

# *Pupil Assessment*



*Informational  
Paper 32*

*Wisconsin Legislative Fiscal Bureau  
January, 2003*

# *Pupil Assessment*



*Prepared by  
Layla Merrifield*

*Wisconsin Legislative Fiscal Bureau  
One East Main, Suite 301  
Madison, WI 53703*



# *Pupil Assessment*

---

This paper provides information on testing programs for elementary and secondary school pupils that are administered or coordinated by the Office of Educational Accountability within the Department of Public Instruction (DPI). The first section of this paper provides background and definitions on assessment alternatives; the second section describes past assessment programs; the following sections outline current assessment programs, previous and current assessment initiatives and federal requirements; and the final section discusses funding for assessment initiatives.

---

## **Background**

---

In recent years, pupil assessment has become the focus of broader educational reforms in response to national reports that the academic performance of U.S. pupils has fallen behind that of other countries, particularly in areas requiring more complex thinking skills. There is evidence that gaps in performance between whites and minorities, economically advantaged and disadvantaged pupils, and males and females have also widened. As a result, greater emphasis has been placed on the purposes and content of pupil assessments and the consequences of test results for teachers, pupils, schools and school districts.

There are three primary purposes of pupil assessment: (1) to evaluate the quality and level of pupil achievement and indicate what pupils, teachers, schools, districts, and states can do to im-

prove their performance; (2) to provide accountability information (the relationship between public investment in education and pupil achievement); and (3) to provide information that can be used by teachers and pupils in decisions relating to remediation, program placement, career paths, and ranking. Different types of assessments are administered depending on the kind of information sought. Below is a description, based on information provided from DPI, on the most widely used types of assessment instruments.

**Standardized tests.** Narrowly defined, standardized tests are tests given to a large number of pupils with identical directions, time limits and questions. Most standardized tests are purchased from commercial publishers. In the past, multiple-choice and true/false questions have been associated with standardized testing. However, recent developments in the field of educational testing have allowed test vendors to include short answer and essay questions in the standardized test as well. Standardized tests are used to measure knowledge of a particular subject or basic aptitude and may or may not be associated with the curriculum.

While standardized tests are available in a variety of skill levels and formats, two types of decisions are commonly made with their result: normative decisions and criterion-based decisions. Normative decisions measure a pupil's performance in relation to a norm group. Tests used to make normative decisions or norm-referenced tests (NRTs) compare the rankings of all pupils taking the test. Results from this type of exam are used to determine where pupils score in comparison to all

other pupils. Test statistics such as percentiles, norm-equivalent scores, and standardized scores are used to make normative decisions.

The second type of decisions made with standardized tests is criterion-based decisions. Test used to make criterion-based decisions or criterion-referenced tests (CRTs) measure how well pupils have learned specific curricular material. Unlike NRTs, a pupil's score is not compared to that of other pupils, but to a minimum standard or criterion. Statistics commonly used with CRTs are pass/fail rates and percent of mastery. Proficiency categories, like those used in Wisconsin, reflect criterion-based decisions. Scores are set for each category and pupils are placed into these categories based on their performance on the tests.

Standardized tests are widely used for accountability purposes because they allow comparisons among pupils, schools, school districts, and states; are easy to administer and score; and are usually the most cost-effective type of test. However, they are frequently criticized as being culturally and/or economically biased and emphasizing less important factual knowledge and rote memorization skills rather than higher-order skills such as problem-solving, writing, and critical thinking. Another criticism is that the pressure to raise standardized test scores encourages schools to adjust their curricula to focus on test material, or "teach to the test," which results in narrowing the curriculum and further encouragement of memorization skills over more complex thought. Norm-referenced tests in particular have been criticized as providing misleading information when the original norm group's scores are as much as a decade old. Critics of criterion-referenced tests dispute the use of standards, which they believe may be arbitrary, and the emphasis placed on passing the standard rather than performing as well as possible.

**Performance Assessments.** To address such criticisms of standardized tests and create assessments which are more authentic, representing situations that pupils may encounter in daily life, and valid, providing true and desired information

about the abilities of pupils, many states and individual school districts have developed or are developing alternative assessments. These include various methods intended to measure not only knowledge of a particular subject, but also the use of complex reasoning and problem-solving skills. Also called performance-based or outcome-based assessments, performance assessments are designed to require pupils to demonstrate what they know and can do and to integrate interdisciplinary knowledge into the accomplishment of a task. Tasks are aligned with the curriculum and may include writing exercises, math problems, science experiments, open-ended multiple-choice questions, or a combination of these. Unlike multiple-choice or true/false questions, in which a pupil selects an answer, performance assessments require pupils to produce an original answer. For example, such assessment could require pupils to write an essay or solve a mathematical problem showing the steps involved in reaching the solution. The U.S. Department of Education classifies performance assessments into three categories: open-ended tasks, extended tasks, and portfolios.

