High Stakes Testing in the 21st Century: Implications for Students in Special Education

Lola Gordon, Ed.S.

Abstract

High-stakes testing has been a part of American education since its inception. The laws that govern the use of high-stakes tests include language that mandates the inclusion of students in special education. These laws play an influential role in the new large-scale assessments aligned with the Common Core State Standards (CCSS). The assessments being implemented in the current 2014-2015 school year include embedded and locally provided accommodations for students with a documented need. For students with the most severe cognitive disabilities alternate assessments are available aligned with the CCSS. The implications of these assessments and the role they play in crucial factors pertaining to students in special education requires additional research. Specific areas for research should include how high-stakes tests can be useful in identifying specific needs, accommodations, and strategies for learning for students in special education, and the role test scores play in retention and dropout rates for this population.

High Stakes Testing in the 21st Century: Implications for Students in Special Education

Throughout history standardized tests have been used to gather data on student achievement. Results from standardized tests are used to measure individual performance of students, schools, and districts, as well as to compare across students, schools, and districts (Salvia, Ysseldyke, & Bolt, 2010). Until recently many students with disabilities were excluded from these tests, or were required to participate inappropriately. Over the past several decades, however, mandates have been implemented which govern how students in special education participate in high-stakes standardized testing, with a focus on appropriate and unbiased participation. The development of new large-scale assessments aligned with the Common Core State Standards (CCSS) addresses issues pertinent to students in special education. Considerations for the use of accommodations and alternate assessments utilizing the principles of universal design have been made, aimed at providing equal access to appropriate assessment opportunities for all students in special education. As students with disabilities begin participating in these new assessments, data gathered about student performance should provide information necessary in identifying the specific needs of students in special education, particularly in relation to the use of appropriate accommodations. Also, examining the role that high-stakes test scores play in retention and dropout rates for students in special education, and whether or not those rates are affected by the new assessments will be important.

Background

The testing of students has a considerable history in American education. Beginning in the early 20th century standardized tests began being used to measure student achievement in basic school subjects. In 1965 the Elementary and Secondary Education Act (ESEA) was created which included a requirement for testing for accountability purposes. Connections between declining

test scores in literacy and math and joblessness in the 1970's led to major reforms in the 1980's. The 1983 report *A Nation at Risk* painted a bleak picture of student performance on academic skills. A new emphasis was on the high-stakes attached to tests, even though the data at the time suggested that little improvement would result from high-stakes alone (Kaestle, 2013). The development of content and performance-based standards with aligned assessments came about in the 1990's, which led into the 21st century, and in 2001, the reauthorization of the ESEA, titled No Child Left Behind (NCLB) placed a high level of importance on frequent administration of high-stakes tests for the purpose of accountability. With rewards and sanctions based on student performance, NCLB mandated that all students, including students in special education, be included in the accountability requirements (Kaestle, 2013).

Prior to NCLB, federal provisions concerning accountability practices were limited in their ability to regulate the assessment of students in special education. NCLB required that most students in special education participate in high-stakes testing aligned with grade-level standards and those with significant cognitive disabilities take an alternate assessment that would be included in the accountability requirement (Salend, 2008). In 2009 the American Reinvestment and Recovery Act was implemented, which included a grant program titled Race to the Top that provided funding to states that adopted common academic standards and assessments aligned with those standards. These common academic standards came for most states in the form of the Common Core State Standards (CCSS) and two aligned assessments were developed by the Smarter Balanced Assessment Consortia (SBAC) and the Partnership for Assessment of Readiness for College and Career (PARCC). In addition, there are two consortia that designed alternate assessments, Dynamic Learning Maps (DLM), and the National Center and State Collaborative (NCSC) for students with severe cognitive disabilities. Test-based accountability has continued to play a vital role in education, and it is more persistent than any other education policy, despite continuing dissatisfaction with student performance (Linn, 2013).

