

Assessing the Modern Urban School System: The Institutionalization of Standardized Testing in Philadelphia, 1925–30

IN 1926, TEACHERS at South Philadelphia High School for Girls faced a problem. Some students were underperforming in their coursework and scoring low on standardized tests. By contemporary measures, educators feared these children would become a future drag on society. Anna Biddle, a South Philadelphia High teacher, pessimistically observed, “Such girls certainly have no place in any four-year high school course,” but the students believed that public education was their best means to secure stable employment, particularly, she noted, “in an office.” Impressed by the students’ stated aspirations, Biddle led a corps of teachers to develop a program for the girls that would take them away from the rest of the student population to receive instruction about the “routine[s] . . . the ideal business girl must know.” The instructors doubted their chances of success, but rationalized that “the state always spends more money on its incompetents than on any others and a small sum spent for prevention can be looked upon as an investment. These girls may become social problems; just now, however, they are teaching problems.”¹

South Philadelphia’s experiment eventually confirmed for Biddle the value of sorting students through standardized testing, validating the belief that testing promoted instructional, administrative, and social efficiency without wasting resources on students perceived to be unteachable. The general school population, for example, “benefit[ed] from having a large number of the less competent pupils isolated from the regular classes.” Students in the training program, meanwhile, were salvaged as the “habit

The author gratefully acknowledges the constructive comments and feedback on earlier versions of this essay from numerous friends and colleagues and from the peer reviewers and editors of the *Pennsylvania Magazine of History and Biography*.

¹ Anna E. Biddle, “Low IQ’s in the High School,” *School Review* 35 (1927): 134–46.

of success" replaced their "habit of failure." The students "blossomed out when they were away from the inhibiting affects of girls with superior ability"; as a result of this intervention, she reported, "their self-respect is restored." Further, society avoided being saddled with the cost of tending to the "social problems" of the underachievers and the unemployable. Businesses had a workforce trained to "contribute toward success in certain non-intellectual occupations." Such remarkable gains compelled Biddle to recommend increased research to identify students' abilities to direct them "into their proper sphere" while simultaneously sparing students "a tremendous amount of disappointment, time, and misspent energy."²

In their attempts to solve perceived problems of school and society, teachers' characterizations of South Philadelphia High's students suggest they held little hope for the girls' success. Such attitudes are not surprising, as teachers in the 1920s had normalized and internalized the rhetorical jargon of education, psychology, and other social sciences of the post-World War I era. Faith in the efficiency of standardized testing and data collection administered and interpreted by technical experts characterized this period. Standardized tests in particular worked well within urban school systems such as Philadelphia's, whose bureaucratic structures provided a "grammar of schooling" that an emerging cadre of educationalists used to maintain order while legitimizing their professional authority.³

Edwin C. Broome, Philadelphia's superintendent of schools, was among this burgeoning group of professionals. Broome was a late convert to standardized student assessments. He had proclaimed his initial skepticism during a national educational conference at the University of Pennsylvania in 1921, stating, "It seems to me it is an open question as to the use of, and the extent to which we shall accept, the results of various educational and intelligence tests," and concluding, "I am not sure as to the extent to which these tests can be safely applied, or the safety with which we can use the

² Ibid. Biddle's experiment likely benefitted from South Philadelphia High's recent introduction of the experimental Dalton Plan. As the school's principal explained, "the fundamental principles of the Dalton Plan, as we at the South Philadelphia High School for Girls interpret it, are: first, individualized instruction, but in a socialized environment, permitting each child to work to capacity, cooperatively, in spite of the individual differences, of which nowadays we are so intensely conscious; and, second, freedom, but with stabilizing responsibility, permitting each child to reach his goal at his own speed and in his own time." Lucy L. W. Wilson, "Experiments in Adolescent Training," *Survey*, June 15, 1926, 368–70.

³ David B. Tyack and Larry Cuban, *Tinkering toward Utopia: A Century of Public School Reform* (Cambridge, MA, 1995); David B. Tyack and Elisabeth Hansot, *Managers of Virtue: Public School Leadership in America, 1820–1980* (New York, 1982), esp. 105–14.

results of such tests to determine the future of the child.”⁴ Four years later, however, Broome authorized the formation of the Division of Educational Research and Results for Philadelphia’s public school system, declaring the need for a staff of trained experts and educational professionals to “collect accurate data [and to conduct] scientific studies of all phases of educational procedure,” including the assessing and grouping of students through widespread standardized testing.⁵

Viewed together, the vocational program at South Philadelphia High School for Girls and Broome’s changed stance toward standardized tests suggests the degree to which professional educators during the early twentieth century believed that such tests could efficiently and effectively sort students within properly organized bureaucratic school systems. The implementation of this belief, however, often had negative consequences. The historiography of standardized testing demonstrates, for example, how the use of scientific assessments resulted in a social hierarchy at the expense of immigrants, blacks, and other marginalized groups.⁶ Yet such explorations generally emphasize the role of intelligence tests and educators’ misuse of these tests in determining individual and group intelligence quotients, or IQ scores. What these analyses tend to overlook is the development of massive testing programs on the local level more generally and how standardized tests came to dominate as they did. In Philadelphia, educators’ and administrators’ rapid institutionalization of a testing program within a brief number of years reveals a desire to create an orderly educational system befitting a modern city and its people.

This essay examines the development of Philadelphia’s testing regimen and the individuals behind it, exploring their rationales for imposing

⁴ Edwin C. Broome, “Address,” in *Annual Schoolmen’s Week Proceedings*, vol. 8 (Philadelphia, 1921), 31–32.

⁵ Edwin C. Broome, “Report of the Superintendent of Schools,” in *Annual Report of the Board of Public Education* (Philadelphia, 1926), 273.

⁶ The literature on this is quite extensive. See, for example, Stephen Jay Gould, *The Mismeasure of Man* (New York, 1981); Clarence J. Karier, “Testing for Order and Control in the Corporate Liberal State,” in *Roots of Crisis: American Education in the Twentieth Century*, ed. Clarence J. Karier, Paul C. Violas, and Joel Spring (Chicago, 1973); Nicholas Lemann, *The Big Test: The Secret History of the American Meritocracy* (New York, 2000); Leon J. Kamin, *The Science and Politics of IQ* (Potomac, MD, 1974); Michael M. Sokal, ed., *Psychological Testing and American Society, 1890–1930* (New Brunswick, NJ, 1987); Mark Snyderman and Stanley Rothman, *The IQ Controversy, The Media, and Public Policy* (New Brunswick, NJ, 1988); Paul David Chapman, *Schools as Sorters: Lewis M. Terman, Applied Psychology, and the Intelligence Testing Movement, 1890–1930* (New York, 1988); and Henry L. Minton, *Lewis M. Terman: Pioneer in Psychological Testing* (New York, 1988). See also Michael M. Sokal’s assessment of this historiography in “Approaches to the History of Psychological Testing,” *History of Education Quarterly* 24 (1984): 419–30.

those testing structures on schools and classroom teachers from the mid to late 1920s. Educational administrators during the early twentieth century equated efficiency with modernization, and the introduction of widespread standardized testing provided them a means to organize big-city school systems based on these principles. Widespread standardized testing during the 1920s may have symbolized a modern urban educational system, but the development and implementation of massive, systemic testing regimens ultimately became ends unto themselves, with the most tangible result of Philadelphia's program being its size and scope. Philadelphia school leaders used the power of the standardized testing program to establish a modern school system—one based on increasing both the numbers of tests administered and the numbers of students tested.⁷ The introduction of massive testing programs at the district level established patterns of educational assessment that would endure in big-city school systems through the remainder of the twentieth century and beyond.⁸

The Structures and Staffing of a Modern Urban School System

Pennsylvania's educators of the early twentieth century saw themselves as heirs to a proud Quaker tradition, championed by men such as Benjamin Franklin and Benjamin Rush, of providing free public education. By the early 1880s, Pennsylvania was considered a leader in expanding public education, and business and community leaders looked to a large public school system to protect the established social hierarchy from the per-

⁷ In developing my idea of the power of standardized testing programs as they developed in Philadelphia, I rely on Ian Hacking's notion of "statistical enthusiasm" borne from "the numerical manipulation of the body politic." See Ian Hacking, "Biopower and the Avalanche of Printed Numbers," *Humanities in Society* 5 (1982): 279–95. I also draw from Joseph Tropea's construction of "backstage organizational order [based on] backstage understandings and rules [that] allowed administrators and teachers, and eventually staff, to respond to many social, legal, and economic vicissitudes while preserving organizational, if not pedagogical, integrity." See Joseph L. Tropea, "Bureaucratic Order and Special Children: Urban Schools, 1890s–1940s," *History of Education Quarterly* 27 (1987): 29–53; and "Bureaucratic Order and Special Children: Urban Schools, 1950s–1960s," *History of Education Quarterly* 27 (1987): 339–61.

