

Using High-Stakes Assessments for Grade Retention and Graduation Decisions

A POSITION STATEMENT OF THE INTERNATIONAL READING ASSOCIATION



Students' knowledge and skills are assessed in multiple ways throughout their elementary and secondary school years. Some of these assessments are considered "high stakes" because a poor performance on them at critical junctures during a student's school career can have life-altering consequences, including a school, district, or state policy to retain the student in a grade or to deny the student a high school diploma. This position statement examines issues and policies related to such uses of assessments and offers stakeholders recommendations concerning these practices.

It is the position of the International Reading Association (IRA) that grade retention and high school graduation decisions must be based on a more complete picture of a student's literacy performance, obtained from a variety of systematic assessments, including informal observations, formative assessments of schoolwork, and consideration of out-of-school literacies, as well as results on standardized formal measures. Further, it is the position of IRA that in addition to these considerations, teachers' professional judgment should be a major factor in such decisions, along with input from students and their families.

The Prevalence of Using High-Stakes Assessments for Grade Retention and Graduation Decisions

It should be noted that few nations outside of the United States mandate using assessment results for purposes of determining grade promotion or graduation from secondary school. For instance, in Singapore, exam results may be used to stream students into particular types of secondary schools, ranging from the advanced express schools to technical programs (OECD, 2011b), or to place elementary-grade students in content learning bands that are designed to meet ranges of student ability (Singapore Ministry of Education, 2013), but results are not used to retain students in the elementary grades or to deny high school diplomas. Further, Japan and Finland rarely practice grade retention (OECD, 2011a), and in its study of retention practices among European countries, the Eurydice Network (2011) found great variation in practices yet concluded that it was primarily educational culture and teacher judgment that determined grade retention and that parents are increasingly included in making decisions. Practice in Latin American countries also differs from that of the United States; in Latin America, a combination of performance measures and minimum attendance requirements (in days) are typically used for grade promotion decision making (Koppensteiner, 2011). Across the United States, however, states are attributing high stakes to statewide tests and end-of-course exams by mandating grade retention and/or withholding high school diplomas on the basis of student performance on these assessments. Thus, the concerns,

recommendations, and data provided here are drawn primarily from the U.S. school experience.

With the inception in the United States of the No Child Left Behind (NCLB) Act of 2001 in early 2002, schools receiving Title I funds (approximately 50% of public schools) were required to demonstrate that their students made adequate yearly progress (AYP) on state-developed assessments that used state standards for reading (literacy) and mathematics, as well as one additional indicator that, for most states, meant an attendance indicator at the elementary- and middle-grade levels and graduation at the high school level (Frey, 2010). States set their own AYP targets each year, but the goal of NCLB was to have 100% of students proficient by 2013–2014. Assessment results, available to the public, had high stakes attached to them. Based on their performance on these assessments, individual students were identified as below basic, basic, proficient, or advanced. Schools had to provide separate reporting of results for economically disadvantaged students, for students with disabilities, for limited English proficient students, and by race/ethnicity. Results were used to identify schools in need of improvement, and those so identified were given up to six years to make AYP before being required to restructure (GreatSchools, n.d.). These were high-stakes assessments.

Most states are now transitioning to new high-stakes assessments to evaluate a student's performance in reading, mathematics, and writing that are aligned with the national Common Core State Standards, a state-led initiative coordinated by the National Governors Association Center for Best Practices and the Council of Chief State School Officers (2010).

Obtaining a Complete Profile of Literacy Performance

Individuals vary in what they know and can do, as well as whether and how they make known their knowledge and skills to others. Student performance on high-stakes assessments is increasingly being used to evaluate a teacher's effectiveness. Thus, literacy instruction frequently centers on teaching to the test, and classroom assessments are often designed to match the high-stakes assessment format. Linn (2003) explains, "It is no surprise that attaching high stakes to test results in an accountability system leads to a narrowing of the instructional focus of teachers and principals" (p. 4). Consequently, students have limited opportunities to demonstrate the full extent of their literacy knowledge.

Curriculum narrowing...reduces many students' chances of being thought talented in school and results in a restriction in the creative and enjoyable activities engaged in by teachers and students. The tests commonly used with

narrower curricula also appear to restrict thinking skills. (Berliner, 2011, p. 287)

In such classrooms, a teacher may erroneously conclude that a student lacks literacy proficiency because such a conclusion is based on limited information.