Open-ended tasks are activities in which pupils respond immediately to a question. Short answer or essay questions are considered open-ended tasks. Extended tasks are activities that required long time periods to complete. For example, a science experiment, which requires several days to complete, would be an extended task. In this example, a pupil might be scored based upon the outcome of the experiment. A portfolio is a file or collection of student projects and/or tasks collected over an extended period of time. Portfolios may include both open-ended and extended tasks such as artistic projects, tasks completed in cooperation with other students, written assignments, and items that the pupil feels represent his or her best effort.

Proponents of alternative assessments argue that because they are intended to be intrinsically valid and authentic, they will automatically improve instruction and emphasize and expedite the achievement of valuable educational goals. How-

ever, significant obstacles to the implementation of alternative assessments exist. Due to the complexity of the tasks involved, alternative assessments are more costly and less efficient to develop and score than traditional tests. Generally, these assessments must be manually scored by trained readers whose work is monitored for consistency. In some programs, each assessment is scored by two or more readers and the results averaged. Because alternative assessments require substantial amounts of time to complete and score, the number of tasks must be limited, which increases the possibility of bias as well as insufficient content coverage. Perhaps the most critical challenge for developers of alternative assessments is how to design and score them in order to provide accountability information such as comparative data for pupils, schools, and school districts.

Nonetheless, DPI notes that performance assessments have many positive characteristics. They do tend to be designed more like the activities and tasks pupils experience in their classes. Because these assessments look similar to the types of activities experienced in the classroom, parents and students believe them to be valid. Teachers who are qualified as readers for performance assessment tend to be more aware of what is tested and thus better able to communicate curricula and standards to their students.

Both standardized tests and performance assessments provide useful assessment information. The Department indicates that many test vendors are creating tests that utilize both performance assessment and standardized tests through "multiple assessments." Vendors now can combine multiple choice, true/false, short answer, and essay questions into one test. These new "multiple-assessments" typically have a larger proportion of multiple-choice items. The combination of both types can provide more complete information on a pupil's education.

---

## Previous Wisconsin Assessment Programs

---

**1. Wisconsin Pupil Assessment Program 1975-1987.** The Wisconsin Pupil Assessment Program was designed to measure pupil achievement in specific skill areas using both criterion- and norm-referenced examinations. The tests were administered to a randomly selected sample of pupils in a group of schools chosen according to their geographic location, district size, and grade enrollment.

**2. Competency-Based Testing Program (CBT) 1985-1992.** The criterion-referenced, curriculum-based CBT exams were designed to test pupils at certain grade levels for minimum standards of proficiency in reading, language arts, and mathematics. Participation by school districts was voluntary, and the district could either develop its own exams, with DPI approval, or use test questions developed by DPI. Participating districts were required to test all pupils once in grades K to 5, once in grades 6 to 8, and once in grades 9 to 11 and were reimbursed by DPI for the costs of printing and scoring the exams.

Districts administering the CBT were required to release test results to pupils' parents or guardians and provide remediation services to any pupil whose test scores did not meet district minimum standards. District scores were reported to the school board with recommendations for curricular changes. Since the results were neither made public, nor provided to DPI, there was no method for comparing the performance of one district to another even if identical tests were used. Each district's scores were, however, included in its annual performance disclosure report required by state law.

**3. Achievement Tests (Standard "s") 1988-1992.** The achievement tests, or standard "s" tests, (named after s. 121.02 (1)(s) of the statutes), were similar to CBTs in that districts were required to test pupils in reading, language arts, and mathe-

matics using curriculum-based tests. Although the standard "s" tests differed from the CBT tests in the frequency of testing required, districts which participated fully in the CBT program automatically met this standard. The results were used to determine if program goals were being met and to monitor pupil achievement. No remediation or parental notification was required.

---

### **Current Wisconsin Assessment Programs**

---

In 1991 Act 269, the CBT program and the standard "s" requirement were repealed. These programs were replaced by a requirement that school districts, beginning in 1993-94, administer "knowledge and concepts" examinations in the 8<sup>th</sup>, and 10<sup>th</sup> grades and beginning in 1996-97 administer a 4<sup>th</sup> grade knowledge and concepts examination. The tests are designed to measure a pupil's knowledge in the subject areas of mathematics, science, social studies, reading and language arts, including an assessment of a pupil's writing ability.

The following section describes the current Wisconsin Student Assessment System (WSAS), which includes the Wisconsin Reading Comprehension Test at 3<sup>rd</sup> grade and the 4<sup>th</sup>, 8<sup>th</sup>, and 10<sup>th</sup> grade knowledge and concepts exams.

**Wisconsin's Reading Comprehension Test (WRCT).** Current law requires all districts to annually administer a standardized reading test, developed by DPI, to 3<sup>rd</sup> grade pupils. Formerly the Third Grade Reading Test, the WRCT is a test given in the spring of each year, intended to: (a) identify marginal readers who may need remediation; (b) provide comparative performance data by school and school district; (c) allow school districts to evaluate their reading programs; and (d) provide data for meeting federal and state requirements regarding student assessment.