⁸ The standardized testing movement considered here formed the basis of today's so-called high-stakes testing of individual students and state and local school districts. Nationally institutionalized in the No Child Left Behind Act, such testing regimens likely will continue to dominate educational reforms for years to come. See, for example, Lemann, *Big Test*; Linda M. McNeil, *Contradictions of School Reform: Educational Costs of Standardized Testing* (New York, 2000); Sam Dillon, "Obama to Seek Sweeping Change in 'No Child' Law," *New York Times*, Feb. 1, 2010, A1. See also W. James Popham, *The Truth about Testing: An Educator's Call to Action* (Alexandria, VA, 2001).

ceived threats presented by an uninformed, indigent, and immigrant citizenry. Yet, by the early 1920s, the results of modernization efforts, such as compulsory student attendance or expanded bureaucratic administration, had plateaued. As the University of Pennsylvania's Frank Graves reported to the Pennsylvania State Educational Association (PSEA) at its annual meeting in Harrisburg, the state ranked near the bottom of the *Index of State School Systems*, "below all, save the Southern and a few of the newer states." Optimism prevailed, however, as Graves inveighed his colleagues to "go forward to new victories and greater achievements than Pennsylvania has yet known. The heights are there for us; let us emerge from the plain and capture them."⁹ The sentiment of progress in Pennsylvania's public schools persisted throughout the decade. In his address to the PSEA in Philadelphia three years later, then governor Gifford Pinchot declared a centralized bureaucracy as the priority in moving Pennsylvania's schools into the future. The success of this bureaucracy rested on "a sound advancing modern plan with full provision for meeting the changing needs of the [educational] situation." In Pinchot's assessment, "Pennsylvania has made progress in public education—real progress—but we cannot let it go at that. We are not yet at the head of the states in our common schools. That is where we belong, and before we are through that is where we are going to be."¹⁰

Such sentiments manifested in Philadelphia schools during the early twentieth century in a belief in social improvement *through* education—that is, that schools could fundamentally alter society by addressing public concerns about health, safety, and welfare in America's growing urban centers. As Philadelphia school superintendents of the early 1900s proclaimed, the "influences of the school reaches up through the children into the home."¹¹ In this way, schooling improved children's home life and enhanced the quality of life for all city residents.

⁹ Walter Licht, *Getting Work: Philadelphia, 1840–1950* (Philadelphia, 1999), 60, 65; Sam Bass Warner Jr., *The Private City: Philadelphia in Three Periods of Its Growth* (Philadelphia, 1986), 111, 123; William H. Issel, "Modernization in Philadelphia School Reform, 1882–1905," *Pennsylvania Magazine of History and Biography* 94 (1970): 358–83; Frank P. Graves, "Educational Pioneers of Pennsylvania," *School and Society* 13, no. 317 (1921): 91–97.

¹⁰ Gifford Pinchot, "The Schools of Pennsylvania," *School and Society* 14, no. 473 (1924): 53–57.

¹¹ Lawrence A. Cremin, *The Transformation of the School: Progressivism in American Education, 1846–1957* (New York, 1962), viii, 85–88; Joel Spring, "Education as a Form of Social Control," in Karier, Violas, and Spring, *Roots of Crisis*, 30–33; Frank V. Thompson, *Schooling of the Immigrant* (1920; repr. Montclair, NJ, 1971); Martin G. Brumbaugh, *Report of the Superintendent of Schools* (Philadelphia, 1912), 9–10; John P. Garber, *Report of the Superintendent of Schools* (Philadelphia, 1916), 26.

The increasing responsibilities of public schools necessitated special departments within a central administrative bureaucracy staffed by an emerging cadre of educational professionals. As such, Philadelphia's system maintained divisions of Compulsory Education, Special Education, Physical Education, Medical Inspection and Nursing Services, Teacher Training, Practical Arts and Vocational Training, Commercial Education, and School Extension Programs to meet its aims. Each division required properly educated and credentialed personnel, drawn from the expanding pool of trained professionals from newly organized schools of education.¹² These professionals and their peers throughout the nation eagerly applied their acquired knowledge of schools and society to solve the problems of an increasingly chaotic urban and industrial society.

School administrators modernized and expanded Philadelphia's system in hopes of socializing the poor and immigrant classes into the city's industrial economy, a goal other leading citizens shared. The Americanization Committee of the Philadelphia Chamber of Commerce, for example, warned: "Loyal and patriotic Philadelphians should have cause for real concern" that only half of the city's foreign-born population was naturalized and only one-third of the other half was proceeding toward US citizenship. "The educational adjustment of the newcomer is plainly the problem of the public school," the committee asserted; as such, the schools needed to "remove illiteracy and all other un-American tendencies." Only then could a foreigner realize his "usefulness [and] economic value."¹³ Educators shared the business community's anxieties regarding increasing foreign populations. School administrators complained that children of foreign-born parents were present "in sufficient numbers to show the magnitude of the problem [of being] inmates of non-English speaking homes [which practiced] only the most meager and imperfect conceptions of American manners and customs." Foreign-born parents and children

¹² David B. Tyack, *The One Best System: A History of American Urban Education* (Cambridge, MA, 1974), 129–32, 182–98; Arthur G. Powell, *The Uncertain Profession: Harvard and the Search for Educational Authority* (Cambridge, MA, 1974), 52–83; Edwin C. Broome, "Report of the Superintendent," in *Annual Report of the Board of Public Education* (Philadelphia, 1922), 36–45; Broome, "Report of the Superintendent," in *Annual Report of the Board of Public Education* (Philadelphia, 1925), 196; Broome, "How Philadelphia Is Solving Its Educational Problems," *Nation's Schools*, Jan. 1930, 26–30; Broome, "Philadelphia's Big Six," *Journal of Education*, Feb. 17, 1930, 186; Robert Wiebe, *The Search for Order* (New York, 1966), 132, 145–49.

¹³ Americanization Committee of the Philadelphia Chamber of Commerce, *Americanization in Philadelphia: A City-wide Plan of Co-ordinated Agencies* (Philadelphia, 1923), 1.

posed “a serious menace to the welfare of our city and state . . . isolated [in their] own colonies . . . as if they really lived in their native lands.”¹⁴

Philadelphia educators matched their worries about poor foreigners with fears about African Americans. In 1920, black students represented 8 percent of the total number of enrolled pupils; by 1930, approximately 14 percent of all students were African American.¹⁵ Elementary school principal Philip A. Boyer articulated educators’ fears when he claimed that “the negro immigrant, like the foreigner, is likely on his arrival in the city to settle first in the congested slum district where housing is poor, tenements are unsanitary, and the general social environment is conducive to ill-health, immorality, and crime.” At times, Boyer observed, “the better negroes move out to the more thinly settled negro sections,” but the practice of taking in lodgers disrupted the home life and denied the family “the opportunity for building up those home interests so essential to the proper development of the negro.” According to Boyer, black migrants’ settlement in concentrated areas of the city created “unsanitary housing, low wages, high rents, lodgers, working mothers, and children left to care for themselves.” Such conditions “disrupt[ed] the recent and only partially organized family life of the negro.” “The members of such families,” he warned, “mingle in the larger social life of the street with its baneful influences”—among them, “Morbidly exciting movies [which] combine their potent influence with that of the street to turn thoughts toward immorality and crime.” Further, migrants “as a whole are woefully ignorant and disrespectful of laws of health,” weakening their “vitality and efficiency,” leading to irregular work and school attendance, “habits of shiftlessness,” and a lowered “moral tone” within the entire black community. Not until the newcomer “has been trained in the exercise of proper health habits,” asserted Boyer, “can we expect to note any great increase in efficiency.”¹⁶

¹⁴ Garber, *Report of the Superintendent of Schools* (1916), 24–26. Much of educators’ anxieties over large numbers of foreigners were unfounded and uninformed. Although parents of school-aged children across the city largely were foreign born in 1916, more than 93 percent of children enrolled in Philadelphia public schools were born in the United States. That figure was more than 98 percent by 1930. See Henry J. Gideon, “Report of the Division of Compulsory Education,” in *Annual Report of the Board of Public Education* (Philadelphia, 1916), 235; Gideon, “Report of the Division of Compulsory Education,” in *Annual Report of the Board of Public Education* (Philadelphia, 1930), 322.