Statement of Problem

Test-based accountability contains the assumption that the tests used are accurate measures of student achievement and can be used as predictors of future success in college and careers. Underperformance on high-stakes assessments has been seen through several factors, including performance on international assessments, the increasing numbers of students needing to take remedial, non-credit-bearing coursework in college, complaints from employers about a lack of preparedness of high school graduates, and the continued prevalence of students dropping out of school (Linn, 2013). Students in special education typically require accommodations to participate appropriately, and the new assessments were designed with consideration of accommodation needs. Whether or not they will be accurate in identifying appropriate accommodations and learning strategies will require more research. In addition, with concern over the dropout rates of students in special education, the question of what role the new assessments will play in changes to those rates will need to be considered.

In the area of needs, accommodations, and strategies, the removal of obstacles that have prevented accurate evaluation of the performance of students in special education was a necessary step in improving testing practices (SBAC.org). When administered appropriately accommodations help to ameliorate the effects of the individual characteristics of students in special education that limit their ability to demonstrate achievement (Geller, Alonzo, Monegan,

& Tindal, 2007). Throughout the history of high-stakes testing, questions regarding what types of accommodations are appropriate, who is qualified to make decisions about accommodations, and how classroom-based accommodations influence the performance of students in special education on high-stakes tests have persisted (Geller, et.al., 2007). Concerns over non-standard administration of assessments threatening validity has prompted more attention being given to the appropriate use of accommodations and in determining their capacity for providing the types of information needed to guide instruction (Fuchs, Fuchs, & Capizzi, 2005). Attention must also be given to the small percentage of students who require alternate assessments to be appropriately and accurately evaluated. Regulations now mandate that students who cannot participate in the regular assessments be offered alternate assessments, however, only 1% of proficient or advanced scores from the alternate assessments are allowed to be counted toward accountability ratings (Slocum, 2005). If the alternate assessments that have been developed prove to be effective in identifying appropriate interventions for students with severe cognitive disabilities the possibility exists that a larger percentage of students taking alternate assessments will score at higher levels, which may lead to a needed change in the 1% rule.

The correlation between scores on high-stakes tests and retention and dropout rates in high school has been documented. The question to be considered is what type of impact the new assessments will have on these rates. Grade promotion tied to performance on high-stakes tests has become more common over the last decade, and the increase in dropout rates due to the emphasis on these types of tests is an unintended consequence of testing. Increasing frustration with school and a lowered academic self-concept tied to scores on high-stakes assessments is a factor in dropout rates (Allensworth, 2005). Most states now require all high school students to pass a general skills exam in order to graduate. There is question as to whether or not the new assessments aligned with the CCSS will serve as exit exams, and for students in special education the question is of particular importance. The wide graduation-rate gaps in many states between students with disabilities and those in regular education is well documented. The most recent U.S. Department of Education data, for 2011-12, shows a four-year graduation-rate gap that ranges from a high of 43 percentage points in Mississippi to a low of 3 percentage points in Montana (Samuels, 2014). Students in special education who experience failure or see little chance of passing assessments and graduating often decide to drop out (Thurlow, Sinclair, & Johnson, 2002). Currently, research does not provide definitive answers to these concerns - regarding whether the new high-stakes assessments will result in growing numbers of students in special education dropping out.

The purpose of high-stakes testing is to improve outcomes for all students by improving ongoing progress monitoring and instruction, but the assessment outcomes need to correspond with real improvements in student achievement (Slocum, 2005). The new wave of computer-based assessments is just beginning to be implemented. The information they will be able to provide for students in special education and the effect they will have on dropout rates will be seen in the coming years as the assessments are evaluated for validity, reliability, and effectiveness (Stephens, 2014).

For students in special education the issue of participation in high-stakes testing has been contentious. The Individuals with Disabilities Education Act (IDEA) and NCLB laid the foundation for the accountability of students in special education by requiring states to include

this population in all state and district assessments and to report their participation and performance (US DOE, 2012). These mandates led to a wider use of accommodations and alternate assessments to allow students in special education to more appropriately participate (Linn, 2013). IDEA 2004, heavily aligned with NCLB, requires that the Individualized Education Plan (IEP) detail specific accommodations students would require for testing, and states that students in special education are no longer exempt from state and district assessments (Katsiyannis, Zhang, Joseph, & Jones, 2007). Unintended consequences of these efforts included an overabundance of test preparation that led to a narrowing of the curriculum, and in the years since NCLB a call for more rigorous standards and expectations along with aligned high-stakes testing (Linn, 2013). In response to these and other failures of the NCLB, the Common Core State Standards (CCSS) were developed, and implementation of the standards began in 2010.