Remedial reading services for pupils in

kindergarten through grade four are required if: (a) a pupil fails to meet the district's reading objectives; or (b) a pupil fails to meet the minimum performance standard on the WRCT and either the teacher and the pupil's parent or guardian agree that the test results accurately reflect the pupil's ability or the teacher determines that based upon other objective evidence of the pupil's reading comprehension, the test results reflect the pupil's reading ability. Only 3<sup>rd</sup> grade pupils identified as limited-English proficient (LEP) or children with disabilities may be excluded from taking the WRCT.

The WRCT was first given in the 1988-89 school year to 56,533 third graders. In 2001-02, 57,259 pupils took the exam. The test is currently comprised of three reading passages, two narrative passages and one expository passage and three types of questions: (a) reading comprehension; (b) prior knowledge; and (c) reading strategies. The exam, which includes approximately 103 questions, is taken over three testing sessions within three weeks. Through 1997, test scores for the reading comprehension questions were placed into three performance categories: above the performance standard, inconclusive and below the performance standard. DPI defined inconclusive as neither clearly above nor below the standard.

Beginning in 1998, test score reporting categories were redefined as four proficiency levels: (a) advanced, defined as distinguished in the content area; (b) proficient, meaning competent in the content area; (c) basic, defined as somewhat competent in the content area; and (d) minimal, meaning limited achievement in the content area. Students who score in the minimal proficiency level must be evaluated further to determine if they are in need of remedial reading or other services. The Department indicates that students who score above the minimal level are said to meet the minimum performance standard, while the proficient and advanced levels are the long-term educational goals for all students. These levels are based on what DPI, in conjunction with teachers across the state, determined a pupil should know

in order to meet the state's recently issued academic standards.

For the 2001-02 test, the maximum score was 67 points with performance scores based on the following: (a) advanced – 63 to 67 points; (b) proficient – 48 to 62 points; (c) basic – 29 to 47 points; and (d) minimal – 28 or less points. Scores for reading strategy and prior knowledge are not included in the performance scores, rather those scores are used to interpret results on the comprehension questions. The statewide average score for the reading comprehension questions was 54.6 points out of 67 total points. Table 1 shows the statewide totals of pupil scores on the WRCT for 2001-02 according to each proficiency level.

**Table 1: 2001-02 Statewide Reading Comprehension Test Results**

Proficiency Level	Number Tested	% of Total
Advanced	16,874	27.5%
Proficient	28,507	46.6
Basic	8,510	13.9
Minimal	3,368	5.5
Subtotal	57,259	93.5%
Excluded	3,962	6.5
Total	61,221	100.0%

The 2001-02 WRCT exam questions were developed by Wisconsin educators in conjunction with MetriTech, Inc. the contractor for WRCT development. A state advisory committee made up of educators reviews all test items and pilot questions. MetriTech, Inc. is also responsible for the production, distribution, scoring and reporting of the results of the WRCT under a separate contract.

Administrative rules provide that DPI will pay for printing, distribution, scoring and reporting the results of the WRCT. Under current law, school districts are allowed to provide the scoring of the exams and DPI is required to reimburse the

districts for such costs, not to exceed the cost to DPI of scoring. The cost of printing, distribution, scoring and reporting the results of the WRCT was \$336,100 in 2001-02. In addition, DPI incurred test development costs of \$419,400 in 2001-02.

**Knowledge and Concepts Examinations.** In 1992-93, DPI was required to make available to districts, at no charge, examinations designed to evaluate the level of knowledge attained by pupils in the 8<sup>th</sup> and 10<sup>th</sup> grades. District participation was voluntary in 1992-93 and required beginning in the 1993-94 school year. A third exam, for pupils in fourth grade, was added under 1995 Act 27. School district participation for the 4<sup>th</sup> grade exam was voluntary in 1995-96 and required beginning in the 1996-97 school year.

Currently, the 4<sup>th</sup>, 8<sup>th</sup>, and 10<sup>th</sup> grade knowledge and concepts examinations are designed to evaluate the level of knowledge attained by pupils in the areas of mathematics, science, social studies, reading and language applications. In 2001-02, the 4<sup>th</sup>, 8<sup>th</sup>, and 10<sup>th</sup> grade tests consisted of multiple choice and short-answer questions in language arts, reading, mathematics, science and social studies and a writing test related to an assigned reading passage. In addition, each 8<sup>th</sup> and 10<sup>th</sup> grade test contains an optional, non-academic section consisting of questions related to the pupil's career interests, intended to aid in pupil guidance counseling and course selection. To familiarize 4<sup>th</sup> grade pupils with test content and format, school districts are required to administer a practice activities test before the pupils take the 4<sup>th</sup> grade knowledge and concepts examination.

