
U.S. Policies Supporting Inclusive Assessments for Students with Disabilities

2

Susan C. Weigert

U.S. Policies Supporting Inclusive Assessment of Students with Disabilities

This overview of U.S. policies supporting inclusive assessment practices traces the policies and supporting educational contexts within the historical framework of key legislation and regulations over the past 50 years.

Assessment Policies in the 1960s and 1970s: Inclusion and ‘Equal Terms’

The history of inclusive assessment policies in the United States has been a product of many political and practical influences, but at the root of their developments has been the central and fundamental tenet of equal protection. Title VI of the Civil Rights Act of 1964 prohibited discrimination on the basis of race, sex, color, or national origin, rather than on the basis of disability. Yet the spirit of the law swept students with disabilities (SWDs) into its strong political current and

promised to ensure and protect the equity of their educational opportunities. The passion for equal access to education as a civil right was best characterized by Chief Justice Warren’s opinion on *Brown v. Board of Education* in 1954:

In these days it is doubtful that any child may reasonably be expected to succeed in life if he is denied the opportunity of an education. Such an opportunity, where the State has undertaken to provide it, is a right that must be made available to all on equal terms. (*Brown v. Board of Education*, 347 U.S. 483 (1954), quoted in Russo & Osborne, 2008 p. 493)

While policies of inclusion constituted a popular solution to problems of inequity in public education, compulsory education statutes during the 1960s and early 70s left the authority to school districts to decide whether SWDs could ‘benefit’ from instruction (Russo & Osborne, 2008). The new inclusion principles were eventually codified in PL 93-112, Section 504 of the Rehabilitation Act of 1973, which prohibited discrimination against individuals with disabilities in federally funded programs, and required reasonable accommodations for students with physical or mental impairments that ‘substantially limited’ them in one or more major life activities, including learning (29 USC § 706 (7)(B)). In addition to physical and sensory handicaps, Section 504 of the Rehabilitation Act applied to persons with ‘mental’ disabilities such as mental retardation, traumatic or organic brain syndromes, emotional disturbance, specific learning

S.C. Weigert (✉)
U.S. Department of Education, Office of Special Education Programs, Alexandria, VA, USA
e-mail: susan.weigert@ed.gov

The views expressed in this chapter are solely those of the author in her private capacity. No official support or endorsement by the U.S. Department of Education is intended nor should it be inferred.

disabilities, and other cognitively disabling conditions (Phillips, 1994).

The Rehabilitation Act further defined the meaning of a free, appropriate, public education (FAPE), and specified that appropriate education included educational services designed to meet the individual education needs of students with disabilities ‘as adequately’ as the needs of non-disabled students were met. Yet the only assessment-related provisions of the Act were those requiring that assessments be provided in a child’s ‘normal mode of communication’ (including native language) unless it was clearly not feasible to do so.

The principle of inclusion in the Rehabilitation Act was later incorporated into the elementary and secondary education act (ESEA) amendments of 1974 (PL 93-380), which also mandated the ‘free appropriate public education (FAPE)’ in the ‘least restrictive environment (LRE).’ These provisions were codified a year later in the Education for All Handicapped Children Act (P.L. 94-142). Yet at the time, FAPE simply meant access to special education and related services—in conformity with individualized academic and behavioral goals stated in the student’s IEP, rather than connoting access to the general education curriculum. A new requirement for inclusive assessment of SWDs in PL 94-142 §612 (5)(C) mandated that testing and evaluative materials used for evaluation and placement not be the sole criterion for determining an appropriate educational program for a child with a disability.

The 1977 regulations amending the Rehabilitation Act of 1973 (§104.35) required that tests and other evaluation materials meet requirements for validity for the specific purpose for which the tests were used, and that they be administered by trained personnel in conformity with test-developer’s instructions. Further, the type of assessments to be used for educational evaluation of SWDs were to include those tailored to assess specific areas of educational needs, not merely those designed to measure a student’s I.Q. Finally, educational tests were to be selected and administered to ensure that the results of testing accurately reflected the student’s educational aptitude or achievement level (or

whatever educational factor was to be measured by the test), rather than merely reflecting the student’s impaired sensory, manual, or speaking skills (except when those skills were the factors measured by the test). The protections against ‘disparate impact’ of assessments for SWDs was restricted to ‘otherwise qualified’ individuals with disabilities, meaning that the student who might have impaired sensory, manual, or speaking skills, still had to be capable of meeting the standards required to pass the test. As U.S. Supreme Court justice Powell commented:

Section 504 imposes no requirement upon an educational institution to lower or to effect substantial modifications of standards to accommodate a handicapped person. (Southeastern Community College v. Davis 442 U.S. 397, 1979 Supreme Court of United States)

The 1980s and 1990s: IEP as Curriculum

As had been the case in the 1977 regulations, regulatory changes to the Rehabilitation Act in 1980 required that assessments used for college admissions constitute validated predictors of college aptitude or college achievement, rather than merely reflecting the applicant’s impaired sensory, manual, or speaking skills. Yet the 1980 regulations essentially permitted use of tests with established disproportionate adverse effects on SWDs, provided that an alternate test with a less disproportionate effect was unavailable. In 1980, when President Jimmy Carter established the Department of Education (ED) as a cabinet-level agency with a mission to ensure that educational opportunities were not denied on account of race, creed, color, national origin, or sex, disability was not included among the list of protected categories. Importantly, Section 103 (a) of the Department of Education Organization Act (PL 96-88) prohibited ED from exercising any control over the curriculum, or any program of instruction, or selection of instructional materials by any school system or educational institution.