The CCSS are intended to be rigorous in promoting higher-order thinking and problem-solving skills, and students who master the CCSS in high school are thought to be well prepared for college and career (Linn, 2013). Students in special education are held to the same content and grade-level standards and are taught with an emphasis on the use of research-based instructional practices. Funding for the two assessment consortia, SBAC and PARCC, was provided from the federal Race to the Top initiative to develop assessments aligned with the CCSS. Initially 44 states signed on to adopt the standards; currently, 23 of the 44 states have sought legislation to repeal, delay, or withdraw from their testing consortia (Stephens, 2014). While states initially felt the idea of common academic standards and standardized assessments directly aligned with those standards was a positive proposition, the states that have withdrawn or are seeking withdrawal cite concerns over the assessments being rushed into implementation without a strong research base supporting their effectiveness. And challenges still exist for students in special education that are not addressed with the implementation of the CCSS or the aligned assessments. Many students in special education require roughly 30-40 more days of instruction to have an equitable opportunity to learn than their peers, and with the technology associated with the new assessments there may not be enough time for students to be taught the standards they will be assessed on (Nirvi, 2012). An additional issue relates to the concern that even after over a decade of comprehensive education reform policies whose purpose was to hold everyone accountable for student performance and outcomes, there are still unacceptably high rates of students in special education dropping out of school (McNeil, Coppola, Radigan, & Heilig, 2008).

The shift in accountability practices since NCLB has raised different perspectives on whether the use of high-stakes assessments for all students regardless of ability is appropriate. Proponents of the CCSS and the next-generation assessments focus on aspects such as the benefit of more continuity among states, that agreed-upon standards will lead to better outcomes for students with increased preparedness for college and careers, and that the assessments were designed utilizing the principles of universal design. Universal Design is a way to provide assessments with tools built into the system, minimizing the need for other accommodations. They also argue that the participation of all students ensures that school districts are responsible for everyone, not just the highest achievers (Salvia, Ysseldyke, & Bolt, 2010). Prior to provisions in NCLB and IDEA the primary source of accountability for students in special education was the IEP, but IEP goals were often not standards-based, making them less reliable and valid as measures of

achievement (McLaughlin & Thurlow, 2010). However, the mandate for all students to be tested according to the same achievement standards is controversial. Opponents of the current reforms cite concerns over too much class time being spent on practice tests, and that assessment policy will drive instructional practices, which has been detrimental for all student learning. There is also concern over educators diverting resources away from students whose scores may not be counted, such as those with severe cognitive disabilities (Salvia, et.al, 2010). Concerns also exist over the achievement gap broadening for those students who are not tech-savvy. A question for the future will be in determining the reliability of the inferences made from high-stakes assessments.

Review of the Literature

The issue of how the results from high-stakes testing can be used to make meaningful decisions for students in special education is well represented in the literature. The question of the correlation between high-stakes tests and dropout rates is also represented in the literature, but both issues require a much larger research base. Research findings about the effects of standardized test-based accountability have been both promising and disappointing. The practices of extensive test-preparation and frequent interim testing as a result of accountability pressures often do not serve the population of students in special education effectively (Ed. Policy, 2009).

The new assessments, including the alternate assessments, have been designed using principles of universal design. Proponents of the new assessments cite components such as universal tools, designated supports, and digitally embedded and locally provided accommodations as being effective tools for not only making the assessments meaningful for all students, but also in providing the necessary information regarding the supports students require for learning (SBAC.org). Components of the assessments that are praised by proponents include the idea that the digital delivery system helps to broaden the availability of tools and accommodations and creates a less restrictive testing environment (SBAC.org). Opponents have concerns about the way accommodations are applied, including the over- identification of accommodations, as well as the choice of specific accommodations that may run counter to the fundamental goal of using them to begin with (Geller, et.al., 2007). These issues may become more prevalent with the new testing format as accommodations embedded in the test delivery system are easily accessible and the tendency to provide more than what is needed may become common. If high-stakes tests are going to provide useful information for guiding students in special education then the choices about accommodations need to be made carefully and consistently, they need to be used in the instructional setting, and they need to be individually-determined, not disability-specific (Salend, 2008). Most importantly, accommodations should be continually evaluated for validity, usefulness, and fairness. The balance of standardization and individualization is an issue that frequently comes up in the literature about this topic and is one of the most crucial factors when planning participation in high-stakes testing for students in special education. A study of the students with disabilities (SWD's) subgroup published in 2012 by EDfacts, a United States Department of Education initiative to collect and place K-12 performance data at the center of policy, management, and budget decisions reported an achievement gap that exceeded 30 percentage points between SWD's and typical peers in 2007-2008. This has led to greater attention being placed on the appropriateness and effectiveness of the types of supports that are