¹⁵ Vincent P. Franklin, *The Education of Black Philadelphia: The Social and Educational History of a Minority Community* (Philadelphia, 1979), 50.

¹⁶ Philip A. Boyer, “The Adjustment of a School to Individual and Community Needs” (PhD diss., University of Pennsylvania, 1920), 24–25, 28–29, 33–34.

Many of Philadelphia's school officials advocated administrative and curricular solutions to the presumed problems of poor, migrant, and African American populations. Educators introduced programs of "[English] language, arithmetic, geography and history [for the] unwashed and unkempt children" and evening classes for adults that dealt "in simple language with matters of sanitation and hygiene [and] the elements of local government and good citizenship." School administrators believed every citizen of Philadelphia needed "to discharge the ordinary duties of life [by knowing] how to speak and read the English language correctly and with facility, to write a legible hand, and be able to apply the rules of arithmetic" in order to secure "positions in the industrial organizations of the community." Schools were responsible for training children to contribute socially and economically to the city by teaching them how "to conform to [local] community regulations rather than . . . municipal and state and national rules of government." Educational institutions that did otherwise, warned Superintendent Edward Brooks, were "not measuring up to the demands of public education."¹⁷

School administrators increasingly found standardized tests to be not only a yardstick by which to measure whether these educational demands were being met but also a means of diagnosing educational problems. As early as the 1870s, for example, psychological scientists such as Francis Galton and James Cattell used standardized tests to varying degrees in Europe and the United States to identify mental problems in children. Beginning approximately in 1908—when Alfred Binet and Théodore Simon developed an "intelligence scale"—and continuing through the World War I era, psychologists demonstrated standardized tests' value to modern society. The Alpha and Beta army tests administered to military recruits demonstrated trained professionals' ability to conduct large-scale testing and provided a means to sort individuals into an established order. The war's end ushered in a new, distinct phase of the testing movement's development as psychologists persuaded education leaders that they could achieve maximum productivity and efficiency by placing students in an educational and social hierarchy using standardized tests.¹⁸

¹⁷ Edward Brooks, *Report of the Superintendent of Schools* (Philadelphia, 1902), 3–4, 121–23; Brooks, *Report of the Superintendent of Schools* (Philadelphia, 1903), 65; Martin G. Brumbaugh, *Report of the Superintendent of Schools* (Philadelphia, 1911), 13–14, 27; Brumbaugh, *Report of the Superintendent of Schools* (1912), 8–9, 11; John P. Garber, *Report of the Superintendent of Schools* (Philadelphia, 1915), 24–25.

¹⁸ Cremin, *Transformation of the School*, 185–89; Chapman, *Schools as Sorters*, 6, 17, 20, 32–34; Minton, *Lewis M. Terman*, 52, 72, 74–76.

By the mid-1920s, educationalists proclaimed that the greatest strengths of standardized tests lay less in measuring students' intellects than in measuring students' abilities to achieve academically. Educators thus aligned tests to specific curricular objectives and established standards of student achievement in particular academic areas. Educators subsequently determined students' knowledge and abilities in those areas by measuring their progress in reaching established standards, ultimately using those results to identify what they believed to be efficient classroom practices in improving instructional methods for specific subjects.¹⁹

Several factors contributed to standardized tests' development, acceptance, and widespread use by educators during the first quarter of the twentieth century. First, psychologists wanted to establish their profession's legitimacy by defining and measuring specific and general abilities of large segments of the general population. Further, testing helped educators legitimize their own professionalism, enabling them to differentiate and categorize growing student populations. Lastly, the sentiments of educational and social reformers of the day—particularly faith in science and trust in academic experts—encouraged the use of tests as a way to improve classroom instruction and school administration. In this context, psychologists needed school superintendents, administrators, and teachers as much as educators needed psychologists. These mutual interests gave educators and psychologists the opportunity to prove their value to schools and society and a reason to use standardized tests in public schools.²⁰

Two of the more influential psychologists in the development and promotion of testing in public schools during the post-World War I era were Edward Thorndike and Lewis Terman. Each man believed that utilizing quantifiable psychology in the schools could improve efficiency, yet they differed on the means of doing so. Thorndike believed improving educational efficiency with the science of psychology would make schools more vital institutions. Psychology informed his advocacy for efficiency

¹⁹ Alexander C. Roberts, "Measuring and Testing in Education," *Journal of the National Education Association* 13, no. 1 (1924): 101; Guy M. Wilson and Kremer J. Hoke, *How to Measure* (1920; rev. ed., New York, 1928), 5; William A. McCall, "Place of Measurement in Education," in *How to Measure in Education* (New York, 1923), 3–18; Virgil E. Dickson, "The Test Controversy," *Journal of the National Education Association* 12, no. 5 (1923): 176; A. R. Gilliland and R. H. Jordan, *Educational Measurements and the Classroom Teacher* (New York, 1925), 25–26, 29–37; William A. McCall and Harold H. Bixler, *How to Classify Pupils* (New York, 1928), 1.

²⁰ Chapman, *Schools as Sorters*, 4–5, 17–18, 39–43; Philip Boyer, "Educational Tests and Measurements: Statistical Treatment of Test Results," *Bulletin of the Division of Educational Research* 68 (Feb. 1928): 5.

in education in two ways. First, psychology could facilitate learning by making teaching methods conform to children's natures. Second, scientists could study the results of these new ways of teaching and evaluate the efficiency of specific teaching methods. Psychology's goal of exploring aspects of human nature that had previously been unknown or considered unimportant could further educators' objectives.²¹ Psychology could improve teaching by clarifying educators' objectives and identifying and measuring the desired student behaviors to be developed through the use of particular teaching methods.

Differences among individuals undergirded Thorndike's views on psychology's utility in the schools. "We may study a human being in respect to his common humanity, or in respect to his individuality," Thorndike wrote, concluding, "In other words, we may study the features of intellect and character which are common to all men, to man as a species; or we may study the differences in intellect and character which distinguish individual men."²² Thorndike acknowledged that large-scale testing was a means of ascertaining these differences:

The superintendents, supervisors, principals and teachers directly in charge of educational affairs have been so appreciative of educational measurements and so sincere in their desire to have tests and scales devised which they can themselves apply, that the tendency at present is very strong to provide means of measurement which are concerned somewhat closely with school achievements, and which can be used by teachers and others with little technical training.

Nevertheless, Thorndike cautioned against the "real danger in sacrificing soundness and principle and precision of result to the demand that we measure matters of importance and measure them without requiring elaborate technique or much time of the measurer." After all, he pointed out: "The danger is that the attention of investigators will be distracted from the problems of pure measurement for measurements sake, which are a chief source of progress in measuring anything."²³

²¹ Geraldine M. Joncich, "Science: Touchstone for a New Age in Education," in *Psychology and the Science of Education: Selected Writings of Edward L. Thorndike*, ed. Geraldine M. Joncich (New York, 1962), 6, 8–9.

²² Edward Thorndike, "Individuality," in Joncich, *Psychology and the Science of Education*, 119.

²³ Edward Thorndike, "The Nature, Purpose, and General Methods of Measurements of Educational Products," in Joncich, *Psychology and the Science of Education*, 154–55.