Soon after the establishment of the new agency, Education Secretary Terrel Bell created a National Commission on Excellence in Education, which produced a report on the status of American education entitled 'A Nation at Risk,' which concluded that the country was threatened by a 'rising tide of mediocrity,' that over 10% of 17-year-olds were functionally illiterate, that SAT scores were declining across the country, and that many students required remediation courses even after entering college. The report concluded that comprehensive strategies to reform education across the country were needed (National Commission on Excellence in Education, 1983). It was believed that the needed reforms could be better comprehended after getting a fuller picture of student performance, by instituting the national assessment of all students (Ginsberg, Noell, & Plisko, 1988). During this decade and throughout the next, however, participation of SWDs in the national assessment (later, the National Assessment of Educational Progress or NAEP) was minimal (Shriner & Thurlow, 1993). In addition, most state assessment programs based their inclusion decisions for SWDs primarily upon those specified by the NAEP or on the basis of time spent in the regular classroom (Thurlow & Ysseldyke, 1993). Some factors that belied high exclusion rates on the NAEP cited by NCEO included unclear participation guidelines, sampling plans that systematically excluded students in separate schools or those not in graded classes, and an 'altruistic' motivation to reduce stress on students not expected to perform well (Zigmond & Kloo, 2009). Some states were simply unwilling to make accommodations to students to permit participation of SWDs in the NAEP (Ysseldyke & Thurlow, 1994a).

On the state assessment front, SWDs were included only slightly more often than on the NAEP. Shriner and Thurlow (1993) document that in the early 1990s, less than 10% of SWDs were being included in state assessments. As a consequence of widespread assessment exclusion policies for the first two decades after the establishment of the Office of Special Education Programs (OSEP), there was very little known about the academic outcomes of

SWDs (Ysseldyke, Thurlow, McGrew, & Shrine, 1994b).

The IDEA was reauthorized in 1990 as PL 101-476, and a focus remained on physical inclusion—greater inclusion in community schools, least restrictive placement of students, and transition services. Placements of SWDs in classes were to be age and grade appropriate with a minimum of placement in self-contained classrooms. Teaching methods for including SWDs in the general education classrooms began to involve cooperative learning and peer-instruction models. Yet an emphasis was placed on the physical inclusion of SWDs over the quality or effectiveness of their academic experience of SWDs (Danielson, personal communication, October 22, 2009). While teaching staff were expected to 'adapt' the curricular content, in doing so, they were encouraged to choose a grade-level curriculum that seemed developmentally most suited to meet each SWDs IEP objectives, rather than to ensure access to grade-level standards (Simon, Karasoff, & Smith, 1991).

IDEA 1990 also funded studies and investigations through which to collect information needed for program and system improvements by states and LEAs. The results of studies, such as the National Longitudinal Transition Study began to be available to OSEP prior to the 1997 authorization, and later shed light on the degree to which inclusion efforts were failing to ensure effective instruction of SWDs (Danielson, personal communication, October 22, 2009).

Prior to the 1993 ESEA reauthorization, Title I funds were to be distributed to schools on the basis of the poverty level and economic needs of students rather than on the basis of performance on State assessments. But the reauthorized 1993 ESEA shifted the focus to assessing outcomes for all children, including students with special needs, in key disciplines—mathematics, science, history, geography, civics, English, the arts, and other languages. The reauthorized ESEA attempted to ensure that 'all students,' including 'special needs' students, met high academic standards, that teaching and learning improved, that government offered flexibility coupled with responsibility for student

performance, that schools work cooperatively with parents and the community, and that Federal aid go to the poorest students (U.S. Department of Education, 1993).

ESEA 1993 endeavored to improve learning through reform approaches similar to those of other countries whose students were thought to be outperforming American students, particularly in the fields of science and mathematics. Thus, closely following the ESEA reauthorization of 1993 was the Goals 2000 Educate America Act (PL 103-227), which was signed into law on March 31, 1994. The essence of Goals 2000 was that by the year 2000 all students would leave grades 4, 8, and 12 with competency in English, mathematics, science, foreign languages, civics and government, economics, arts, history, and geography. Every student was to be prepared for responsible citizenship, postsecondary learning, and productive employment. The reforms of Goals 2000 were grounded in the expectation that States develop more challenging content and performance standards, design instruction and assessments aligned to those standards, and participate in accountability reporting on the extent to which schools and students were meeting the State standards (The White House, 1990; National Academy of Education, 1998). For some states, this was the first effort at trying to develop a broad framework for a general curriculum (National Academy of Education, 1998). Ultimately, under the ESEA Title I requirement, all states were expected to have valid, reliable, and aligned assessments based on their new content standards in the four core academic subjects by school year 2000–2001.

At the same time, among disability advocates, it was well understood that there was both an education gap as well as an ‘assessment gap’ for SWDs (Danielson, personal communication, October 22, 2009). The National Center on Educational Outcomes (NCEO) publicly posed the question of whether SWDs were seriously being considered in the standards-based reform movement, and pointed out that when identifying sources of data for monitoring progress toward the national goals, in 1991 the National

Education Goals Panel identified data collection programs that had excluded up to 50% of SWDs (McGrew et al., 1992). An NCEO Synthesis report ended with ‘Our nation in its quest to become first in the world has forgotten many of its students’ (Thurlow & Yesseldyke, 1993).

The Council for Exceptional Children testified to Congress in 1992 that the standards themselves should be constructed so as to accommodate all students, and it called for an investigation into alternative forms of assessments as well as ways to ensure that when educators worked on standards for assessments that at least one member be included who had expertise in working with individuals with disabilities (CEC testimony before House Subcommittee on Elem, Sec, and Vocational Education, 1992). By the time IDEA 1997 was reauthorized, most states had established content standards in the four core content areas, yet the question of which students with disabilities could access these standards, and participate in assessments based upon them, was a subject of debate. Moreover, the type of tests that was being constructed posed barriers over and above the content standards. States began moving away from flexible, or ‘authentic assessments,’ which held greater promise for inclusiveness of a range of student ability levels, in order to fulfill the pragmatic requirements of large-scale testing. Such ‘authentic’ assessments, popular during the era, were difficult to standardize across large numbers of diverse students. Moreover, while most special educators believed that performance-based assessments provided more accurate descriptions of student progress and were more helpful in informing classroom practice, their administration was expensive and overly time consuming for consideration in accountability testing.