in place for students in special education when participating in high-stakes testing, and research will need to continue to determine how high-stakes tests can be useful tools in providing the types of information about students needed for effective instruction.

Research from the past decade shows that retained students are more likely to drop out of school due to lowered self-efficacy, compounding feelings of failure, and a negative attitude toward school (Allensworth, 2005). The question about there being a connection between high-stakes testing and dropout rates is one that has been examined and will require continued research with the implementation of new assessments. For students in special education the dropout rate is twice that of other students, and they are among the lowest performing students on high-stakes tests (Thurlow, Sinclair, & Johnson, 2002). The U.S. Department of Education, Office of Special Education Programs reported in 2006 that 37.6% of all students with disabilities dropped out of school at age 14 or over. Of that 37.6%, 61.2 % were students with EBD, 35.8% were students with a speech/language disorder, and 35.4% were students with a learning disability.

The role that high-stakes testing may play in these dropout rates is difficult to determine. One study, published in 2008 looked at the extraordinarily high rates of dropout under Texas's system of high-stakes, test-based accountability. Since the model for NCLB came from the Texas system the correlation between the Texas system and high dropout rates may be predictive of the rest of the U.S. under NCLB. The study showed systematic incentives to take administrative action that encouraged low-scoring students (the lowest scoring being students in special education) to drop out or be retained to keep their scores from counting toward accountability, and a relationship between the dropout of increasing numbers of students and rising accountability ratings was established (McNeil, et.al., 2008). And, as reported by McNeil, et al. (2008), the problem was not limited to Texas. Other published studies found higher rates of retention and dropout in states and cities that have instituted more stringent graduation requirements and exit exams. Data from the National Educational Longitudinal Survey found that graduation exit exams increased the probability of dropping out among the lowest-ability students, typically students in special education (McNeil, et.al, 2008). By 2010, 28 states had implemented exit exams and 24 of the states used these exams for meeting graduation accountability mandates under NCLB. A number of people have filed lawsuits that have challenged the use of high-stakes tests as graduation requirements and legal issues arise for students in special education as the use of these exams runs contrary to the provisions set forth in IDEA (Yell, Katsiyannis, Collins, & Lasinski, 2012). The role that these exit exams play in students in special education dropping out requires further investigation. How the new assessments that are aligned with the CCSS will address this issue has yet to be determined.

Questions/Hypotheses

The new large-scale assessments aligned with the CCSS have been developed to include more accessible accommodations for all students, and are claimed to be more appropriate as they are aligned with the standards that students are required to meet through classroom instruction. However, the construction of the assessments using principles of universal design does not guarantee a more authentic testing experience for students in special education, nor does it guarantee a decrease in dropout rates. The possibility that the new assessments may exacerbate

the present issues surrounding accommodations and dropout rates due to the increased pressure for performance exists. Continued research is needed in both of these areas.

In the area of accommodations several factors need to be considered when determining if the assessments will have a positive impact on identifying appropriate accommodations to be used in both instruction and testing. There is an absence of a firm research base in determining accommodations for students in special education (www.sbac.org). In addition, the tools and accommodations provided in the new assessments were implemented without a thorough amount of baseline data being collected (Lane, 2013). Additionally, results of pilot tests were not used to inform practitioners or families about student performance or other factors, such as difficulties with the testing experience. These types of issues may conflict with both NCLB and IDEA compliance and result in a disproportionate representation of which students receive accommodations, and may interfere with the decision-making process (Salend, 2008). The Smarter Balanced consortium has stated that since the digitally-delivered tools and accommodations are new, additional research is needed as part of the validation process for the assessments. One concern over the accommodations provided is that rather than being tools that allow students in special education to appropriately demonstrate their abilities they may instead pose barriers because of the amount of technological knowledge required to navigate them (Geller, et.al, 2007). Research should include factors related to specific accommodations and their effectiveness in allowing students in special education to use them appropriately to demonstrate achievement. A proper research base for these components will allow for more effective decision-making regarding accommodations and strategies used in both instructional and testing domains.