Lewis Terman's promotion of the use of tests in public schools complemented Thorndike's ideas. Objective testing helped educators provide appropriate instruction for children, for, as Terman argued, "it is time that the school should ask not only what it would like to do, but what it can do for a given pupil." Terman claimed that standardized tests were "an indispensable aid" to educators in diagnosing educational problems "for the simple reason that these problems cannot be dissociated from the quality of material with which the school works."²⁴ Terman thus advanced the idea of schools as factories, using raw materials to produce commodities befitting an industrialized society. Industries, including public schools, needed to employ scientific methods to improve the manufacturing of their products. Terman repeated this theme when he described how testing "subjected . . . the material with which the school works . . . to the same cold analysis as the products of farm, factory, or mine. Nothing is taken for granted, everything must be proved. The spirit of educational research rules the day." Terman's faith in that "spirit" was unwavering. Believing the triumph of science over the problems of school and society to be inevitable, Terman expressed confidence "that the opponents of the scientific movement in education [would not] be able seriously to retard its progress." As he saw it, "There is every likelihood that such opponents of the inevitable will lose whatever opportunities they might have had to shape the course of modern educational currents."²⁵

Proving Administrative Efficacy through Standardized Testing

In the spirit of educational research, the post–World War I era witnessed a proliferation in the publication and dissemination of nationally standardized tests for both intelligence and achievement. An examination of the kinds of tests deployed at the local level reveals how educators often used a variety of tests to ascertain students' abilities to recall information or perform certain tasks or to determine their intelligence. Ease of use and ease of interpretation of the results often were valued more than any other aspect of the tests.²⁶

²⁴ Lewis M. Terman, "The Use of Intelligence Tests in the Grading of School Children," *Journal of Educational Research* 1 (1920): 30–32.

²⁵ Lewis M. Terman, "Research and the Problems of Educational Readjustment," *Journal of Educational Research* 1 (1920): 138–39.

²⁶ The following descriptions of assessments come from Wilson and Hoke, *How to Measure*, and Walter S. Monroe, James C. DeVoss, and Fredrick J. Kelly, *Educational Tests and Measurements* (1917; rev. ed., Boston, 1924). Considered one of the "pioneering" texts on the subject (Roberts, "Measuring

Intelligence assessments saw their heaviest use in early grades, as teachers attempted to evaluate the youngest of students. The Detroit Tests of Intelligence for kindergarteners and first graders, for example, were designed to be completed in “seven to twelve minutes . . . by the average teacher with a little practice and careful study of the directions.” The Pressy Intermediate Classification Test required third graders to complete 96 separate tasks and was “simply constructed so that it can be easily applied by the teacher.” The Haggerty Intelligence Test for second and third graders contained 12 questions that assessed students’ ability to take and follow directions, copy designs, complete partially drawn pictures, draw pictures freehand, and work with “simple digits,” all within thirty minutes. Instructors of ninth through twelfth graders, meanwhile, encountered the “simplicity of the mechanics” of the Terman Group Tests of Mental Ability; “the definiteness of the instructions for giving them make it possible for any teacher with a small amount of study” to accurately administer and interpret these tests. Ninth through twelfth graders could also complete the Otis Classification Tests. Results from the 115 questions about history and civics, grammar, physiology and hygiene, geography, music, and art gave classroom teachers “a fairly accurate index” of students’ mental ability after the thirty minutes allotted for completion.²⁷

While knowledge of predetermined levels of hygiene presumably indicated students’ intelligence, reading competencies supposedly indicated students’ academic abilities. The Thorndike-McCall Reading Scale assessed reading comprehension of third through twelfth graders. Students silently read a series of “isolated” paragraphs and then answered questions based on each individual selection. The test was “simple in its nature so that any teacher [could] apply it with accuracy.” Further, the content of the Thorndike-McCall was “fairly representative of reading in general,” though what determined “reading in general” remained vague. Teachers

and Testing in Education,” 101), *How to Measure* attempted to evaluate the strengths and weaknesses of various achievement tests “which on account of their use, purpose, and adaptability have been found to be most serviceable to the classroom teacher” (Wilson and Hoke, *How to Measure*, v). *Educational Tests and Measurements* provided descriptions “to enable [the teacher] to choose wisely in selecting a test” (Monroe, DeVoss, and Kelly, *Educational Tests and Measurements*, vii). Both texts generally found assessments that promoted uniform methods of instruction, required minimum levels of teacher expertise, and evaluated minimum levels of student competencies to be among the best available.

²⁷ Philip A. Boyer, “Report of the Division of Educational Research and Results,” in *Annual Report of the Board of Public Education* (Philadelphia, 1926), 490; Boyer, “Report of the Division of Educational Research and Results,” in *Annual Report of the Board of Public Education* (Philadelphia, 1930), 528; Wilson and Hoke, *How to Measure*, 229, 328–29, 333, 342–43, 348–49.

used the Thorndike-McCall because it offered “a method in the direction” of classroom instruction and suggested to individual teachers a “selection of materials” to be used on a regular basis.²⁸ The Monroe Silent Reading Tests evaluated comprehension and reading speed of ninth and tenth graders. As with the Thorndike-McCall, students read paragraphs selected “from school readers and the books which children read.” After each selection, students underlined one word from a provided list that best described the meaning of each paragraph. The assessment contained a range of difficulty and variations of materials from “prose, poetry, narration and description,” but teachers could easily administer the test in a short amount of time.²⁹ Finally, the Haggerty Reading Exam for seventh graders and ninth through tenth graders served as three tests in one. Tests of vocabulary asked students to underline the best definition of words in questions—for example: “Minister (Servant, Preacher, Agent, To Assist).” The sentence reading tests asked simple but value-laden “yes-no” questions such as “Can good children make promises?” Tests of paragraph reading asked students a question related to a selected passage:

Underline the one phrase which tells what Rip did not like to do
Run errands
Work at home
To hunt
To fish

Based on student responses to these kinds of questions, teachers determined students’ abilities and were encouraged to select “suitable reading material” within the scope of those abilities.³⁰

The advocacy of scientific testing by Thorndike, Terman, and others, however, only partially explains the proliferation of national tests and the testing movement’s triumph in public schools at the local level. Testing programs in urban school systems such as Philadelphia’s gained further momentum from the school survey movement. Between 1910 and 1925, hundreds of state and local boards of education and school superintendents commissioned educational experts—usually university professors,

²⁸ Wilson and Hoke, *How to Measure*, 128–29; Monroe, DeVoss, and Kelly, *Educational Tests and Measurements*, 118–21.

²⁹ Wilson and Hoke, *How to Measure*, 132–34; Monroe, DeVoss, and Kelly, *Educational Tests and Measurements*, 99–102.

³⁰ Wilson and Hoke, *How to Measure*, 137, 139–40; Monroe, DeVoss, and Kelly, *Educational Tests and Measurements*, 118.

state education authorities, and administrators from either other public school systems or newly formed educational research bureaus—to assess the numerous features of public school systems, including teaching methods, employee salaries, quality of building structures, and student achievement.³¹ Philadelphia was no exception. As early as 1917, city leaders publicly promoted the need for a comprehensive survey to identify and solve problems of instructional inefficiency and financial waste in the schools. Almost simultaneously, state legislators empowered the recently reorganized Pennsylvania Department of Public Instruction to undertake steps for statewide educational improvements with the legal authority to initiate surveys and the personnel to conduct them. Years of negotiations between Philadelphia's Board of Education, community leaders, and the state agency resulted in State Superintendent for Public Instruction Thomas Finegan initiating a survey of the city's schools in May 1920. State agents concluded their work in March 1922.³² Among the surveyors' final recommendations was the call for further testing of students as a means of promoting greater efficiency in school organization and administration.

The *Report of the Survey of the Public Schools of Philadelphia* contains several noteworthy aspects. First, it revealed Philadelphia educators' views about school efficiency and organization. Surveyors discovered that some individual principals had employed standardized tests, but no significant systemwide effort "to classify pupils according to ability" existed. State officials concluded that without a "systematic attempt," the administration of tests was inefficient and wasteful, lacking "organization and direction." For greater efficiency and productivity in the schools, state agents recommended the "scientific classification of pupils"—sorting students early and often. "The principle of classification according to ability [should] be adopted at once in Philadelphia," state authorities urged, and "should begin in the first grade" through any of the publicly available "group" tests

³¹ Chapman, *Schools as Sorters*, 35–37; Tyack, *One Best System*, 191–96; Powell, *Uncertain Profession*, 84–107; Leonard P. Ayers, "School Surveys," *School and Society* 1, no. 17 (1915): 577–81. The influence of Edward Thorndike over the school survey movement is worth noting. As Geraldine Joncich observes, "Leading figures in the surveys, like George D. Strayer of Teachers College and Ellwood P. Cubberley of Stanford University, received their statistical training and their faith in the power of quantification in Thorndike's courses in educational measurement." See Joncich, "Science," 15.