In the midst of the standards movement, there was a pervasive concern that norm-referenced assessments were not appropriate to the goals of standards-based reforms, not just in the case of SWDs, but for all students. More importantly, norm-referenced tests were not well aligned to the curricula that students were to be taught

under the new standards movement. During the mid-1990s, there had been much concern about the fairness and the validity of norm-referenced scoring approaches for use with special populations. Many states also justified the exclusion of SWDs from standardized testing on the basis of fairness—that students had not received an opportunity to learn the material assessed on general assessments—and on the basis of utility—arguing that the results of assessment scores did not provide useful information about the academic performance or educational needs of SWDs. Advocates complained that the use of norm-referenced testing, in which SWDs were usually ranked lowest, led to the perpetuation of assumptions that SWDs were incapable of making any academic progress. Yet, excluding them from participation provided no information at all about their academic performance and, some argued, denied FAPE.

While many in the special education field were divided as to how SWDs should participate in the standards-based accountability movement of the 1980's, most later came to agree, as one state policymaker commented, that 'the removal of special education students from the "accountability track" resulted in their removal from the "curriculum track"' (Koehler, 1992).

Following Goals 2000, as most states embarked on a full-scale revision of their assessment systems and attempted to define 'what all students should know' in their new content standards, nearly all states shifted to the use of criterion-referenced assessments and began including a percentage of SWDs in these new assessments. In order to assist SWDs in accessing these new criterion-referenced assessments, States developed a list of 'standard accommodations.' The four classes of accommodations included the following: *presentation format*, which were changes in how tests were presented and involved accommodations like providing Braille versions of the tests or orally reading the directions to students; *response format*, which were changes in the manner in which students gave their responses and included accommodations such as having a student point

to a response or use a computer for responding; *setting of the test*, which could be alone, or in small groups; and finally, *timing of the test*, which could include extending the time allowed, or providing more breaks during testing.

In response to questions about the attainability of performance standards for all students with disabilities, ED advised states to implement alternative assessments for a 'small number' of students and to implement accommodations to ensure an 'equal playing-field' for those students, stating, 'Assessment accommodations help students show what they know without being placed at a disadvantage by their disability' (U.S. Department of Education, 1997). However, claims about the capacity of accommodations alone to overcome the disadvantages created by a student's disability were considered true for students with sensory, manual, or speaking skills, but not for SWDs with cognitive impairments.

The implementation of testing participation guidelines for SWDs was the subject of considerable controversy among disability advocates across states. Policy experts maintained, often based upon the Supreme Court ruling in *Southeastern Community College v. Davis* that standards could never not be lowered for SWDs taking accountability assessments, even if those assessments were also to be used to guide instruction (e.g., Philips, 2002). Yet most advocates maintained that, as students with disabilities were not included when the standards were developed, it seemed inappropriate to hold them to the standards.

Prior to the passage of the Americans with Disabilities Act of 1990 test developers were most familiar with the provision of accommodations for students with sensory impairments. However, following the passage of the ADA, advocates for the disabled argued that federal law should ensure the availability of testing accommodations and modifications for mental disabilities such as dyslexia and other learning disabilities. Yet policymakers responded again that the effect of accommodations for cognitive disabilities undermined the valid interpretation of a student's test score (e.g., Philips, 1993).

IDEA 97 and Options for Alternate Assessment

OSEP hoped, by providing funding opportunities, to spur the Research and Development Community to go beyond what had been considered technically feasible, to respond to the increasing demand for teaching tools and new approaches to the assessment of SWDs. The reauthorized IDEA of 1997 provided for the development of new assessments to both identify areas of academic need and to measure academic progress for children with disabilities. The new assessments were also to be used in educational program planning and for placement in special education, related services, and/or early intervention under § 641, (1)(G) of the law. Funds were to be made available for the development of alternative assessments for the inclusion of non-native English speakers and other minority students, to prevent misidentification of such students as SWDs. The mandates of IDEA 97 called for assessments to meet requirements that could not be entirely met by inclusion of SWDs in large-scale state assessments alone. Scientific measurement of student progress called for new types of classroom assessments to monitor academic progress (e.g., Fuchs & Fuchs, 2004), in response to intensive and evidence-based interventions. These new ‘curriculum-based’ assessments (CBM) were inherently inclusive, as they could be individualized for students working across a broad continuum of skill levels. The new CBM measures also represented a significant improvement from previous classroom assessments which relied on ‘mastery measurement,’ or summative, end-of-unit assessments, as the new measures permitted the monitoring of incremental progress over time by teachers (Stecker, 2005).

While progress-monitoring assessments permitted inclusiveness and informed decision-making in the classroom, some advocates maintained they had the disadvantage of isolating SWDs from the general education curriculum:

These assessments frequently were conducted in isolation from the larger general education curriculum. The assessments focused on immediate

and discrete skill deficits and IEPs often were a collection of isolated skill objectives that led to isolated instruction...Too often, the IEP *became* the curriculum for the student, instead of a tool for defining how to implement a general education curriculum. (Nolet & McLaughlin, 2000, p. 10)

Overall, the 1997 reauthorization was most significant for its endorsement of the participation of all children with disabilities in state assessments and for the requirement that alternate assessments be made available, by July 2000, for any of those students who could not meaningfully participate in the regular state assessment even with accommodations (U.S. Department of Education, 2000). The sanctioning of alternate assessments was to shift the views of federal and state policymakers on what constituted fair assessment practices by moving away from the principle that a single standard of performance on state standards, even for purposes of accountability, would by necessity apply to ‘all students.’