However, in classrooms where high-stakes assessments do not drive the curriculum or test design, teachers are able to obtain a much richer profile of a student's performance. In these classrooms, students have opportunities to demonstrate their literacy achievement throughout the school year through such varied activities as completing reading and writing tasks, taking teacher-made tests, engaging in reading and writing conferences, completing independent and group projects, or performing in plays. Many teachers use these activities as formative assessments. Teachers' focused observations and anecdotal records document student reading and writing performance during authentic learning tasks throughout the school day and also contribute to formative assessment. Such assessments inform instruction and are thus purposeful, collaborative, dynamic, and descriptive; they contribute to the continuous improvement of teaching and student learning (IRA, 2013). These frequent, ongoing, low-stakes assessments allow teachers to monitor student learning and enable them to adjust instruction, and these assessments provide opportunities for students to reassess their learning goals. By design, as they review results of formative assessments, teachers can also consider how the unique attributes of individual learners, such as home language, culture, strengths, and needs (Risko & Walker-Dalhouse, 2010), affect student performance. The information obtained from formative assessments can make an important contribution to teacher decisions about student needs and progress, including decisions about retention and graduation.

Limitations of High-Stakes Assessments for Grade Retention/ Graduation Decision Making

Summative assessments may include standardized tests, state assessments, end-of-course exams, and other assessments such as those being developed by the Smarter Balanced Assessment Consortium and the Partnership for Assessment of Readiness for College and Careers to evaluate student performance in relation to the Common Core State Standards (Doorey, 2012). Unlike formative assessments, summative assessments may be described as "assessments of learning" because they assess cumulative learning of specific benchmarks, criteria, or standards over a longer period of time than do formative assessments, and thus they cannot be used on an ongoing basis to guide instruction. Because they assess cumulative learning,

summative assessments may also be used to evaluate the effectiveness of curricula and instructional programs (Afflerbach, 2012). Increasingly, however, results of student performance on summative assessments are being used inappropriately as tools to reward, evaluate, or punish students and/or teachers. Thus, they are often referred to as *high-stakes* assessments.

Although they can provide effective program evaluation information, high-stakes assessments do not provide the qualitative, diagnostic, and formative information that helps teachers plan for differentiated instruction and learning opportunities for each of their students, in *all* of their diversity (Flippo, 2014). High-stakes tests are constructed based on prescribed standards. That is not only their strength but also their limitation. They do not typically consider a number of important student dimensions of learning that affect student performance, including the following (Flippo, 2014):

- Depth and range of prior knowledge and other cognitive experiences
- Language diversity and experiences
- Family, community, cultural, and sociocultural experiences
- Linguistic and cognitive strategy use to get important information
- Motivations, aspirations, and goals

Penfield (2010) notes that grade retention based on the results of a single high-stakes test does not meet professional standards for fair and appropriate test use. This view is corroborated by Standard 8 in *Standards for the Assessment of Reading and Writing* (Joint Task Force on Assessment of the IRA and the National Council of Teachers of English, 2010), which advises that the assessment process should involve multiple perspectives and sources of data. Because of the complex nature of reading and writing skills and strategies, a single test cannot measure all aspects of literacy accurately. High-stakes assessments do not provide sufficient information regarding students' performance and thus are not appropriate to use for making grade retention or graduation decisions. As the Joint Task Force on Assessment states, "Such decisions are simply too important to make on the basis of a single measure, evaluation tool, or perspective" (p. 25).

Current Policies and Practices Linking Retention and Graduation to High-Stakes Assessments

Extensive use of high-stakes assessments for determining a student's advancement to the next grade or to high school graduation remains very evident in the United States,

despite the fact that beginning in 2011, many states applied for and have received waivers from the U.S. Department of Education's AYP requirement of the NCLB. Such waivers, according to the U.S. Department of Education (2011), allow flexibility such that "states will begin to move beyond the bubble tests and dumbed-down standards that are based on arbitrary standards of proficiency," and will provide states with opportunities to "make accountability decisions based on student growth and progress, as well as other measures of student learning and school performance. They will consider more than a single test score measured against an arbitrary proficiency level" (p. 2). A review of current policies suggests that states have not taken advantage of this flexibility.

Retention in the Early Grades

Policies and discussions in the United States about grade retention are often directed at grade 3, perhaps because of the conventional thinking that if a child has not learned to read by the end of third grade, he or she will struggle throughout school. Torgesen, Wagner, Rashotte, Alexander, and Conway (1997) found, however, that with appropriate research-based intervention, older readers develop the skills they missed in the primary grades and can reach grade level in one to two years. Nonetheless, in 2012, the Education Commission of the States identified that 14 states and the District of Columbia required third-grade students to "be proficient in reading, attain a specific score on a state-wide reading exam, or otherwise meet a defined literacy benchmark in order to be promoted to 4th grade" (Rose, 2012, p. 3). According to the commission's 2013 survey, the number of states using high-stakes assessments for third-grade retention decisions had increased to 15, with four additional states considering adopting this requirement (Lu, 2013). Many children are affected by such policies. In Arizona, for instance, "The Arizona Department of Education estimates that the law taking effect this fall will force about 1,500 children to repeat third grade next year. Another 17,000 third graders are at risk of being held back under the new rule" (Reid, 2013, para. 2).