School boards can decide to exclude from testing limited English-proficient pupils, may permit such pupils to be examined in his or her native language or can modify the format and administration of the tests for these pupils. Districts must include children with disabilities in the tests, with appropriate modifications where necessary or alternative assessments for those children who cannot participate in the assessment. If a district excludes certain children with

disabilities from the assessment, then a statement explaining why that assessment was not appropriate and how the pupil will be assessed through alternative means must be included in the pupil's individualized educational program. In addition, a statement must be included in a pupil's program indicating any modifications that were made to the pupil's assessment. Any 4<sup>th</sup>, 8<sup>th</sup>, or 10<sup>th</sup> grade pupil may be excused from taking the tests upon the request of the pupil's parent or guardian.

The full battery of tests, for each grade level, requires approximately six hours to complete and is usually administered over a three-day period. In 1993-94, the first year of required administration, 55,570 8<sup>th</sup> grade pupils, or 86% of the total 8<sup>th</sup> grade enrollment, completed all subject area tests. At the 10<sup>th</sup> grade level, 50,561 pupils, or 80% of the total 10<sup>th</sup> grade enrollment, completed all subject area tests. In 2001-02, approximately 60,735 4<sup>th</sup> grade pupils (95.8% of the total enrollment), 64,818 8<sup>th</sup> grade pupils (97.9%) and 69,962 10<sup>th</sup> grade pupils (98.0%) completed each subject area test. According to DPI, participation rates vary widely across schools. A three-week testing window is provided to allow local flexibility in scheduling for make-up testing. Despite this fact, many students are not tested who are required to be. The Department notes that economically disadvantaged and minority students tend to have lower participation rates than other student groups.

Starting in 1997-98, results of the knowledge and concepts examinations are reported by proficiency categories. Separate results are reported for each test area: reading, mathematics, science, social studies, language arts, and writing. Scores on the writing sample were formerly combined with scores on the language arts test. These combined scores were called enhanced language scores. The Department indicates that these scores were difficult to explain and delayed reporting, and as a result scores for the writing sample and language arts tests have been reported separately since 1998-99. Writing scores are not used for performance scoring purposes. The proficiency levels are categorized as minimal performance, basic, proficient

and advanced and generally defined the same as the WRCT.

Proficiency summaries are reported for all students who have been enrolled in the school or district for a full academic year, as well as for a partial year, regardless of disability or English-proficiency status. Previously, scores were reported only for students who took the test. Under the new proficiency levels reporting, those pupils not tested are listed under the not tested category and are not included in proficiency level scoring. Beginning in 1998-99, DPI also reports the percentage of students not excluded or excused from the test. As in the past, scores of students tested are also reported as percentile rankings that compare each pupil's performance to that of their peers statewide and nationwide in each subject area.

Wisconsin's statewide test results for each subject area of the 4<sup>th</sup>, 8<sup>th</sup>, and 10<sup>th</sup> grade tests for 2001-02 are provided in Table 2. The normative data is based on all students tested while the statewide proficiency scores are reported for all students enrolled. The table shows, for each grade level tested and by each test area, the percentage of students enrolled in Wisconsin public schools that scored at each proficiency level and the percentage of students that were not tested. In addition, Table 2 provides the national percentile rank for each test area according to grade level. The national percentile rank compares each pupil's performance with the performance of a national norm group of pupils. For example, the national percentile rank for the 4<sup>th</sup> grade reading test in 2001-02 was 67, which means that the average student in Wisconsin scored as well as or higher than approximately 67% of the students in the national norm group. The national percentile ranks range from 1 to 99 with the average rank in the national norm group of 50.

Federal law requires that the results must also be reported by gender, race/ethnicity, by English proficiency status, by students with disabilities as compared to non-disabled students and by economically disadvantaged students as compared to students who are not economically disadvantaged.



**Table 2: 2001-02 Statewide Knowledge and Concepts Exam Results (Percent of Pupils in each Proficiency Level)**

	Reading	Language	Mathematics	Science	Social Studies
<b>4th Grade</b>					
Not Tested	5%	6%	4%	4%	3%
Minimal	5%	4%	4%	4%	5%
Basic	11%	17%	23%	15%	10%
Proficient	61%	42%	44%	58%	43%
Advanced	18%	31%	25%	19%	39%
Nat'l Percentile Rank	67	67	64	62	68
<b>8th Grade</b>					
Not Tested	3%	3%	4%	3%	4%
Minimal	11%	6%	17%	11%	5%
Basic	12%	21%	35%	26%	11%
Proficient	56%	51%	28%	43%	45%
Advanced	18%	19%	16%	17%	35%
Nat'l Percentile Rank	65	64	71	65	67
<b>10th Grade</b>					
Not Tested	7%	5%	6%	6%	6%
Minimal	10%	13%	29%	17%	12%
Basic	23%	20%	22%	29%	17%
Proficient	37%	44%	27%	35%	40%
Advanced	23%	18%	16%	13%	25%
Nat'l Percentile Rank	69	71	75	69	67

The 2001-02 results of the 4<sup>th</sup>, 8<sup>th</sup>, and 10<sup>th</sup> grade exams for all pupils, by school district, school, and by demographic group within the district or school are available on the Department's website [[www.dpi.state.wi.us/dpi/oea/stkce.html](http://www.dpi.state.wi.us/dpi/oea/stkce.html)]. The Department advises that school and district national percentile ranks based on students tested should be interpreted with caution when test participation rates are low, because student groups with the lowest achievement levels typically have the lowest test participation rates while school results based only on students tested would be higher for schools that do not test lower achieving students.