³² "Address of Honorable Franklin Spencer Edmonds," in *Report of the Survey of the Public Schools of Philadelphia*, 4 vols. (Philadelphia, 1922), 1:5–10; "Address by Thomas E. Finegan," in *ibid.*, 1:11–30; *Journal of the Board of Public Education, School District of Philadelphia* (1918): 57; *Journal of the Board of Public Education* . . . (1920): 93; William Rowen, "Report of the President," in *Annual Report of the Board of Public Education* (Philadelphia, 1921), 12.

of intelligence.” Thereafter, schools were to group students homogenously and “readjust [these groupings] throughout the entire course” of a student’s career. “Progressive” schools found such classifying and continuous adjustment to be helpful “for the sake of better teaching and the greater retention of pupils,” state agents declared.³³ Further, the report demonstrated educators’ views on the role of schools in society. State authorities, echoing Lewis Terman, spoke of efficiency of instruction, organization of administration, and schools taking in children “much as they are—bright, average, dull, quick or slow, energetic or apathetic” and properly training them according to their needs.³⁴ Finally and most importantly, the report initiated the proliferation of educational testing in Philadelphia public schools by recommending classification of students based on test results. Although local educational authorities did not begin “at once,” as state agents advocated, widespread standardized testing of Philadelphia students did begin in 1925 with the creation of the Division of Educational Research and Results.³⁵

*Broome, Boyer, and Philadelphia’s
Division of Educational Research and Results*

The two men responsible for the development of the Division of Educational Research and Results were part of a second wave of administrative progressives—educational careerists who advocated a “new educational order” of bureaucratic efficiency while working to legitimize their own professional authority.³⁶ The standardized testing program that Edwin Broome and Philip Boyer implemented in Philadelphia during the 1920s through the new division culminated educational modernization efforts begun earlier in the century and placed administrative control of schools in the hands of what historian Walter Issel characterizes as an “efficiency-minded upper-class [of] university-trained, educational experts.”³⁷

³³ Pennsylvania Department of Public Instruction, *Report of the Survey of the Public Schools of Philadelphia*, 2:246–47, 285–89.

³⁴ *Ibid.*, 2:287.

³⁵ *Journal of the Board of Public Education* . . . (1925): 227; “Establish School Research Bureau,” *Philadelphia Evening Bulletin*, Sept. 8, 1925, “Philip A. Boyer” envelope, and “Dr. Boyer Named to \$5000 Post,” *Philadelphia Public Ledger*, Sept. 9, 1925, George D. McDowell Philadelphia Evening Bulletin Newsclipping Collection, Special Collections Research Center, Temple University Libraries, Philadelphia (hereafter Evening Bulletin Clipping Collection).

³⁶ Tyack and Hansot, *Managers of Virtue*, 94–129.

³⁷ Issel, “Modernization in Philadelphia School Reform,” 381–83. See also Robert H. Weibe, “The Social Functions of Public Education,” *American Quarterly* 21 (1969): 147–64.

As superintendent of schools, Edwin Broome oversaw the creation of the new bureau. Broome viewed as self-evident the reasons for the division's existence. It aided the superintendent while benefiting students and the community. Comparing the city's schools to military, industrial, and business organizations, Broome proclaimed that Philadelphia's educational system "serves the public and must anticipate social and economic changes and prepare for them." To do so, the school system needed a "trained agency to make constant and scientific studies of all phases of educational procedure" so the superintendent could both know "at all times [the] present tendencies" of students and "anticipate and clearly formulate future needs" of the students and the schools. Educational Research and Results was to accomplish this mission by collecting statistics and data to guide the superintendent in creating educational policy; preparing that information for publication and public dissemination; conducting standardized tests throughout the school system to improve teacher instruction; studying the "classification and promotion of pupils"; recommending models of school organization that "affect the efficiency of instruction"; and continually re-viewing "the work of the schools."³⁸

Broome appointed Philip Albert Boyer head of the new division. Boyer epitomized the educational professional of the early twentieth century and, as such, was particularly qualified for his new position. As Boyer rose through the ranks of Philadelphia public schools as student, teacher, and administrator, he solidified his beliefs in efficiently organized urban school systems based on student assessment through extensive standardized testing.

Boyer graduated from Philadelphia's prestigious Central High School in 1903 and the Philadelphia School of Pedagogy two years later.³⁹ Between

³⁸ Ralph D. Owen and LeRoy A. King, "Volume II: Central Administrative Organization, Finance and School Business, Educational Research and Results," *Philadelphia Public School Survey* (Philadelphia, 1937), 233–34; William Rowen, "Report of the President," in *Annual Report of the Board of Public Education* (1925), 35; Edwin C. Broome, "Report of the Superintendent of Schools," (1926), 273, 295; Philip A. Boyer, "Educational Measurements: The Contributions of Educational Research to Teaching Practices," in *Annual Schoolmen's Week Proceedings*, vol. 13 (Philadelphia, 1926), 371; "Establish School Research Bureau" and "Dr. Boyer Named to \$5000 Post," Evening Bulletin Clipping Collection.

³⁹ Philip Albert Boyer file, box 241, Office of Alumni Records Biographical Records, 1750–2002, UPF 1.9 AR, University Archives, University of Pennsylvania, Philadelphia (hereafter Boyer Alumni File). Throughout much of its history, Central High School's promotion of an educational meritocracy through a rigorous entrance examination and a tradition of classical instruction often clashed with reformers' attempts to consolidate educational programs, including vocational instruction, in a comprehensive high school. The tensions at Central between vocational and classical curriculums undoubtedly influenced Boyer, who, as a public school administrator, rejected written exams in favor of

1905 and 1914, he taught at various public schools across the city while continuing his professional training, earning his bachelor's degree in social sciences from Temple University in 1912. He was a principal at different schools for the next eleven years while furthering his professional development, earning his master's degree in sociology and economics in 1915 and his PhD in education in 1920, all from the University of Pennsylvania. In sociology courses such as "Social Debtor Classes," "American Race Problems," "American Criminology," and "Eugenics and the Family," Boyer studied groups and individuals whom theorists believed contributed to society little else than crime, vice, and other social problems. Boyer combined this learning with the theories of pedagogy and efficient organization he studied in courses like "Educational Research" and "School Administration." Boyer's resultant ideas were that urban schools could and should categorize students in order to reach "maximum efficiency" in classroom instruction and pupil advancement. He believed that "scientific management has entered the educational field" and that homogeneously grouping school children increased student promotion rates, saved school systems tens of thousands of dollars annually, and spared the individual pupil "the loss . . . in confidence in his own ability to achieve."⁴⁰

Boyer asserted in his dissertation that the "doctrine of efficiency in industry . . . has direct bearing upon the organization and administration of schools."⁴¹ To demonstrate this, Boyer applied scientific management principles to two predominantly black elementary schools in one of Philadelphia's poorer neighborhoods, arguing that efficiently managed educational programs—characterized by a rigorous testing program and cooperation between schools, homes, and community service organizations—could reverse the affects of slum life and "do much to strengthen the influence of the school and the effectiveness of its work." Further, Boyer theorized:

standardized objective assessments and advocated the grouping of students within schools. For more on the history of Central, see David F. Labaree, *The Making of an American High School: The Credentials Market and the Central High School of Philadelphia, 1838–1939* (New Haven, CT, 1988).

⁴⁰ Boyer Alumni File; Philip Albert Boyer record sheet, box 9, Graduate School Record Sheets, Sept. 30, 1913, to June 17, 1914, Graduate School of Arts and Sciences Student Records, 1896–1982, UPB 7.62, University Archives, University of Pennsylvania; "Dr. P. A. Boyer, Educator, 85, Dies in Hospital," *Philadelphia Evening Bulletin*, Sept. 21, 1971, in Evening Bulletin Clipping Collection; "Assoc. Supt. Retires," *School News and Views* 4, no. 10 (1952), in Evening Bulletin Clipping Collection; *Bulletin of the Graduate School of the University of Pennsylvania*, 1913 to 1920, passim, University Archives, University of Pennsylvania; Philip A. Boyer, "Class Size and School Progress," *Psychological Clinic* 8 (1914): 82–90.