Determination of a connection between high-stakes testing and retention and dropout rates will require a detailed examination of the factors that are involved. One factor, that retained students are more likely to drop out due to lowered self-efficacy and a negative feeling about school, is an issue that has been clearly documented (Allensworth, 2005). Recurring failure in school is one of the most significant predictors of dropout. Because the dropout rate of students in special education is twice that of other students this population is at greater risk of experiencing consistent failure and is more likely to give up on school (Thurlow, et al., 2002). Proponents of the new assessment system believe that since there is a stronger connection between what students are exposed to in their instruction and what they are assessed on there will be a clearer purpose to the testing, and higher standards will increase students motivation to do well (Linn, 2013). However, opponents will cite those same factors as reasons that more students will experience failure. Because the new standards are more rigorous, students in special education may experience more difficulties in learning, especially if the accommodations provided are not useful or effective, and since the stakes are so much higher for performance, struggling students may be more inclined to give up. The heightened expectations and new performance standards will result in many students, particularly those in special education, being identified as not yet ready for college or career (Jones & King, 2012). The other factor related to dropout rates tied to high-stakes testing is the systematic encouragement from administration for low-performing students to be retained or to drop out so their scores do not negatively impact accountability ratings (McNeil, et al., 2008). The lowest performing students on high-stakes tests are students in special education, and these students may end up being the targets of this process. Further

research is needed to uncover these types of practices in schools so that equity in educational opportunities can be realized for students in special education.

Conclusion

High-stakes testing has been a strong presence in American education, particularly in the last few decades. For students in special education, high-stakes testing is not simply a requirement to be fulfilled, but often a determinant of their future. High-stakes assessments administered in the past were not fair or accurate representations of the abilities of students in special education. With the advent of the Common Core State Standards and the aligned assessments, changes have been made to the assessment process that aims to provide more and better access for students in special education so that the results of high-stakes tests can be considered accurate and fair representations of the abilities of this population. The utilization of universal design is meant to avoid the previous and ineffective practice of trying to retrofit tests for students in special education (www.udlcenter.org). However, many questions regarding the appropriate use of accommodations and how the testing process itself can provide useful information to guide classroom instruction for students in special education have been proposed (Salend, 2008). There is needed research in the area of accommodations, and the results from the first round of next-generation assessments will need to be examined to identify specific components of the assessments involving accommodations and accessibility tools. One method for gathering this information should include student response data in which students in special education are given an opportunity to discuss their experience using specific tools and accommodations.

The issue of retention and dropout for students in special education and how high-stakes testing plays a role is an area that has been of concern in education for some time. Students in special education are often the ones who perform lowest on high-stakes assessments which puts them at greater risk of either being retained due to their poor performance or dropping out. Research shows that students who experience consistent failure are at the highest risk for dropping out, and students in special education are typically the ones to experience the most failure in school (Thurlow, et al., 2002). The new assessments, while designed to attempt to meet the needs of a broader range of students, may contribute to higher rates of dropout as students are identified as not being on-track for college and career. As the assessments are implemented data regarding the outcome on retention and dropout rates will need to be carefully examined, as well as the specific contributing factors.

High-stakes testing will continue to be a part of the American educational landscape. Having an adequate research base regarding the appropriate use of accommodations in instruction and testing, and researching and identifying the factors that contribute to the retention and dropout rates of students in special education is required. The role that high-stakes testing plays in these specific areas requires further research if they are going to be considered useful and valid aspects of the educational experience for students in special education.

References

Allensworth, E. (2005). Dropout rates after high-stakes testing in elementary school: A study of the contradictory efforts of Chicago's efforts to end social promotion. *Educational Evaluation and Policy Analysis, 27*(4), 341-364.

Common Core State Standards Initiative. Retrieved from http://corestandards.org.