The states had concluded that, if all students were to be included in the new accountability systems, new assessments based on the standards would have to be developed for a small percentage of students with the most severe disabilities—generally students labeled with ‘severe-profound disabilities and trainable mentally handicapped’ (Quenemoen, 2009). According to Browder, & Wakeman, & Flowers (2009), prior to IDEA 97 there were three classes of SWDs: (a) those who pursued a general education curriculum with expectations for grade-level achievement, (b) those who required a remedial curriculum (e.g., a 7th grader working on 4th grade math), and (c) those who required functional life skills to prepare for independent living. While prior to 1997, teachers expected that only the first group of SWDs would participate in state assessments, after the 1997 reauthorization, the inclusion of SWDs in the new alternate assessments (Browder et al., 2009) constituted a major shift in assessment inclusion policies, permitting all subgroups of SWDs to participate with validity. The new alternate assessments were to be ‘aligned with the general curriculum standards set for all students and should not be assumed

appropriate only for those students with significant cognitive impairments' (34 CFR §200).

In spite of the inclusion accomplished by the 1997 IDEA, the disability community was torn as some advocates continued to maintain that exclusion from state assessments and substitution of measurements of progress toward IEP goals were the only appropriate responses to the new standards movement. Others contended that the substitution of IEP goals for state and National assessment participation would violate the spirit of inclusion—especially considering that IEP goals were often chosen from a list of useful skills of 'everyday life' (Browder et al., 2009), rather than being designed to provide equitable access to the full scope of the State content standards.

In response to the mandate for alternate assessments, states developed a variety of types of assessments which came in a variety of forms, including teacher checklists of functional skills, reports on progress toward IEP goals, portfolios of student work, and performance tasks moderately aligned to grade-level content standards (Thompson & Thurlow, 2000).

Significant policy input into the 1997 IDEA reauthorization came through David Hoppe, an aid to Senator Trent Lott, and a parent of a person with disabilities, who exercised a gentle touch in bringing diverse stakeholders to consensus. While the field grew to be unanimous in believing accountability and assessments based entirely on IEP rubrics did not make sense, at the same time, there continued to be debates about the validity of the scores of SWDs, especially those with cognitive impairments, who had not been exposed to the curriculum being assessed. Hoppe convinced policymakers to sidestep the partisanship in reauthorizing IDEA 97 and urged Congress to come up with a bill to please both sides of the debate (Danielson, personal communication, October 22, 2009).

The critical new elements in the 1997 IDEA amendments were accountability for inclusion of SWDs in general state and district-wide assessments, with appropriate accommodations and modifications, if necessary, and the establishment of performance goals for SWDs as a condition

of funding under IDEA Part B. In the infrequent cases when an IEP team or Section 504 team determined that standard assessments, even with reasonable accommodations, did not provide a student with an opportunity to demonstrate his or her knowledge and skill, the State or school district was to provide an alternate assessment. Yet whatever assessment approach was taken, the scores of students with disabilities were to be included in the assessment system for purposes of public reporting and school and district accountability.

The reauthorized IDEA 97 also required the consideration of assistive technology needs for participation, as well as the communication needs of children who were deaf, hard of hearing, or those with limited English language proficiency. A further requirement for assessments used for the evaluation of SWDs was the inclusion of information that was 'instructionally relevant' in the evaluation, in order to help a child become involved in and make progress in the general education curriculum. Assessment instruments used for disability evaluation were also required to be technically sound to assess the 'relative contributions' of both cognitive and behavioral factors, in addition to physical or developmental factors, on students' academic performance.

States were, once again, required to administer assessments to SWDs in a child's native language, or typical mode of communication. Importantly, any standardized tests given to a child were required to be validated for the specific purposes for which they were to be used. In addition, assessment tools and strategies that directly assisted teachers in determining the educational needs of the child were also to be provided under IDEA 97.

In response to the mandate for new alternate assessments, OSEP funded a variety of projects, such as supporting computer-adaptive assessments aligned to the state standards that would be capable of identifying, through 'dynamic' assessment techniques, learning issues of students with learning disabilities or other 'gap students' in order to uncover the instructional gaps they were manifesting in general education settings (e.g., see Tindal, 2008). It was known that

the population of students with specific learning disabilities (SLD) consisted of slow learners who could eventually address all content standards, though not necessarily in the time frame required to participate fairly in the summative end-of-year assessments. OSEP struggled with how to balance the learning needs of high-incidence SLD students with the mandate to include them in state assessments, as well as how to overcome the historical problem of low expectations. The answer OSEP arrived at to best address this was to mandate that instruction be provided by skilled teachers specifically trained to work with the SLD population. OSEP later funded the Access Center to help teachers adapt and individualize instruction aligned to standards that were appropriate for the student's grade and age-level, rather than 'out-of-level' standards, as had been a common teaching practice prior to 1997. Yet, to many advocates in the field, assessments based on standards that many SWDs could not master in the same time frame were also considered 'out-of-level' assessment, since such assessments required that a typical SLD student would need to make more than a year's worth of average progress in a year to learn enough grade-level material to be fairly assessed on the full scope of material being tested (Danielson, personal communication, October 22, 2009).

The effect of the 1997 IDEA was to shift reform efforts to the IEP, envisioning it not as a guide to what SWDs were to be learning, but rather rendering it into a tool to ensure inclusion and progress in the grade-level general education curriculum by defining each student's present level of performance, including how the student's disability affected his or her ability to be involved in and make progress in the general education curriculum. Additionally, the law required a statement in the IEP about the program modifications and supports to be used by school personnel to enable the child to be involved in and make progress in the general education curriculum and to participate with his or her non-disabled peers.