⁴¹ Boyer, "Adjustment of a School to Individual and Community Needs," 13.

The adjustments above indicated, culminating in a wholesome, vigorous school spirit, based upon a [student's] thorough appreciation of the ideals of the school and a willingness to cooperate in their achievement, would result in a strength of character and fixedness of purpose so necessary for sound individual progress, especially for those pupils who by reason of their race are destined to be harassed by many obstacles.⁴²

Boyer believed schools needed to promote social reforms and conformity by “develop[ing] in each individual, the knowledge, habits and attitudes that should be possessed in common by all members of society” by inculcating “unsanitary [and] immoral” newcomers to the city with a faith in education. Those who required such schooling needed to share this belief if they were to contribute socially and economically to the urban society. Students who failed to adopt these values, Boyer asserted, bred social and economic disorder.⁴³

Appointed the head of Philadelphia's Division of Educational Research and Results in 1925, Boyer made widespread student assessment the bureau's top priority. Adopting the language of leading educators, Boyer promulgated that standardized testing served multiple purposes efficiently and affordably. Tests established minimum standards of academic attainment against which educators could evaluate their students. Students grouped according to their test results then could reach educational standards more easily than children in heterogeneous groups, because the so-called slower or mentally inferior individuals did not hold back the more capable students. Moreover, tests measured students' proficiencies in particular subject areas and “improved” classroom instruction by indicating to teachers what topics needed review and which students needed additional attention. Finally, individual subject tests aided in “educational guidance” of students for the myriad employment opportunities in Philadelphia's growing business sector by indicating what commercial or industrial areas best suited students' futures.⁴⁴

⁴² Ibid., 106.

⁴³ Ibid., 13–14. Boyer's analysis of black Philadelphians lacked any acknowledgment of how an increasingly segregated city and the beginnings of a two-tiered public educational system based on race contributed to the conditions that concerned him the most. See Franklin, “Politics, the Public Schools, and the Black Community in the 1920s,” in *Education of Black Philadelphia*, 60–86, for an exploration of these conditions.

⁴⁴ Boyer Alumni File; *Journal of the Board of Public Education* . . . (1925): 227; “Establish School Research Bureau,” and “Dr. Boyer Named to \$5000 Post,” Evening Bulletin Clipping Collection; Philip A. Boyer, “Report of the Division of Educational Research and Results,” *Reports of Special Divisions of the Department of Instruction* (Philadelphia, 1934), 32; Boyer, “Report of the Division of Educational

Boyer advocated the use of simple objective testing strategies that emphasized students' informational learning and recalling of factual knowledge acquired through classroom instruction. Boyer favored true-false, one-word answer, sentence completions, and multiple-choice tests over written exams that required students' critical thinking and analytical skills. Written examination, according to Boyer, wasted excessive amounts of time as students completed the exams and as teachers graded them. Further, Boyer believed written exams often focused too narrowly on individual topics or subject areas. Objective tests, on the other hand, were easily available, directly related to individual class content, and provided a "definite check on class and individual progress . . . promptly and economically." Boyer also believed that consistent use of achievement tests raised teacher professionalism and that regular administration and interpretation of assessments improved classroom pedagogy. Boyer used every opportunity to promote the public image of teachers as educational experts capable of diagnosing definite causes of student success and failure through standardized testing.⁴⁵

The testing program Boyer oversaw was extensive (Table 1). By the end of its first year of operation, the Division of Educational Research and Results administered thousands of tests within two categories. Some were of the widely available variety authored and published by nationally leading psychologists such as Thorndike, Terman, and their contemporaries; the division's own personnel developed others. Within the first five years under Boyer's leadership, the division quickly increased the number of tests administered and students evaluated. While Educational Research and Results depended heavily on nationally standardized tests between 1925 and 1928, bureau officials introduced more tests of their own design beginning in 1927, increasing the number of division-designed tests every

Research and Results," in *Annual Report of the Board of Public Education* (Philadelphia, 1925), 560–562; Boyer, "Educational Research and the Commercial Teacher," in *Annual Schoolmen's Week Proceedings*, vol. 17 (Philadelphia, 1930), 392–99; Boyer, *Report of the Division of Educational Research and Results* (Philadelphia, 1931), 13; Harriet M. Barthelmess and Philip A. Boyer, "An Evaluation of Ability Grouping," *Journal of Educational Research* 26 (1932): 284–94. See also Thompson, *Schooling of the Immigrant*, 220–38.

⁴⁵ Boyer, "Educational Measurements," 360–70, 374; J. Crosby Chapman, "Home-made Objective Examinations for Everyday Use," in *Annual Schoolmen's Week Proceedings*, vol. 12 (Philadelphia, 1925), 291–96; Monroe, DeVoss, and Kelly, *Educational Tests and Measurements*, 1–11, 487; Gilliland and Jordan, *Educational Measurements and the Classroom Teacher*, 8–15; Wilson and Hoke, *How to Measure*, 6; McCall, *How to Measure in Education*, v–vi; Dickson, "Test Controversy," 176.

Table 1: Nationally standardized and division-designed tests administered by the Division of Educational Research and Results, October 1925 to June 1930^a

<i>October 1925 to June 1926</i>	<i>Number of Tests Used</i>	<i>% Within Total</i>	<i>Number of Student Tests Administered</i>	<i>Total Student Population</i>	<i>Number of Tests per Student^b</i>
Nationally Standardized Tests	52	78.79	258,159	230,596	1.1
Division-Designed Tests	14	21.21	738,526	230,596	3.2
Total	66	100.00	996,685	230,596	4.3
<i>September 1926 to June 1927</i>					
Nationally Standardized Tests	51	83.61	542,080	232,455	2.3
Division-Designed Tests	10	16.39	1,038,886	232,455	4.5
Total	61	100.00	1,580,966	232,455	6.8
<i>September 1927 to June 1928</i>					
Nationally Standardized Tests	65	69.89	310,045	234,257	1.3
Division-Designed Tests	28	30.11	1,283,347	234,257	5.5
Total	93	100.00	1,593,392	234,257	6.8
<i>September 1928 to June 1929</i>					
Nationally Standardized Tests	60	65.93	205,687	233,689	0.8
Division-Designed Tests	31	34.07	986,001	233,689	4.2
Total	91	100.00	1,191,688	233,689	5.1
<i>September 1929 to June 1930</i>					
Nationally Standardized Tests	46	41.44	101,502	234,861	0.4
Division-Designed Tests	65	58.56	1,500,072	234,861	6.4
Total	111	100.00	1,601,574	234,861	6.8

a. Author's calculations based on the number of tests and the number of students tested, reported in "Report of the Division of Educational Research and Results," in *Annual Report of the Board of Public Education* (Philadelphia, 1926–30), passim; and the total student population reported in "Report of the Division of Compulsory Education," in *ibid.* b. Rounded to the nearest tenth.

year thereafter until in-house-developed assessments were the majority by 1929. This shift toward increased use of division-designed tests reflected educators' belief in emphasizing particular subject matters and instructional methods but also limited the number of subject areas classroom teachers evaluated.

A main feature of Boyer's testing program was educators' increasing use of locally created tests to evaluate the greatest number of students. Between 1925 and 1926, for example, nearly three times as many school-children were assessed with division-designed tests as those students evaluated with nationally standardized tests. By 1930, however, nearly fifteen times as many students were assessed with division-designed tests as those evaluated with nationally standardized tests. The percentage of students who took nationally standardized tests during that time (with many students taking more than one test) ranged from approximately 7 percent to 137 percent while the percentage of the student population who took division-designed tests during that time ranged from approximately 35 percent to 220 percent (see Tables 2 and 3). During the 1929–30 school year alone, the division administered 1.5 million student assessments using 65 division-designed tests. The practice of using a limited number of division-designed tests to assess the greatest number of students multiple times per year casts doubt on the assessments' validity and reliability. While by most indications, educators used standardized assessments for the recommended grade levels, many of the tests were applicable to multiple grades simultaneously. It is reasonable to expect, therefore, that students could see the same tests in the seventh grade as they saw in the fifth or sixth grades.