Schools are held accountable for achievement and progress in each subject area. Low achievement in reading, for example, is not offset by high achievement in math. Current law prohibits using the results of the knowledge and concepts tests to evaluate teacher performance, discipline teachers, or as a reason for non-renewal of their contracts. Further, a district's scores may not be used to determine its general or categorical school aids. The

tests are also required, to the extent possible, to be free from bias.

The contract costs for printing, scoring and reporting the results of these assessments was \$1,803,200 in 2001-02, and is budgeted to be \$2,622,600 in 2002-03. DPI currently provides these examinations through a six-year contract with a testing vendor, CTB/McGraw-Hill.

Under 1997 Act 237, starting in 1998-99 a school board operating elementary grades may develop or adopt its own examination designed to measure pupil attainment of knowledge and concepts in 4<sup>th</sup> and 8<sup>th</sup> grades. If a school board develops or adopts its own examination it is required to notify DPI. In addition, the board must provide the State Superintendent with statistical correlations of those examinations with the 4<sup>th</sup> and 8<sup>th</sup> grade knowledge and concepts examinations adopted or approved by the State Superintendent and the federal Department of Education must approve the examination.

---

## Previous State Assessment Initiatives

---

**The Educational Goals Committee.** This committee, which consisted of 12 members including the Governor, State Superintendent, President of the UW System, Director of the Wisconsin Technical College System and members appointed jointly by the Governor and State Superintendent was required to hold 12 regional conferences and one statewide conference to allow school boards and the public to submit their recommendations. A final list of 28 goals, was submitted to the Legislature in September, 1993. The goals were placed in three categories: (a) learner goals which indicate expectations of students; (b) institutional goals which refer to school staff and environment; and (c) societal goals which focus on conditions outside of the educational community. Although the new goals were intended to replace those provided in the statutes, no statutory changes that would accomplish this were made.

**The State Superintendent Assessment Advisory Committee (SSAAC).** Comprised of teachers, parents and other interested persons appointed by the State Superintendent, SSAAC was charged with advising the State Superintendent on utilizing the new educational goals in the development of a new pupil assessment program. The State Superintendent was required to submit a report to the Legislature by January 1, 1994, on plans for implementing such an assessment program in the 1996-97 school year. The report recommended a program consisting of three types of assessments to be based on the first three of the proposed new learner goals which state that a pupil: (a) build a substantial knowledge base; (b) develop thinking and communication processes; and (c) apply knowledge and processes.

While the assessment program would have been based on the first three learner goals, the advisory committee recommended that more detailed "learner outcomes" be used to guide the development of the assessment items. The 17

learner outcomes, which were developed by DPI with input from Wisconsin educators at several meetings in 1992 and 1993, stated more precisely what students should be able to do, for example, develop and test a hypothesis.

The Department began developing WSAS following the recommendations of the SSAAC, and requested that funding be included in the 1995-97 biennial budget. The proposed assessment program included limited response tests, performance assessments, portfolios and gateway assessments at the 10th grade level. Funding for these changes in the assessment program was not authorized by the Legislature; however, the Legislature did require DPI to study the utility of administering technology-based performance assessments.

**Governor's Advisory Taskforce on Education and Learning.** In January, 1996, the Governor created the Governor's Advisory Taskforce on Education and Learning by executive order to address policies surrounding educational standards, assessment and accountability. Specifically, the Taskforce was directed to: (a) identify which educational functions should be performed by the state and which should be performed at the local level through school districts or cooperative educational service agencies (CESAs); (b) evaluate the current use of state resources to determine the best method to assist school districts in providing quality educational opportunities; and (c) identify the tools which need to be strengthened or utilized to achieve the goal of greater student learning.

Appointed by the Governor and composed of the Secretaries of the Departments of Administration and Revenue (DOR) and members of the state's business community, education professions and general public, the Taskforce created several subcommittees, including a Subcommittee on Standards and Assessment. This subcommittee, which included the State Superintendent, the Secretary of DOR, the Lieutenant Governor and parents, teachers and school administrators studied topics such as a high school graduation test, detailed statewide

standards and various testing options. The Taskforce released its final report in March, 1996.

---

### Recent State Assessment Initiatives

---

**Governor's Council on Model Academic Standards.** By executive order in January, 1997, the Governor created the Governor's Council on Model Academic Standards. The Council consisted of the Lieutenant Governor who served as chair, the State Superintendent of Public Instruction, the chairs and ranking minority members of the Senate and Assembly Education Committees and one public member appointed by the Governor. The Council was responsible for working on the development of academic standards for all pupils in English language, arts, mathematics, science, and social studies at grades 4, 8, and 12.