African American and Hispanic students and students living in poverty are most affected by grade retention practices that use the results of high-stakes assessments for decision making. Achievement patterns reveal wide disparities between the achievement of white students and that of African Americans and Hispanics (National Center for Education Statistics, 2011); thus, it follows that there would be similar differences in the number of students retained in each subgroup. In 2009–2010, African American students represented 49% and 56% of the third and fourth graders who were retained, respectively, which was disproportionate to their representation in those grades; Hispanic students were twice as likely to be retained than their white counterparts (Adams, Robelen, &

Shah, 2012). In *The Condition of Education 2013*, Aud and colleagues (2013) also report,

the percentages of repeating kindergartners were higher for students living below the poverty threshold (10 percent) or for those living between 100 percent and 199 percent of the poverty threshold (6 percent) than for students living in households that were at 200 percent or more of the poverty threshold (4 percent). (p. 37)

High School Exit Exams

High-stakes tests are also used frequently to determine whether a student will graduate from high school on time. NCLB legislation required the establishment of an assessment system for high school students but did not require test results as a criterion for graduation. Prior to the passage of NCLB in 2002, 18 states already required passing scores on exit tests for graduation (Snyder & Dillow, 2012). By 2012, this increased to 25 states, with 69% of all high school students being affected by this practice. An even greater percentage of minority students are enrolled in high schools using exit exams to determine graduation (McIntosh, 2012). Some states are beginning to reassess the value of exit exams for graduation, with several (e.g., Alabama, Arizona, Georgia, Nevada, South Carolina, Tennessee) recently announcing that they were eliminating these exams, phasing them out, or replacing them with more frequent course exit exams. Yet, nearly half of the states using exit exams for graduation determinations have not placed reconsideration of this practice on their legislative agenda (Barnett, 2013). Further, two states (Connecticut and Rhode Island) were adding an exit exam in 2020 and 2014, respectively (Swanson & Lloyd, 2013). It should be noted that despite this focus on requisite literacy proficiency for graduation, only 16 states include resources in their budgets for remediation for struggling students who did not pass exit, or end-of-course, exams (Swanson & Lloyd, 2013).

As with the outcomes of third-grade retention policies, African Americans, Hispanics, and students living in poverty are most affected by the use of high-stakes assessments for diploma decisions.

While progress is encouraging, a deeper look at the data reveals that gains in graduation rates and declines in dropout factory high schools occurred unevenly across states and subgroups of students....[Further], large "graduation gaps" remain in many states among students of different races, ethnicities, family incomes, disabilities and limited English proficiencies. (Balfanz, Bridgeland, Bruce, & Fox, 2013, p. 5)

Policymakers may believe that linking grade retention and high school graduation to students' results on high-stakes assessments will motivate students to perform better, but instead, evidence indicates that these practices have harsh and lasting consequences for students academically,

psychologically, socially, and economically (Baker & Lang, 2012; Jimerson, 1999; Jimerson, Anderson, & Whipple, 2002; Norton, 2011; Walker & Madhere, 1987; Yamamoto & Byrnes, 1984).

Conclusion

IRA recognizes the need to improve the literacy achievement of many students. However, high-stakes assessments do not provide a complete picture of students' literacy knowledge and accomplishments and should not be used to make decisions about student grade retention or high school graduation.

Recommendations

To address concerns about the use of high-stakes assessments for making grade retention and graduation decisions, IRA recommends the following actions:

1. Grade retention and graduation decisions should be based on multiple assessments, including teacher professional judgment, results of formative assessments, and student and family input, as well as results from standardized literacy assessments.
2. Schools, school districts, and policymakers should be guided by the expertise of professional associations and literacy professionals when making decisions about how to best utilize results obtained from high-stakes literacy assessments.
3. Professional development should be available for teachers on assessment strategies for obtaining a complete picture of a student's literacy performance.

