger, N. (2012). Implementing standards-based assessment effectively: Incorporating discussion of exemplars into classroom teaching. *Assessment & Evaluation in Higher Education*, *37*(2), 149-161. https://doi.org/10.1080/02602938.2010.515014

Horsford, S., Scott, J., & Anderson, G. (2019). The Politics of Education Policy in an Era of Inequality Possibilities for Democratic Schooling. *Routledge*.

Hurwitz, M. & Lee, J. (2018). Grade Inflation and the Role of Standardized Testing. *Measuring Success: Testing, Grades, and the Future of College Admissions, Johns Hopkins University Press,* 64.

Johnson, S. (2021). Schools must give standardized tests this year, Biden administration says. EdSource. Retrieved from https://edsource.org/2021/schools-must-give-standardized-tests-this-spring/649751

Kaden, U. (2020). COVID-19 school closure-related changes to the professional life of a K–12 teacher. Education Sciences, 10(6), 165. https://doi.org/10.3390/educsci10060165

Kamenetz, A., Sreenivasan, H. (2015, Jan. 22). Is 'The Test' failing American Schools? PBS News Hour. Retrieved from: https://www.pbs.org/newshour/show/test-failing-american-schools

Kaukab, R., & Mehrunnisa, S. (2016). History and evolution of standardized testing–a literature review. *International Journal of Research-Granthaalayah*, *4*(5), 126-132. https://doi.org/10.29121/granthaalayah.v4.i5.2016.2688

Kokkinos, C. M. (2007). Job stressors, personality and burnout in primary school teachers. British journal of educational psychology, 77(1), 229-243. https://doi.org/10.1348/000709905X90344

Kohn, A. (2013). The Case Against Grades. *Counterpoints*, 451, 143-153. http://www.jstor.org/stable/42982088

Labov, W. (1972). *Language in the inner city. Philadelphia*. The University of Pennsylvania Press.

Ladson-Billings, G. (2006). From the achievement gap to the education debt: Understanding achievement in US schools. *Educational researcher*, *35*(7), 3-12. https://doi.org/10.3102/0013189X035007003

Lepp, L., Aaviku, T., Leijen, Ä., Pedaste, M., & Saks, K. (2021). Teaching during COVID-19: The Decisions Made in Teaching. Education Sciences, 11(2), 47. https://doi.org/10.3390/educsci11020047

Levin, H. M. in Seymour, M. (2004). *Educating for humanity: Rethinking the purposes of education*. London: Paradigm Publishers.

Lewis, A., & Smith, D.C. (1993). Defining higher order thinking. *Theory Into Practice: Teaching for Higher Order Thinking*. 32(3), 131-137.

Meier, D., & Knoester, M. (2017). *Beyond testing: Seven assessments of students and schools more effective than standardized tests*. Teachers College Press.

Mercer, J. & Lewis, J. (1978). *SOMPA: System of multicultural pluralistic assess- ment. Conceptual and technical manual*. Riverside, California: Institute for Pluralistic Assessment Research and Training.

Middleton, K. V. (2020). The Longer‐Term Impact of COVID‐19 on K–12 Student Learning and Assessment. *Educational Measurement: Issues and Practice*, *39*(3), 41-44. https://doi.org/10.1111/emip.12368

Moody, J. (2020). Navigating Test-Optional Admission Amid Covid-19. US News & World Report: https://www.usnews.com/education/best-colleges/articles/how-the-coronavirus-is-pushing-colleges-to-go-test-optional

New York Performance Standards Consortium (Alternative to traditional Standardized Assessments) - http://www.performanceassessment.org/

Papageorge, N. & Gershenson, S. (2018) The Power of Teacher Expectations: How Racial Bias Hinders Student Attainment. *Education Next 18*, 65–70, http://educationnext.org/power-of-teacher-expectations-racial-bias-hinders-student-attainment/

Pogere, E. F., López-Sangil, M. C., García-Señorán, M. M., & González, A. (2019). Teachers' job stressors and coping strategies: Their structural relationships with emotional exhaustion and autonomy support. Teaching and Teacher Education, 85, 269-280. https://doi.org/10.1016/j.tate.2019.07.001