Subsequent to the IDEA Part B regulations in 1999, which mandated inclusion of all SWDs in standards-based reform programs, however, many SEAs did not succeed in ensuring that local

education agencies (LEAs) and schools taught SWDs the grade-level curriculum.

2001 No Child Left Behind Act

Under the 1994 ESEA, States were required to test only three times during a student's tenure in the K-12 educational system. For policymakers crafting the reauthorized ESEA, this left too many intervening years in which children's academic difficulties could go unaddressed, with the result that many children were being 'left behind,' academically. Under the 'No Child Left Behind Act' (NCLB) of 2001, States were obliged to enhance their existing assessment systems to include annual assessments in reading/language arts and mathematics for all public school students in grades 3 through 8 and at least once in grades 10 through 12 by the 2005–2006 school year. Additionally, by the 2007–2008 school year, all States were to annually assess their students in science at least once in grades 3 through 5, once in grades 6 through 9, and once in grades 10 through 12 (U.S. Department of Education, 2003).

The NCLB required annual testing in reading and mathematics, the demonstration of 'adequate yearly progress' against state-specified performance targets, and the inclusion of all students in annual assessments. Secretary of Education, Rod Paige, later succeeded by White House domestic policy advisor, Margaret Spellings, emphasized that the purpose of the NCLB provisions was to ensure that every child was learning 'on grade level.' The accountability for the SWD subgroup also required steps to recruit, hire, train, and retain highly qualified personnel, research-based teaching methods, and the creation of improvement programs to address local systems that fell short of performance goals.

During the same year, President Bush created the President's Commission on Excellence in Special Education, a program designed to improve the dropout rate among SWDs, who were leaving school at twice the rate of their peers, and whose enrollment in higher education

was 50% lower. Moreover, the SLD subgroup had grown over 300% since 1976, and 80% of those with SLD reportedly had never learned to read (President's Commission on Education, 2002). The claim was made that few children in special education were closing the achievement gap to a point where they could read and learn like their peers. A major thrust of the Commission was that although special education was based in civil rights and legal protections, most SWDs remained at risk of being left behind. Several findings of the Commission included criticisms that the reauthorized 1997 IDEA placed process above results and compliance above student achievement and outcomes. Further, Special Education did not appear to guarantee more effective instruction. The identification of students for special education services was criticized for being based upon a 'wait-to-fail' model. The criticism was launched that ED had become two separate systems instructionally, when it was critical that general education and special education share responsibilities for the education of SWDs. Among the recommendations of the report was a call for improved assessment policies to prevent exclusion from State and district-wide assessments, still a common practice in 2001.

2002–2003 Title I Regulations Permitting Alternate Achievement Standards in Accountability

The ESEA regulations of 2002 implementing the assessment provisions of NCLB authorized the use of alternate assessments in accountability assessments and required that States make available alternate assessments for any student unable to participate in the State's general assessments, even with accommodations. The subsequent ESEA regulations of 2003 permitted states to develop alternate achievement standards for students with the most significant cognitive disabilities. The 2003 regulations required that the alternate assessment be aligned with the State's academic content standards, promote access to the general curriculum, and reflect professional judgment of the highest achievement

standards possible (34 CFR§200.1). These regulations effectively forced most states to begin to revise the assessments they had originally created in response to the more liberal 1997 IDEA requirements. While the due date for the development of alternate assessments was 2000, there was little knowledge in the field of the appropriate academic content to base such tests on. While there had been some early work on how to teach general education curriculum content to students with severe disabilities (e.g., Downing & Demchak, 1996), the mandate for alternate academic achievement standards aligned to the state's academic content standards was to become a critical force in transforming the curriculum for SWDs with severe disabilities. Over the next decade, problems with developing a coherent academic curriculum appropriate to the diverse population of students with significant cognitive disabilities were prevalent. Among the most significant contributing factors in the delay in developing valid and aligned alternate assessments under the NCLB was the belief among special educators charged with developing the assessments that grade-level academic content standards were not relevant to these students, and that the appropriate content for these students consisted of 'life-skills' for independent living (Wallace, Ticha, & Gustafson, 2008). In response to this, ED set new standards for technical adequacy that alternate assessments were required to meet. Over the next decade, while the quality of many alternate assessments on alternate achievement standards (AA-AAs) improved, by 2010, many states still did not have technically adequate, peer-reviewed alternate assessments. At the same time, research carried out in the field indicated that students eligible for the AA-AAS could use symbolic communication systems and could learn to read and reason mathematically (Browder, Wakeman, Spooner, Ahlgrim-Delzell, & Algozzine, 2006; Kearns, Towles-Reeves, E., Kleinert, H. L., & Kleinter, 2009).

Students with significant cognitive disabilities continue to be inappropriately excluded from participation in alternate assessments, although one study of excluded students found that many could be using augmentative and assistive technologies

to speak or otherwise communicate, and that approximately 75% were learning sight words and using calculators to perform mathematical calculations (Kearns et al., 2009). Kearns recommends that to meaningfully include the population of students with significant cognitive impairment in state assessments, future assessments must include ‘authentic’ demonstrations of skills and knowledge aligned to the grade-level content standards, within the assessment context such students require. Scaffolding may be required for some within this population to enable them to show what they know and can do. The inclusiveness of alternate assessments for students with significant cognitive impairment depends largely upon whether or not the teacher has an understanding of how to teach a standards-based curriculum appropriate to the students within this diverse population (Kearns et al., 2009).