As part of the 1997-99 budget (1997 Act 27), a Standards Development Council under the Office of the Governor was statutorily created that was nearly identical to the Governor's Council. Statutorily, the Council was required review to the Governor's proposed pupil academic standards in mathematics, science, reading and writing, geography, and history. Through 1997, the Council held various public meetings on the proposed standards. The Council's final recommendations on the standards were provided to the Governor in December, 1997. The Governor then had 30 days to approve or disapprove of the Council's recommendations. In January, 1998, the Governor approved the recommended standards and issued the standards as Executive Order 326. By August 1, 1998, each school board had to either adopt these statewide academic standards or develop their own. The Council is required to review the issued pupil academic standards periodically. If the Governor approves any subsequent modifications to the standards recommended by the Council, the changes can be issued as an executive order.

**High School Examinations.** Under 1997 Act 27,

DPI was required to design a state high school graduation test that local school districts may use if they have adopted the Model Academic Standards as issued and approved under Executive Order 326. DPI included in its 2001-03 budget request approximately \$4.6 million annually above its base level \$2.5 million for completion of development of the high school graduation test and its implementation. Funding was not included in the final budget, 2001 Act 16. Instead, statutory requirements for its implementation were delayed by two years under 2001 Act 109, the 2001-03 budget adjustment act. Under 2001 Act 109, each school district that operates a high school is required to adopt a written policy by September 1, 2004, specifying criteria for granting a high school diploma. Beginning in 2004-05, a high school graduation test must be used by school boards as one of several criteria for graduation, which should also include a pupil's academic performance and teacher recommendations. Beginning September 1, 2005, a high school diploma cannot be granted to any pupil unless the pupil has satisfied the school board's criteria. The test may be administered only in grades 11 and 12, and must be offered twice each year. In addition, a board must excuse a pupil from the high school graduation exam upon the request of a parent or guardian. These provisions apply to charter schools as well.

A school board must adopt a high school graduation exam that measures whether pupils meet pupil academic standards adopted by the school board. If the board adopts the statewide standards issued by executive order, the board could adopt the high school graduation exam developed by DPI. If a school board develops and adopts its own high school graduation examination, it is required to notify DPI.

In 2001 Act 109, base level funding of \$2.5 million for DPI's administration and development of the high school graduation test was deleted in 2002-03. The Department did not include funding for the high school graduation test in its 2003-05 biennial budget request.

School boards are required to establish alternative criteria upon which to determine qualification for high school graduation if a pupil has been excused from the high school graduation exam. In order to graduate from high school, a pupil who was excused from the examination must satisfy the alternative criteria.

**4<sup>th</sup>, 8<sup>th</sup>, and 10<sup>th</sup> Grade Knowledge and Concepts Examinations.** Under 1999 Act 9, beginning with the 2002-03 school year, school districts must administer the state's 4<sup>th</sup>, 8<sup>th</sup>, or 10<sup>th</sup> grade examination or develop and administer its own examinations to measure pupil attainment of knowledge and concepts in the respective grades. School boards were required to provide a pupil with at least two opportunities to take the 4<sup>th</sup> or 8<sup>th</sup> grade examination; however, this requirement was deleted under 2001 Act 16.

Under Act 9, school boards and charter schools were required to devise written policies for promoting pupils from grade four to grade five and from grade eight to grade nine by September 1, 2002. The knowledge and concepts examination score, unless the pupil has been excused from taking the exam by a parent or guardian, is one of several criteria to be used to make the promotion decision, including the pupil's academic performance and teachers' recommendations, along with any other criteria the school board or charter school operator chooses. Beginning September 1, 2002, a school board or charter school operator cannot promote a 4<sup>th</sup> or 8<sup>th</sup> grade pupil unless the pupil satisfies the board's criteria for promotion.

A school board may determine not to administer an examination to a pupil enrolled in a special education program or a limited-English proficient pupil, and a school board may modify the format and administration of an examination for these pupils or permit a pupil to be examined in his or her native language. Additionally, school boards are required to excuse a pupil from taking the 4<sup>th</sup> or 8<sup>th</sup> grade examination upon the request of the pupil's parent or guardian.

---

## Federal Assessment Programs and Requirements

---

This section provides a discussion of a national assessment program in which Wisconsin voluntarily participates and recent changes to federal law that directly affect pupil assessment in Wisconsin.

**National Assessment of Educational Progress (NAEP).** The NAEP, commonly referred to as the Nation's Report Card, is intended to provide a continuous national survey of educational achievement and trends. The program is administered by the Commissioner of Educational Statistics, who heads the National Center for Education Statistics in the U.S. Department of Education. The independent National Assessment Governing Board, appointed by the Secretary of Education, governs the program and is responsible for selection of subject area to be assessed, development of assessment methodology, standards, testing procedures and reporting. Under NAEP, objective-referenced tests are administered periodically to representative, randomly selected national and state samples of 4th-, 8th-, and 12th-grade pupils. Items included in the NAEP are fixed-response, machine-scorable, multiple-choice questions and open-ended questions.