REFERENCES

- Adams, C.J., Robelen, E.W., & Shah, N. (2012). Data show retention disparities. *Education Week*, 31(23), 18–20. Retrieved from www.edweek.org/ew/articles/2012/03/07/23data_ep-2.h31.html
- Afflerbach, P. (2012). *Understanding and using reading assessment*, K–12 (2nd ed.). Newark, DE: International Reading Association.
- Aud, S., Wilkinson-Flicker, S., Kristapovich, P., Rathbun, A., Wang, X., & Zhang, J. (2013). *The condition of education 2013* (NCES 2013-037). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Retrieved from nces.ed.gov/pubs2013/2013037.pdf
- Baker, O., & Lang, K. (2012). *The effect of high school exit exams on graduation, employment and wages*. Retrieved from people.bu.edu/obaker/files/Olesya_Baker_JMP.pdf
- Balfanz, R., Bridgeland, J.M., Bruce, M., & Fox, J.H. (2013). *Building a grad nation: Progress and challenge in ending the high school dropout epidemic—2013 annual update*. Washington, DC: Civic Enterprises; Baltimore, MD: Everyone Graduates Center, School of Education, Johns Hopkins University; Washington, DC: America's Promise Alliance & Alliance for Excellent Education. Retrieved from www.civicerprises.net/MediaLibrary/Docs/Building-A-Grad-Nation-Report-2013_Full_v1.pdf
- Barnett, R. (2013, May 22). Exit exams may be on their way out. *USA Today*. Retrieved from www.usatoday.com/story/news/nation/2013/05/22/exit-exams-high-school-diploma/2351009
- Berliner, D. (2011). Rational responses to high stakes testing: The case of curriculum narrowing and the harm that follows. *Cambridge Journal of Education*, 41(3), 287–302. doi:10.1080/0305764X.2011.607151
- Doorey, N.A. (2012). Coming soon: A new generation of assessments. *Educational Leadership*, 70(4), 28–34.
- Eurydice Network. (2011). *Grade retention during compulsory education in Europe: Regulations and statistics*. Brussels, Belgium: Education, Audiovisual and Culture Executive Agency. Retrieved from eacea.ec.europa.eu/education/eurydice/documents/thematic_reports/126EN.pdf
- Flippo, R.F. (2014). *Assessing readers: Qualitative diagnosis and instruction* (2nd ed.). New York, NY: Routledge; Newark, DE: International Reading Association.
- Frey, D. (2010). *Measuring adequate yearly progress: What "other" indicators count besides reading and math?* Denver, CO: Education Commission of the States. Retrieved from www.ecs.org/clearinghouse/87/98/8798.pdf
- GreatSchools. (n.d.). *No Child Left Behind (NCLB) requirements for schools*. Oakland, CA: Author. Retrieved from www.greatschools.org/definitions/nclb/nclb.html
- International Reading Association. (2013). *Formative assessment: A position statement of the International Reading Association*. Newark, DE: Author. Retrieved from www.reading.org/Libraries/position-statements-and-resolutions/ps1080_formative_assessment_web.pdf
- Jimerson, S.R. (1999). On the failure of failure: Examining the association between early grade retention and education and employment outcomes during late adolescence. *Journal of School Psychology*, 37(3), 243–272. doi:10.1016/S0022-4405(99)00005-9
- Jimerson, S.R., Anderson, G.E., & Whipple, A.D. (2002). Winning the battle and losing the war: Examining the relation between grade retention and dropping out of high school. *Psychology in the Schools*, 39(4), 441–457. doi:10.1002/pits.10046
- Joint Task Force on Assessment of the International Reading Association and the National Council of Teachers of English. (2010). *Standards for the assessment of reading and writing* (Rev. ed.). Newark, DE: International Reading Association; Urbana, IL: National Council of Teachers of English. Retrieved from www.reading.org/General/CurrentResearch/Standards/AssessmentStandards.aspx
- Koppensteiner, M.F. (2011). *Automatic grade promotion and student performance: Evidence from Brazil* (Working Paper No. 11/52). Retrieved from www.le.ac.uk/ec/research/RePEc/lec/leecon/dp11-52.pdf
- Linn, R.L. (2003). *Accountability: Responsibility and reasonable expectations* (CSE Report 601). Los Angeles: Center for the Study of Evaluation, National Center for Research on Evaluation, Standards, and Student Testing, Graduate School of Education and Information Studies, University of California, Los Angeles. Retrieved from www.cse.ucla.edu/products/reports/r601.pdf
- Lu, A. (2013, November 15). States insist on third grade reading proficiency. *Stateline*. Retrieved from www.pewstates.org/projects/stateline/headlines/states-insist-on-third-grade-reading-proficiency-85899520263
- McIntosh, S. (2012). *State high school exit exams: A policy in transition*. Washington, DC: Center on Education Policy, Graduate School of Education and Human Development, George Washington University. Retrieved from www.cep-dc.org/displayDocument.cfm?DocumentID=408
- National Center for Education Statistics. (2011). *Reading 2011: National Assessment of Educational Progress at grades 4 and 8* (NCES 2012-457). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.
- National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards for English language arts and literacy in history/social studies, science, and technical subjects*. Washington, DC: Authors. Retrieved from www.corestandards.org/assets/CCSSI_ELA%20Standards.pdf