Fuchs, L., Fuchs, D., & Cappizzi, A. (2005). Identifying appropriate test accommodations for students with learning disabilities. *Focus on Exceptional Children, 37*(6).

Geller, L. Alonzo, J., Monegan, J., & Tindal, G. (2007). Recommendations for accommodations: Implications of (in)consistency. *Remedial and Special Education, 28(*4), 194-206.

Hager, K. & Slocum, T. (2005). Using alternate assessment to improve educational outcomes. *Rural Special Education Quarterly, 24*(1), 24-29.

Jones, A. & King, J. (2012). The common core state standards: A vital tool for higher education. *Change*. Retrieved from http://www.changemag.org.

Kaestle, C. (2013). *Testing policy in the United States: A historical perspective* (The Gordon Commission on the Future of Assessment in Education). Retrieved from http://www.gordoncommission.org.

Katsyiannis, A., Zhang, D., Joseph, R., & Jones, J. (2007). High-stakes testing and students with disabilities. *Journal of Disability Policy Studies, 18*(3), 160-167.

Lane, S. (2013). The need for a principled approach for examining indirect effects of test use. *Measurement, 11*, 44-46.

Linn, R. (2013). *Test-based accountability* (The Gordon Commission on the Future of Assessment in Education). Retrieved from http://www.gordoncommission.org.

McLaughlin, M., & Thurlow, M. (2010). Educational accountability and students with disabilities: Issues and challenges. *Educational Policy, 17*(4), 431-451.

McNeil, L., Coppola, E., Radigan, J., & Helig, J. (2008). Avoidable losses: High-stakes accountability and the dropout crisis. *Education Policy Analysis Archives, 16*(3).

National Academy of Education (2009). *Standards, assessment, and accountability* (Education Policy White Paper). Washington, DC: Author.

National Center on Universal Design (2014). Retrieved from http://www.udlcenter.org.

Nirvi, S. (2012). Challenges seen in testing special ed. Pupils on common core. *Education Week, 31*(33).

Salend, S. (2008). Determining appropriate testing accommodations: Complying with NCLB and IDEA. *Teaching Exceptional Children, 40*(4), 14-22.

Salvia, J., Ysseldyke, J., & Bolt, S. (2010). *Assessment: In special and inclusive education.* Wadsworth: California

Samuels, C. (2014). Graduation disparities loom large for students with special needs. *Education Week, 33*(19).

Smarter Balanced Assessment Consortium (2014). Retrieved from http://smarterbalanced.org.

Stephens, W. (2014). Testing under the microscope: How aligned assessments place demands on time, technology & connectivity. *Knowledge Quest, 43*(1), 31-35.

Thurlow, M., Sinclair, M., & Johnson, D. (2002). Students with disabilities who drop out of school: Current challenges in secondary education and transition. *National Center on Secondary Education and Transition Issue Brief, 1*(2).

U.S. Department of Education, National Center for Education Evaluation and Regional Assistance (2012). *Inclusion of students with disabilities in school accountability systems* (NCEE Publication No. 2012-4056). Retrieved from http://www.ed.gov/ies.ed.gov.

U.S. Department of Education, Office of Special Education and Rehabilitation Services, Office of Special Education Programs (2006, April). *26th Annual (2004) Report to Congress on the Implementation of the Individuals with Disabilities Education Act, Vol. 1*. Washington, DC: Author.

U.S. Department of Education (2012). Race to the Top Initiative. Retrieved from http://www2.ed.gov/racetothetop.

Yell, M., Katsiyannis, A., Collins, J., & Losinski, M. (2012). Exit exams, high-stakes testing, and students with disabilities: A persistent challenge. *Intervention in School and Clinic, 48*(1), 60-64.

About the Author

Lola Gordon, Ed.S., has been working in the field of special education for 12 years. She has taught students with various disabilities in both public school and alternative school settings, focusing mainly on students with emotional and behavioral disorders. She has also taught at the college level, instructing courses about children and adults with disabilities. Lola is currently a full-time doctoral student at the University of Connecticut, working toward her PhD in Educational Psychology with a focus on behavioral supports. She also has a strong interest in special education law, and is focusing her studies on bringing more awareness regarding the behavioral and mental health needs of students with disabilities into education policy.