Since 1969, assessments have been conducted periodically in reading, mathematics, writing, science, history/geography or other areas including music, art, computer competence and civics. The NAEP has used the results to track changes in national student achievement levels over time and collect information on pupil performance by gender, race/ethnicity and other variables intended to indicate the pupils' instructional experiences. In 1990, NAEP began administering trial state-level assessments with an 8th-grade mathematics assessment, which represented the first national program designed to provide state-by-state comparisons of pupil achievement. Previously, NAEP was specifically

prohibited from reporting results at the state level.

In 1992, 42 states (including Washington, D.C.) participated in the Trial State Assessment. Wisconsin's average proficiency scores on all of the three individual assessments were above the national and regional averages. Although NAEP did not rank state scores, no state had an average proficiency score which was statistically significantly higher than Wisconsin's on any of the assessments. In 1994, Wisconsin ranked third, behind Maine and North Dakota, out of the 41 states and the District of Columbia, the Department of Defense Overseas Schools and Guam that participated in the 4th grade reading proficiency examinations. Overall, 71 percent of Wisconsin pupils who participated, scored at or above the basic reading level, compared to 65 percent for states that participated in the central region (Indiana, Iowa, Michigan, Minnesota, Missouri, Nebraska, North Dakota and Wisconsin) and 59 percent nationally.

In 1996, Wisconsin took part in the 8<sup>th</sup> grade state-level science and the 4<sup>th</sup> and 8<sup>th</sup> grade state-level mathematics examinations. Forty-four states, the District of Columbia, Guam and the Department of Defense participated in the 1996 state-level assessment program. For Wisconsin public school 8<sup>th</sup> grade students the average science score was 160 compared to 155 for states that participated in the central region and 148 nationally, out of a possible 300 points. In 1996, Wisconsin 4<sup>th</sup> grade students' average score on the mathematics examination was 231, out of a possible 500. Nationally the average was 222 and 230 in the central region. Overall, 27 percent of Wisconsin pupils who participated scored at or above the proficient level, compared to 20 percent nationally. Eighth grade students' average math score was 283, compared to the national average of 271 and central region average of 276. Thirty-two percent of the Wisconsin students that participated performed at or above the proficient level compared to 23 percent nationwide.

The 1997-98 national assessment consisted of civics, reading, and writing examinations at 4th,

8th and 12th grades. The state assessment, no longer considered a trial assessment, included a reading examination for 4th and 8th grades and a writing test for 8th grade pupils. In 1997-98, NAEP randomly selected 138 Wisconsin school districts to participate in at least one portion of the national assessment program, including 108 schools from 80 districts for the 4th grade reading test and 113 schools from 89 districts for the 8th grade reading and writing exam. Approximately 2,500 students in each grade and subject in Wisconsin participated in the 1997-98 assessment.

The NAEP 1998 state-by-state writing assessment included a sample of 2,006 eighth-graders from Wisconsin from 80 public schools, which represented three percent of all Wisconsin eighth-graders. The average scale score for pupils from Wisconsin was 153, compared to 148 nationally, on a scale of 0 to 300. The NAEP 1998 state reading assessment used a sample of 2,071 pupils in Wisconsin for grade four and 1,918 pupils in grade eight. The average score for fourth graders was 224 compared to a national average of 215, while the average score for eighth graders was 266 compared to a national average of 261, on a scale of 0 to 500.

Wisconsin participated in the NAEP 2000 assessments of mathematics and science for grades four and eight. The 2000 science assessment used a sample of 1,393 grade four Wisconsin pupils and 1,811 pupils in grade eight. The average score for fourth graders was 157 compared to a national average of 148, while the average score for eighth graders was 162 compared to a national average of 149, on a scale of 0 to 300. The mathematics assessment used a sample of 1,455 fourth graders and 1,760 eighth graders. The average fourth grade score was 229 compared to a national average of 226, while eighth graders scored 287 on average, compared to a national average of 274, on a scale of 0 to 500. Wisconsin participated in the NAEP 2002 state assessments, but test results are not yet available.

In order to administer the exams consistently nationwide, Westat, the NAEP sampling and field

administration subcontractor, is responsible for conducting the building-level assessments.

**No Child Left Behind Act of 2001.** In 2001, Congress reauthorized the Elementary and Secondary Education Act of 1965 (ESEA), renaming it the No Child Left Behind Act (NCLB). Under the reauthorized legislation, schools receiving Title I funds are subject to extensive new accountability provisions. NCLB requires that students be tested in reading and math each year in grades three to eight by 2005-06, with science assessments once each in elementary, middle, and high school beginning in 2007-08. States select and design their own assessments, but the tests must be aligned with the state's academic standards. A sample of 4<sup>th</sup> and 8<sup>th</sup> graders in each state must participate in NAEP in reading and math every other year to provide a point of comparison of the state's results on its own tests.