A closer examination of the kinds of tests administered by the Division of Educational Research and Results reveals the bureau's increased utilization of tests assessing individual students' abilities in particular subject areas (Table 2). The division promoted the use of subject-area tests because such assessments were simple to administer, relied on uniform methods of instruction, and needed minimum levels of competencies for students' successful completion. Despite Boyer's pronouncements of teachers developing into educational experts through consistent use of standardized testing, the research bureau promoted objective assessments for individual instructors' use in a variety of classroom settings without requiring high levels of expertise for interpretation. Minimum skill levels and knowledge characterized nationally standardized achievement tests such as the

Table 2: Top three nationally standardized tests administered by the Division of Educational Research and Results by year, October 1925 to June 1930^a

<i>October 1925 to June 1926</i>	<i>Number of Tests Used</i>	<i>Grades Tested</i>	<i>Number of Tests Administered (% of Student Population Tested)^b</i>
Reading	9	7, 9, 3–12	133,702 (57.98)
Intelligence	10	Kgn, ^c 1–2, 2–3, 9–12, 3–6	38,807 (16.83)
Form Tests ^d	17	7–9, 7–12	37,005 (16.05)
<i>September 1926 to June 1927</i>			
Reading	9	9, 3–12	318,439 (136.99)
English Form Tests	18	7–9, 9–12, 8–9	63,954 (27.51)
Intelligence	11	Kgn, 1, 2–3, 1–2, 3–6, 9–12	59,837 (25.74)
<i>September 1927 to June 1928</i>			
Reading	8	1–3, 9, 3–6, 7–12, 3–12	203,339 (86.80)
English Form Tests	28	7–9, 9, 9–12, 7–12, 10	73,485 (31.37)
Intelligence	10	Kgn, 1, 2–3, 5–8, 4–6, 1–2, 3–6, 9–12	25,348 (10.82)
<i>September 1928 to June 1929</i>			
Reading	7	1–4, 9, 3–6, 7–12, 12	157,680 (67.47)
English Form Tests	24	7–9, 9, 9–12	17,955 (7.68)
Arithmetic	1	4–7	15,951 (6.83)
<i>September 1929 to June 1930</i>			
Reading	7	1–2, 3–6, 4–5, 7–9, 9–10, 7–12, 12	54,304 (23.12)
Intelligence	16	Kgn, 1, 2–3, 4–6, 4–8, 5–9, 6–8, 9, 9–12, 13	21,517 (9.16)
Form Tests	7	9–12	19,320 (8.23)

a. Author's calculations based on the number of national standardized tests, the number of students tested, and the grades tested, reported in "Report of the Division of Educational Research and Results," *Annual Report of the Board of Public Education* (Philadelphia, 1926–30), passim. From this information is derived the three most tested subject areas based on the percent of students tested from the total school population reported in Table 1 above. b. Rounded to nearest hundredth. c. Kindergarten. d. English-language assessments of grammar, vocabulary, and punctuation.

Table 3: Top three division-designed subject area tests administered by the Division of Educational Research and Results by year, October 1925 to June 1930^a

<i>October 1925 to June 1926</i>	<i>Number of Tests Used</i>	<i>Grades Tested</i>	<i>Number of Tests Administered (% of Student Population Tested)^b</i>
Arithmetic	9	3, 4, 5, 6, 7, 8	321,181 (139.28)
Spelling	2	1–8, 2–8	288,983 (125.32)
Handwriting	1	2–8	80,000 (34.69)
<i>September 1926 to June 1927</i>			
Spelling	1	2–8	429,920 (184.95)
Arithmetic	2	4–5, 6–7	308,011 (132.50)
Handwriting	1	2–8	218,000 (93.78)
<i>September 1927 to June 1928</i>			
Spelling	2	2–8	459,075 (195.97)
Arithmetic	9	3, 4, 5, 4–5, 6, 6–7	302,851 (129.28)
Handwriting	2	2–8	269,360 (114.98)
<i>September 1928 to June 1929</i>			
Arithmetic	11	2, 3, 4, 5, 6, 7	434,591 (185.97)
Spelling	1	2, 8	221,964 (94.98)
English	6	5, 6, 7, 8, 9	116,660 (49.92)
<i>September 1929 to June 1930</i>			
Handwriting	1	2–8	515,735 (219.59)
Arithmetic	32	2, 3, 4, 5, 6, 7, 8	493,806 (210.25)
Spelling	1	2, 8	141,735 (60.35)

a. Author's calculations based on the number of division-designed tests and the grades tested, reported in "Report of the Division of Educational Research and Results," in *Annual Report of the Board of Public Education* (Philadelphia, 1926–30), passim. From this information is derived the three most tested subject areas based on students tested from the total school population reported in Table 1 above.

b. Rounded to nearest hundredth.

Thorndike-McCall Reading Scale, the Monroe Silent Reading Tests, and the Haggerty Reading Exam.⁴⁶

While creating an extensive standardized testing program based on student achievement, Boyer simultaneously worked to decrease the bureau's reliance on nationally standardized tests by developing assessments within the research division. Two reasons guided this effort. First, the school district spent less money reproducing in-house tests than buying commercially available tests. More importantly, Boyer believed tests developed in-house were more valid and reliable than nationally standardized tests to evaluate curricular content and instructional methods specific to Philadelphia schools. The division staff concentrated much of its efforts during its first years on establishing the "validity and reliability" of these tests on the student population of Philadelphia schools.⁴⁷ Contrary to claims of improving Philadelphia-specific curricula and instruction, division-designed tests assessed minimum skill competencies within a few subjects. Division personnel, for example, designed evaluations for students in the lower elementary and junior high school grades and for pupil achievement in arithmetic, spelling, and handwriting (Table 3). Arithmetic assessments focused on mathematical computation of simple numbers, asking students to solve problems of addition, subtraction, multiplication, and division. Tests of arithmetic reasoning were slightly more sophisticated, yet narrowly focused, requiring students to apply mathematical functions to short word problems, such as:

Before Albert went to the country, he weighted $82\frac{3}{4}$ pounds, but when he came back after the vacation his weight was $96\frac{1}{2}$ pounds. How much had he gained in weight?

Frank's mother gave him a five-dollar bill to buy 5 lbs of nuts at \$.35 a pound and a bag of flour for \$.65. What change should he return to her?

Mr. Brown is offered a 5-gallon can of auto oil for \$4.25. If he buys it by the quart he must pay \$.30 for each quart. How much does he save by buying the 5-gallon can?

⁴⁶ Boyer, "Report of the Division of Educational Research and Results," in *Annual Report of the Board of Public Education* (Philadelphia, 1926), 490; Boyer, "Report of the Division of Educational Research and Results," in *Annual Report of the Board of Public Education* (Philadelphia, 1930), 528.

⁴⁷ Boyer, "Report of the Division of Educational Research and Results" (1926), 492–93; Boyer, "Report of the Division of Educational Research and Results," in *Annual Report of the Board of Public Education* (Philadelphia, 1934), 32; George A. Works, "Volume I: Summary of Findings and Recommendations," in *Philadelphia Public School Survey* (Philadelphia, 1937), 50.

Students received credit for the number of questions they attempted compared to the number of correct answers and for their demonstrated reasoning.⁴⁸

In addition to illustrating educationalists' beliefs in using standardized measures in public schools, Philadelphia's expanding testing program demonstrated a narrowing emphasis on select subjects. Educators believed students needed to be proficient in just a few areas to make productive contributions to industry and society. Additionally, extensive testing resulted in uniform and didactic teaching methods by classroom instructors. Educators determined student achievement in English and arithmetic, for example, by using so-called "teaching tests." That is, teachers provided instruction on material specifically relevant to the test of achievement and later used those same tests to demonstrate student competency and quality classroom practices.⁴⁹ Administrators like Boyer encouraged such methods, describing national standardized tests such as the Briggs English Form Test as illustrative of "the value of the test-teach-test procedure [for] instruction and drill." Boyer further claimed that "differing pupil capacity" explained variation in student results in arithmetic, which teachers could overcome by providing "review and drill" and "motivation for completely accurate work." Indeed, Boyer argued that "the drill lesson, in recent years neglected, has a legitimate place in modern teaching practice. The testing program has tended to emphasize the importance of drill and therefore some teachers have modified their methods accordingly."⁵⁰ Test-drill-test models of instruction and uniform teaching methods contributed to increased promotion rates of Philadelphia schoolchildren over time (Table 4). School administrators thus openly praised widespread testing as "evidence of increasing efficiency" and for continued improvement in classroom instruction throughout the entire school system.⁵¹

"Increased efficiency" arguably improved student promotion rates. Although students in all categories of nonpromoted students were a negligible percent of the entire student population prior to the introduction

⁴⁸ Boyer, "Educational Tests and Measurements," 37–45.