An additional problem with the implementation of alternate assessments for students with the most significant cognitive impairments has been the continued difficulty of establishing appropriately challenging achievement standards for the full range of ability levels manifested in this population. While states are permitted to set more than one achievement standard in order to make an alternate assessment more inclusive, teachers and schools have been cautious about assigning students to the more challenging achievement standard on a test used for making school accountability determinations—often assigning both higher- and lower-functioning students in the 1% population to the lowest achievement standard, to safeguard a favorable outcome in accountability under NCLB. As a result of these widespread practices, the proficiency rates on alternate assessments have been much higher compared to proficiency rates of SWDs taking the general assessment across the majority of states, suggesting that alternate assessments are simply not challenging enough for the majority of students taking them.

Currently available alternate assessments vary in the extent to which they inform parents and teachers about student academic progress. While advances in the development of classroom-based formative assessments, such as CBM, have been

widely used among students with high-incidence disabilities since IDEA 1997, comparable formative assessments have rarely if ever been available for teachers of students in the 1% population, though OSEP encouraged their development (but see Phillips et al., 2009). Assessment measures designed to measure a student’s level of mastery and progress on both prerequisite and grade-level content and skills at a challenging level for such students would greatly assist and support instruction of students with significant cognitive disabilities and help to demonstrate their capacity for progressing in an academic curriculum. Fundamentally, the central problem with alternate assessments has arisen from the need to understand and develop an appropriate academic curriculum for eligible students—one that is aligned to the same grade-level content standards intended for all students, yet reflects content and skills at an appropriate level of difficulty for each unique student, and is also capable of indicating progress toward an expected level of knowledge and skill.

IDEA 2004 and Assessments Measuring Responsiveness to Intervention

The reauthorization of the IDEA 2004 (PL 108-446) reiterated the NCLB mandate for inclusion of all SWDs in State and district-wide assessments, and clarified that IEP goals were not to be the only form of assessment. The most important assessment-related changes made in the IDEA 2004 reauthorization were those that pertained to requirements for eligibility determinations of SLD. These changes permitted states to move away from the IQ-performance discrepancy model of SLD identification that had been the subject of criticism in the Commission report. Under IDEA 2004, new assessments were to be developed for the purpose of assisting with the identification of students with SLD through the assessment of a student’s response to tiered, evidence-based instructional interventions (RTI). Such assessments were to be of several types, those to screen students in

basic skills (e.g., literacy or mathematics), and those to help define and monitor responsiveness to evidence-based interventions designed to help the child progress in significant areas of academic weakness and to inform a decision about the need for more intensive remedial instruction.

2007 Joint Title I IDEA Regulations Permitting Modified Academic Achievement Standards in Accountability

Nearing the end of the George W. Bush presidential term, Secretary of Education Margaret Spellings grew concerned at the slow pace with which states were advancing to the goal of ‘universal proficiency’ by 2014 intended by the NCLB. States and the special education community expressed concerns that a small group of SWDs who were enrolled in general education classes were, due to the nature of their disabilities, unable to demonstrate the same extent of academic progress on their state’s general assessment by the end of the school year. In 2006, the administration responded by announcing a new assessment policy option to provide states with the flexibility to include the scores of ‘persistently low-performing’ students with disabilities in alternate assessments based on modified academic achievement standards (AA-MAS). While states had struggled to improve teaching practices for a subgroup of low-performing students, both those with and without disabilities, it was widely accepted that some students, because of the effects of a disability, required additional time to learn content standards to mastery. State rationales for the new assessment included a need to develop academic achievement standards inclusive of this small group of SWDs, the majority of whom were enrolled in general education classes, yet had cognitive impairments such as mental retardation, autism, specific learning disability, or other health impairment for which accommodations alone could not ensure an ‘equal playing field.’

After the publication of the 2007 Joint IDEA Title I regulations, the field became split over

the potential consequences of permitting states to assess a portion of SWDs against a lower standard of performance. Concerns of some advocates in the disability community centered on the potential lowering of standards and educational tracking of students with SLD predominantly for the sake of making ‘Adequate Yearly Progress’ toward assessment performance targets specified under Title I accountability. States differed in the degree to which they included the SLD population in the new modified assessments—some argued that the SLD population in particular had been receiving below grade-level instruction and that the availability of such a test would unnecessarily lower academic expectations for these students, who, by definition, have normal or above average intellectual abilities. Others argued that problems manifested by this group of students in accessing general assessment items and their inability to master the expected content in the same time frame constituted the direct effects of their disabilities. They maintained that the new modified assessments, especially if used as a temporary measure, could help to illuminate the academic progress of this group of students, to help ensure that they received instruction aligned to grade-level content standards.

Subsequent to substantial investments by ED in the development of the AA-MAS, many states began investigating the population of SWDs who were persistently low achieving and conducted item accessibility and ‘cognitive lab’ studies to develop more accessible test items. Elliot et al., (2010) reported the results of a study in four states that indicated that certain modifications made to general test items improved the accessibility of the items and improved the accuracy of measurement for SWDs eligible for the AA-MAS. Modifications to regular test items did not change the content to be tested nor did they reduce the complexity of the test item (‘depth of knowledge’). The modifications made to test items were straightforward—including removing unnecessary words, simplification of language, addition of pictures and graphics, breaking long paragraphs into several, and bolding of key words. The study showed that an additional boost in student performance on the modified items

occurred when reading support (audio versions of text, except for the key vocabulary words being tested) was added to item modification. Results also showed a large boost in mathematics performance for the eligible group (Elliott et al., 2010).

‘Race To The Top’ Assessment Initiatives

In 2009, during the initial months of the Obama administration, a ‘new generation’ of standards and assessments was envisioned by which policymakers hoped all American students would become more competitive in the global marketplace. Many of the goals of the RTT assessment initiative reiterated those of the Goals 2000 era in this research. The RTT initiative endeavors to create assessments aligned to ‘fewer, higher, and clearer’ state standards held in common by most states, as well as becoming more inclusive and informative for students who typically perform at lower achievement standards, including SWDs.