- Norton, M.S. (2011). Please, not another push to get tough on student retention. *Planning and Changing*, 42(3/4), 209–223.
- OECD. (2011a). *Lessons from PISA for the United States: Strong performers and successful reformers in education*. Paris, France: Author. Retrieved from www.oecd.org/pisa/46623978.pdf
- OECD. (2011b). Singapore: Rapid improvement followed by strong performance. In *Lessons from PISA for the United States: Strong performers and successful reformers in education* (pp. 159–176). Paris, France: Author. Retrieved from www.oecd.org/countries/singapore/46581101.pdf
- Penfield, R.D. (2010). Test-based grade retention: Does it stand up to professional standards for fair and appropriate test use? *Educational Researcher*, 39(2), 110–119. doi:10.3102/0013189X10363007
- Reid, J. (2013, July 24). Reading, writing, 'rithmetic—retention? New 3rd-grade standards. *TucsonSentinel.com*. Retrieved from www.tucsonsentinel.com/local/report/072413_reading_standards/reading-writing-rithmetic-8211-retention-new-3rd-grade-standards
- Risko, V.J., & Walker-Dalhouse, D. (2010). Making the most of assessments to inform instruction. *The Reading Teacher*, 63(5), 420–422. doi:10.1598/RT.63.5.7
- Rose, S. (2012). *Third grade reading policies*. Denver, CO: Education Commission of the States. Retrieved from www.ecs.org/clearinghouse/01/03/47/10347.pdf
- Singapore Ministry of Education. (2013). *Subject-based banding in primary schools: Catering to your child's abilities*. Singapore: Author. Retrieved from www.moe.gov.sg/education/primary/files/subject-based-banding.pdf
- Snyder, T.D., & Dillow, S.A. (2012). *Digest of education statistics 2011* (NCES 2012-001). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Retrieved from nces.ed.gov/pubs2012/2012001.pdf
- Swanson, C.B., & Lloyd, S.C. (2013). Graduation rate approaching milestone. *Education Week*, 32(34), 22–28. Retrieved from www.edweek.org/media/downloads/files/educationweek_diplomas_count_2013.pdf?r=1598499473
- Torgesen, J.K., Wagner, R.K., Rashotte, C.A., Alexander, A.W., & Conway, T. (1997). Preventive and remedial interventions for children with severe reading disabilities. *Learning Disabilities*, 8(1), 51–61.
- U.S. Department of Education. (2011). *What ESEA flexibility means for students, teachers, and parents: Answering the public's questions*. Washington, DC: Author. Retrieved from www2.ed.gov/policy/elsec/guid/esea-flexibility/index.html
- Walker, E.M., & Madhere, S. (1987). Multiple retentions: Some consequences for the cognitive and affective maturation of minority elementary students. *Urban Education*, 22(1), 85–102.
- Yamamoto, K., & Byrnes, D.A. (1984). Classroom social status, ethnicity, and ratings of stressful events. *Journal of Educational Research*, 77(5), 283–286.

International Reading Association High-Stakes Assessments Board Task Force 2013–2014

Chair

Jill D. Lewis-Spector, New Jersey City University, Jersey City, New Jersey, USA

Members

Laurie A. Elish-Piper, Northern Illinois University, DeKalb, Illinois, USA

Rona F. Flippo, University of Massachusetts Boston, Boston, Massachusetts, USA

Maryann Manning, University of Alabama at Birmingham, Birmingham, Alabama, USA

Suzanne N. Nakashima, Lincrest Elementary School, Yuba City, California, USA

IRA Board of Directors at Time of Adoption:

Maureen McLaughlin, *President*
Jill D. Lewis-Spector, *President-elect*

Heather I. Bell
Steven L. Layne
William H. Teale
Douglas Fisher
Rona F. Flippo
Shelley Stagg Peterson
Bernadette Dwyer
Laurie A. Elish-Piper
Julianne Scullen

This position statement is available in PDF form for free download through the International Reading Association's website: www.reading.org.

© 2014 International Reading Association

INTERNATIONAL
Reading Association
800 BARNSDALE ROAD, PO BOX 8139
NEWARK, DE 19714-8139, USA
www.reading.org

Suggested APA Reference

International Reading Association. (2014). *Using high-stakes assessments for grade retention and graduation decisions* [Position statement]. Newark, DE: Author.