Poletti, M. (2020). Hey teachers! Do not leave them kids alone! Envisioning schools during and after the coronavirus (COVID-19) pandemic. Trends in Neuroscience and Education. 10.1016/j.tine.2020.100140

Powers, J. M., Brown, M., & Wyatt, L. G. (2020). SPARK-ing innovation: a model for elementary classrooms as COVID-19 unfolds. Journal of Professional Capital and Community. https://doi.org/10.1108/JPCC-06-2020-0036

Ratteray, J. (1974). *The testing of cultural groups. A paradigmatic analysis of the literature on testing and a proposition*. Santa Monica, California: Rand Corporation.

Richmond, G., Bartell, T., Cho, C., Gallagher, A., He, Y., Petchauer, E., & Curiel, L. C. (2020). Home/school: Research imperatives, learning settings, and the COVID-19 pandemic. https://doi.org/10.1023/A:1021323509290

Savage, F. & Adair, A. (1977) Testing minorities: Developing more culturally relevant assessment systems. The Negro Educational Review. 28, 219-228. Retrieved from: https://files.eric.ed.gov/fulltext/ED574461.pdf

Scott, T. (2004). Teaching the ideology of assessment. Radical Teacher, 71 (4), 30-37.

Simpson, J. C. Distance learning during the early stages of the COVID-19 pandemic: Examining K-12 students' and parents' experiences and perspectives. http://www.mifav.uniroma2.it/inevent/events/idea2010/doc/46_2.pdf

Smith, V., Szymanski, A. (2013). Critical Thinking: More than Test Scores. *International Journal of Educational Leadership Preparation*. 8(2), 16-25. Retrieved from: https://files.eric.ed.gov/fulltext/EJ1016160.pdf

Soh, K. (2011). Grade Point Average: What's Wrong and What's the Alternative? *Journal of Higher Education Policy and Management.* https://doi.org/10.1080/1360080X.2011.537009

Starr, J. (2020). Pandemic Offers Opportunity to Reduce Standardized Testing. EducationNext. Retrieved from https://www.educationnext.org/pandemic-offers-opportunity-reduce-standardized-testing/

Strauss, V. (2020). School in a coffee shop? A different approach to teaching and learning during the pandemic. Washington Post: https://www.washingtonpost.com/education/2020/07/17/school-coffee-shop-different-approach-teaching-learning-during-pandemic/.

Sullivan, F., Hillaire, G., Larke, L., & Reich, J. (2020). Using teacher moments during the COVID-19 pivot. Journal of technology and teacher education, 28(2), 303-313. Retrieved from https://www.learntechlib.org/primary/p/216171/.

Toppo, G. (2021). States Grapple with Standardized Testing during Pandemic. *EducationNext*. Retrieved from https://www.educationnext.org/states-grapple-standardized-testing-during-pandemic/

Toth, M. D. (2020). Achievement Gaps and the Lost COVID-19 Generation. https://www.learningsciences.com/wp-content/uploads/2020/09/LSI50-09-Achievement-Gaps-COVID.pdf

Ujifusa, A., (2012). Standardized Testing Costs States $1.7 Billion a Year, Study Says. *Education Week*. Retrieved from https://www.edweek.org/teaching-learning/standardized-testing-costs-states-1-7-billion-a-year-study-says/2012/11?tkn=VLMFJUQpeyvKkTzwuCHPd%2FuQG%2BPWLRrD1lNp&cmp=clp-edweek

Ujifusa, A. (2020). States Push to Ditch or Downplay Standardized Tests During Virus Surge. EducationWeek. Retrieved from https://www.edweek.org/teaching-learning/states-push-to-ditch-or-downplay-standardized-tests-during-virus-surge/2020/11

U.S. Department of Education (2015). Every Student Succeeds Act, Washington, D.C. Retrieved from https://www.ed.gov/essa?src=ft

Vigdor, N., & Diaz, J. (2020). More colleges are waiving SAT and ACT requirements. *The New York Times*: https://www.nytimes.com/article/sat-act-test-optional-colleges-coronavirus.html