⁴⁹ Carmon Ross, "Discussion," in *Annual Schoolmen's Week Proceedings*, vol. 7 (Philadelphia, 1920), 147, warned of "teaching tests" that did not vary and were administered "again and again until both pupils and teachers became thoroughly familiar with the contents of the tests."

⁵⁰ Boyer, "Report of the Division of Educational Research and Results," in *Annual Report of the Board of Public Education* (Philadelphia, 1927), 581; Boyer, "Report of the Division of Educational Research and Results," in *Annual Report of the Board of Public Education* (Philadelphia, 1928), 549, 492–93.

⁵¹ Boyer, "Report of the Division of Educational Research and Results" (1930), 177–78.

Table 4: Student promotion rates, 1925 to 1930^a

<i>January</i>	<i>1920</i>	<i>1921</i>	<i>1922</i>	<i>1923</i>	<i>1924</i>	<i>1925</i>	<i>1926</i>	<i>1927</i>	<i>1928</i>	<i>1929</i>	<i>1930</i>
Senior High School	79.4	80.2	81.4	78.6	79.7	81.6	80.8	78.4	83.0	85.0	85.0
Junior High School	83.9	87.7	91.2	90.1	92.8	91.8	91.2	88.7	89.0	90.0	90.0
Elementary School	83.6	83.9	84.4	84.1	85.5	87.0	87.4	86.9	87.9	87.9	87.5
<i>June</i>											
Senior High School	78.6	78.2	76.2	78.6	77.7	76.5	76.7	74.1	83.0	84.0	84.0
Junior High School	86.7	89.4	89.2	89.3	90.4	89.9	88.3	89.3	89.0	90.0	90.0
Elementary School	85.0	84.5	84.9	85.7	86.3	87.3	86.8	87.2	87.5	87.5	87.7

a. As reported in "Report of the Superintendent of Schools," in *Annual Report of the Board of Public Education* (Philadelphia, 1921–25), passim; "Report of the Division of Educational Research and Results," in *Annual Report of the Board of Public Education* (Philadelphia, 1926–30) passim.

of widespread testing, the number of students *not* promoted from year to year decreased following educators' efforts to group students according to ability, further legitimizing claims of efficiency through standardized testing (Table 5).

Philadelphia's Dubious Legacy as a Modern Urban School System

By 1930, several elements of the specialized course at South Philadelphia High School for Girls characterized Philadelphia's public school system as a whole, and many aspects of latter-century urban education had coalesced in Philadelphia's public schools. Primary among them was an emphasis on pupils' minimum competencies—demonstrated on numerous standardized tests—of the few subject areas relevant to students' social conformity, economic potential, and future citizenship. Further, the importance educators placed on students' performances on standardized tests encouraged teachers to employ routine methods in their classroom instruction. Finally, administrators advocated instructional uniformity and narrowed curricular options for students. The "ultimate aim" of educational research, Philip Boyer averred, was "more effective guidance" for students, ensuring their proper placement in "specialized forms of training [to meet] the qualifi-

Table 5: Cause, number, and percent of student population of nonpromoted students, 1920 to 1927^a

	1920	1921	1922	1923	1924	1925	1926	1927
School Population	215,862	226,230	225,810	227,306	229,942	230,529	230,536	232,455
Mental Deficiency ^b	87 (0.04)	86 (0.04)	81 (0.04)	61 (0.03)	72 (0.03)	57 (0.02)	50 (0.02)	74 (0.03)
Backwardness ^c	738 (0.34)	720 (0.32)	721 (0.32)	676 (0.30)	575 (0.25)	518 (0.22)	525 (0.23)	571 (0.25)
Foreign Parentage ^d	72 (0.03)	66 (0.03)	81 (0.04)	69 (0.03)	65 (0.03)	33 (0.01)	29 (0.01)	36 (0.02)
Irregular Attendance ^e	368 (0.17)	334 (0.15)	347 (0.15)	465 (0.20)	429 (0.19)	339 (0.15)	318 (0.14)	241 (0.10)
Other Causes ^f	118 (0.05)	93 (0.04)	86 (0.04)	86 (0.04)	67 (0.03)	65 (0.03)	50 (0.02)	47 (0.02)
Total	1,383 (0.64)	1,299 (0.57)	1,316 (0.58)	1,357 (0.60)	1,208 (0.53)	1012 (0.44)	972 (0.42)	969 (0.42)

a. Author's calculations based on the number of nonpromoted students and the causes for their non-promotion reported in "Report of the Superintendent of Schools," in *Annual Report of the Board of Public Education* (Philadelphia, 1921–25), passim; "Report of the Division of Educational Research and Results," in *Annual Report of the Board of Public Education* (Philadelphia, 1926–27), passim; and the total school population in "Report of the Division of Compulsory Education," in *Annual Report of the Board of Public Education* (Philadelphia, 1931), 301. From this is derived the percentage of the total student population represented in each category. The Division of Educational Research and Results stopped reporting the nonpromoted statistics after 1927 without comment. b. Defined as "Defective, weak mentality, sub-normal, feeble-minded." c. Defined as "Dull, slow development." d. Defined as "Non-English speaking, foreign home." e. Defined as "Non-attendance, truancy." f. No definition provided.

cation set by industry as requirements for particular occupations." Boyer understood that the schools system's success in fulfilling this obligation to business and society through testing and "guidance" depended heavily on the cooperation of educators at the school level. Fortunately for Boyer, school principals and classroom teachers reported high levels of satisfaction with the results and expressed their appreciation for the ways they perceived objective assessments improved instruction and raised teachers' consciousness of individual students' abilities.⁵²

⁵² Boyer, "Educational Measurements," 369, 371; Boyer, "Report of the Division of Educational Research and Results" (1927), 562; Boyer, "The Philadelphia Experiment in Homogeneous Grouping," in *Annual Schoolmen's Week Proceedings*, 17:252; Boyer, *Report of the Division of Educational Research and Results* (Philadelphia, 1926), 28–29; *Report of the Division of Educational Research and Results* (Philadelphia, 1927), 36–37.

The work of Philadelphia's Division of Educational Research and Results during the mid to late 1920s thus illustrates the powerful influence the institutionalization of standardized testing had on several aspects of the modern urban school system. Urban educationalists responded to community leaders' demands to address the perceived threats of social and economic instability presented by the increased presence of newcomers to the city. The strategies city educators employed to alleviate such anxieties were administrative in nature. The Philadelphia school system expanded its specialized bureaus in the early twentieth century to solve the perceived problems of an increasingly diverse community, based the administration of these special divisions on principles of efficient management, and proved their success in objectively collected data embodied in standardized test results and evidenced by increasing student promotion rates. Educationalists such as Boyer and Broome believed so-called scientific assessments enabled educators to determine appropriate employment and educational opportunities for students, which in turn was the best way of publicly demonstrating successful educational reform initiatives. As Broome, Boyer, and the Division of Educational Research and Results oversaw the expanding number of administered tests and the number of tested students, however, they emphasized measurement as a means to maintain the educational bureaucracy and to legitimize their professional authority within that order. Today's teachers and students who labor under high-stakes testing regimens borne from and institutionalized by educational policies such as No Child Left Behind, Race to the Top, and Common Core are heirs to their legacy.⁵³

Northeastern Illinois University

RENÉ LUIS ALVAREZ

⁵³ Javier C. Hernández and Al Baker, "A Tough New Test Spurs Protest and Tears," *New York Times*, Apr. 19, 2013, A24; Motoko Rich, "Debut of School Standards Is Rocky, and the Critics Are Pouncing Left and Right," *New York Times*, Aug. 16, 2013, A11.