Additionally, under NCLB states are required to report the performance of districts in making "adequate yearly progress" (AYP), as defined under Title I and measured by academic assessments. AYP must be calculated by race, disability, economic, and limited English proficiency status. States must attain academic proficiency, as defined by each state, for each subgroup of students within 12 years. States must raise the level of proficiency gradually, but in equal increments over time, as compared to a minimum performance threshold based on the lowest-achieving schools or student subgroups. At least 95 percent of each subgroup must take the assessments in order for the school to make AYP. Districts receiving Title I funds must identify a school that fails to make AYP for any subgroup as a school in need of improvement and notify the parents of students enrolled in such schools.

If a school fails to make AYP for two consecutive years, then it is identified for improvement. The school district must provide technical assistance to the school and transportation for students who choose to attend other district schools until the school is no longer identified for improvement. In providing such an option, priority must be given

to the lowest achieving students from low-income families. The district must use five percent of its Title I funds to pay for that option. After a third year of failure to make AYP, the district must also make tutoring and other supplemental educational services available to low-income students still enrolled in the school identified for improvement. Private and public, nonprofit, and for-profit entities may provide these services if they agree to various criteria, including that all content and instruction are secular, neutral, and non-ideological, and are consistent with the district's instructional program. The district must use five percent of its Title I funds to pay for that option. Unless a smaller amount is needed to satisfy all requests, up to 20 percent of a district's Title I funds are required to be spent on either or both of these options. After a fourth year of failure to make AYP, the district must implement corrective actions such as replacing school staff, implementing a new curriculum, providing professional development, or otherwise restructure the school and enable it to make AYP. After a full year of corrective action and continued failure to make AYP, the district must implement major restructuring of the school, including reopening as a public charter school, contracting with a different entity to operate the school, or turning operation over to the state. Requirements related to school improvement, corrective action, or restructuring end if the school makes AYP for two consecutive school years. Title I-A implements similar oversight requirements for states over districts as a whole.

In 2001-02, 1,053 schools in 394 districts in Wisconsin received Title I funding totaling approximately \$149.7 million. A total of 108 Title I schools were identified for improvement (failed to make AYP for at least two consecutive years) in 2001-02 based on 2000-01 assessment results. DPI also received approximately \$7.0 million to cover test development and expansion costs related to NCLB.

Some changes to the Wisconsin Student Assessment System will be necessary to fully comply with the NCLB. DPI indicates that additional test items will be added to the WKCE in all subject ar-

eas to more fully assess state model academic standards. Some standards not assessed by the WKCE will instead be measured and reported at the local district level for Title I accountability purposes. In addition, the WKCE will be administered in the fall rather than spring, in order to be included in promotion decisions, to comply with "no social promotion" provisions of the NCLB.

CTB/McGraw Hill for the knowledge and concepts exams and with MetriTech, Inc. for the WRCT.

2. Contract costs for updates to the Wisconsin reading comprehension test and knowledge and concepts exams.

3. Program operations costs. In 2002-03, the Office of Educational Accountability within DPI consists of 11.0 authorized positions, which are directly responsible for assessment-related activities. Federal funds support 6.65 of these positions. The supplies and services budget includes items such as data processing, printing, travel, space rental, postage, conferences and consultant expenses.

All items included in Table 3 are funded with state general purpose revenue (GPR) except where noted.

### Funding for Pupil Assessment

Table 3 provides a breakdown of total funding provided to DPI for pupil assessment programs from 1999-00 to 2002-03. The table identifies costs in three areas:

1. Printing, scoring and reporting costs. This includes the cost of the contracts with

**Table 3: Expenditures for DPI Pupil Assessment Programs**

	1999-00 Actual	2000-01 Actual	2001-02 Actual	2002-03 Budgeted
Printing, Scoring and Reporting				
Reading Comprehension Test	\$352,700	\$365,400	\$336,100	\$337,000
Knowledge and Concepts Exams	1,526,800	2,154,800	1,803,200	2,622,600
Development				
Reading Comprehension Test	223,800	393,200	419,400	420,000
High School Graduation Test	1,300,000	2,343,100	962,100	0
Knowledge and Concepts Exams	0	0	0	2,100,000 FED
Program Operations				
Salaries and Fringe Benefits	679,500	620,500	507,600	277,900
	73,600 FED	140,800 FED	273,900 FED	437,300 FED
			64,800 PR	64,700 PR
Supplies and Services	364,200	64,400	64,100	56,000
TOTAL	\$4,520,600	\$6,082,200	\$4,431,200	\$6,315,500
Permanent Positions (FTE)	10.6	10.6	5.60	3.60
			6.65 FED	4.65 FED
			0.75 PR	0.75 PR
Project Positions (FTE)	4.0	4.0	4.00	2.00 FED
	2.4 FED	2.4 FED	1.00 FED	---
TOTAL	17.0	17.0	18.00	11.00