In the words of Secretary Duncan,

The majority of students with disabilities take the regular state tests based on the state’s standards for all students, with appropriate accommodations to ensure that their results are valid. Students with the most significant cognitive disabilities can take alternate tests based on alternate standards and other students with disabilities may take an alternate test based on modified standards. (Duncan, 2010).

Our proposal would continue to hold schools accountable for teaching students with disabilities but will also reward them for increasing student learning. . . The Department plans to support consortia of states, who will design better assessments for the purposes of both measuring student growth and providing feedback to inform teaching and learning in the classroom. All students will benefit from these tests, but the tests are especially important for students with disabilities. (Duncan, 2010).

Policies supportive of inclusive assessments under the new administration promise to support the development of a new generation of assessments aligned to a range of achievement levels required for inclusion of all SWDs and to include measures of growth for students who make

progress at different rates. The history of inclusion in standards-based assessments shows that thoughtful inclusion decisions promote meaningful academic progress for SWDs. In the era of Goals 2000, 32% of SWDs graduated high school with a regular diploma, compared to nearly 60% of students in 2007. In 1987, only one in seven SWDs enrolled in postsecondary education, compared to a third of students in 2007 (Duncan, 2010). The challenge for growth-based accountability is to ensure that conceptions of growth are academically meaningful, in addition to being ‘statistically’ meaningful. To demonstrate consequential validity, the new assessments must provide results that specify clearly the knowledge and skills at each grade level that the student has attained and those requiring additional focus or individualized instructional intervention. Promising research and development since 2008 has resulted in empirically based assessment tools which can help to specify the accessibility of test items and to separate construct-relevant from irrelevant or extraneous cognitive processing demands of test items (Elliott, Kurz, Beddow, & Frey, 2009). To maximize inclusion of SWDs in the next generation of state assessments, a renewed commitment will be required by the disability community, to build upon the lessons learned over the history of inclusive assessment, and to methodically investigate and empirically establish expectations for their academic achievement and progress. Maximal inclusiveness in assessment will always be a function of successful inclusion in the general education curriculum, and of empirically-derived expectations for academic achievement and progress.

Acknowledgments The author gratefully acknowledges the contribution and support of Lou Danielson, who provided his unique historical perspective on assessment-related events during the first 40 years of OSEP.

References

- Browder, D., Wakeman, S., & Flowers, C. (2009). Which came first, the curriculum or the assessment? In W. D. Shafer & R. W. Lissitz (Eds.), *Alternate assessments based on alternate achievement standards: policy, practice, and potential*. Baltimore, MD: Paul H Brookes Publishing Co.

- Browder, D., Wakeman, S., Spooner, F., Ahlgrim-Delzell, L., & Algozzine, B. (2006). Research on reading instruction for individual students with significant cognitive disabilities. *Exceptional Children, 72*, 392–408.
- Brown v. Board of Education, 347 U.S. 483 (1954).
- Council for Exceptional Children. (1992). *Statement prepared for testimony to the house subcommittee on elementary, secondary, and vocational education*. Reston, VA: Author.
- Council for Exceptional Children. (1992). *Statement prepared for testimony to the technical advisory panel on uniform national rules for NAEP testing for students with disabilities*. Reston, VA: Author.
- Downing, J.E., & Demchak, M. (1996). First steps: Determining individual abilities and how best to support students. In J. E. Downing (Ed.), *Including students with severe and multiple disabilities in typical classrooms: Practical strategies for teachers* (pp. 35–61). Baltimore, MD: Paul H. Brookes Publishing Co.
- Duncan, A. (2010). *Keeping the promise to all America's children*. Remarks made to the council for exceptional children, April 21, Arlington, VA, www.ed.gov
- Elementary and Secondary Education Act of 1965, PL 89-10 (1965).
- Elliott, S. N., Kurz, A., Beddow, P., & Frey, J. (2009, February 24). *Cognitive load theory: Instruction-based research with applications for designing tests*. Paper presented at the annual convention of the national association of school psychologists, Boston.
- Elliott, S. N., Kettler, R. J., Beddow, P. A., Kurz, A., Compton, E., McGrath, D., et al. (2010). Effects of using modified items to test students with persistent academic difficulties. *Exceptional Children, 76*(4), 475–495.
- Fuchs, L. S., & Fuchs, D. (2004). Determining adequate yearly progress from kindergarten through grade 6 with curriculum-based measurement. *Assessment for Effective Instruction*.
- Ginsberg, A. L., Noell, J., & Plisko, V. W. (1988). Lessons from the wall chart. *Educational Evaluation and Policy Analysis, 10*(1), 1–10.
- Hehir, T. (2005). *New directions in special education: Eliminating Ableism in policy and practice*. Cambridge, MA: Harvard Education Press.
- Kearns, J. F., Towles-Reeves, E., Kleinert, H. L., & Kleinter, J. (2009). Who are the children who take alternate achievement standards assessments? In W. D. Schafer & R. W. Lissitz (Eds.), *Alternate assessments based on alternate achievement standards: Policy, practice, and potential*. Baltimore, MD: Paul H Brookes Publishing Co.
- Koehler, P. D. (1992). Inclusion and adaptation in assessment of special needs students in Arizona. In M. L. Thurlow & J. E. Yesseldyke (Eds.) (1993), *Can "all" ever really mean "all" in defining and assessing student outcomes?* (Synthesis Report No. 5). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.
- McGrew, K. S., Thurlow, M. L., Shriner, J. G., & Spiegel, A. N. (1992). *Inclusion of students with disabilities in national and state data-collection programs* (Technical Report 2). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.
- National Academy of Education. (1998). *Goals 2000: Reforming education to improve student achievement*. Washington, DC: National Academy of Education.
- National Commission on Excellence in Education. (1983). *A Nation at risk: The imperative for educational reform*. Washington, DC: Government Printing Office.
- Nolet, V., & McLaughlin, M. J. (2000). *Assessing the general curriculum: Including students with disabilities in standards-based reform*. Thousand Oaks: Corwin Press.
- Phillips, S. E. (1993). Testing accommodations for disabled students. *Education Law Reports, 80*(9).
- Phillips, S. E. (1994). High stakes testing accommodations: Validity versus disabled rights. *Applied Measurement in Education, 7*(2), 93–120. Lawrence Erlbaum, Associates, Inc.
- Phillips, S. E. (2002). Legal issues affecting special populations in large-scale assessment programs. In G. Tindal & Thomas M. Haladyna (Eds.), *Large-scale assessment programs for all students: Validity, technical adequacy, and implementation*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Phillips, G. W., Danielson, L., & Wright, L. (2009). *Psychometric advances in alternate assessments*. Washington, DC: American Institute for Research.
- President's Commission on Excellence in Special Education. (2002). *A new Era: Revitalizing special education for children and their families*. Washington, DC: Author
- Quenemoen, R. (2009). The long and winding road of alternate assessments: Where we started, where we are now, and the road ahead. In William D. Schafer & Robert W. Lissitz (Eds.), *Alternate assessments based on alternate achievement standards: Policy, practice, and potential*. Baltimore, MD: Paul H Brookes Publishing Co.
- Russo, C., & Osborne, A. (Eds.). (2008). *Essential concepts and school-based cases in special education law*. Thousand Oaks: Corwin Press.
- Shriner, J., & Thurlow, M. L. (1993). *State special education outcomes: A report on state Activities at the end of the century*. Minneapolis, MN: National Center on Educational Outcomes, University of Minnesota.
- Simon, M., Karasoff, P., & Smith, A. (1991). *Effective practices for inclusion programs: A technical assistance planning guide*. Unpublished paper supported by U.S. Department of Education Cooperative Agreements #GOO87C3056-91 and #GOO87C3058-91.
- Skirtic, T. M. (1991). The special education paradox: Equity as the way to excellence. *Harvard Educational Review, 61*(2), 148–206.
- Southeastern Community College v. Davis, 442 U.S. 397, 413 (1979).
- Stecker, P. M. (2005). Monitoring student progress in individualized educational programs using curriculum-based measurement. In U.S. Department of Education, Office of Special Education: IDEAS that Work:

- Toolkit on teaching and assessing students with disabilities. National Center on Student Progress Monitoring.
- The White House. (1990). *National educational goals, office of the press secretary*. Washington, DC: Author.
- Thompson, S. J., & Thurlow, M. L. (2000). *State alternate assessments: Status as IDEA alternate assessment requirements take effect* (Synthesis Report 35). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.
- Thurlow, M. L., & Ysseldyke, J. E. (1993). *Can "all" ever really mean "all" in defining and assessing student outcomes?* (Synthesis Report No.5). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.
- Tindal, G. (2009). Reflections on the alternate assessment in Oregon. In W. D. Shafer, & R. W. Lissitz (Eds.), *Alternate assessments based on alternate achievement standards: Policy, practice, and potential*. Baltimore, MD: Paul H Brookes Publishing Co.
- U.S. Department of Education. (1993). *The reauthorization of the elementary and secondary education act, executive summary*. Washington, DC: U.S. Department of Education.
- U.S. Department of Education. (1997). *Elementary and secondary education. Guidance on standards, assessments and accountability*. Retrieved October 2, 2009 from http://www2.ed.gov/policy/elsec/guid/standardsassessment/guidance_pg4.html#disabilities3
- U.S. Department of Education. (1998). *Goals 2000: Reforming education to improve student achievement, April 30, 1998*. Washington, DC: U.S. Department of Education.
- U.S. Department of Education. (2000). *Summary guidance on the inclusion requirement for Title I final assessments*. Retrieved March 13, 2010 from <http://www2.ed.gov/policy/elsec/guid/inclusion.html>
- U.S. Department of Education. (2003). *Standards and assessments non-regulatory guidance March 10, 2003*. Washington, DC: Author.
- Wallace, T., Ticha, R., & Gustafson, K. (2008). *Study of general outcome measurement (GOMs) in reading for students with significant cognitive disabilities: Year 1* (RIPM Technical Report #27). Minneapolis, MN: Research Institute on Progress Monitoring, University of Minnesota.
- Ysseldyke, J., & Thurlow, M. L. (1994a). *Guidelines for inclusion of students with disabilities in large-Scale assessments* (Policy Directions # 1). Minneapolis: National Center on Educational Outcomes, University of Minnesota.
- Ysseldyke, J., Thurlow, M. L., McGrew, K. S., & Shriner, G. J. (1994b). *Recommendations for making decisions about the participation of students with disabilities in statewide assessment programs: A report on a working conference to develop guidelines for statewide assessments and students with disabilities* (Synthesis Report 15). Minneapolis, MN: National Center on Educational Outcomes.
- Zigmond, N., & Kloof, A. (2009). The 'two percent students': Considerations and consequences of eligibility decisions. *Peabody Journal of Education*, 84(4), 478–495.



<http://www.springer.com/978-1-4419-9355-7>

Handbook of Accessible Achievement Tests for All
Students

Bridging the Gaps Between Research, Practice, and
Policy

Elliott, S.N.; Kettler, R.J.; Beddow, P.A.; Kurz, A. (Eds.)

2011, XIV, 341 p., Hardcover

ISBN: 978-1-4419-9355-7