



Pupil Assessment

Informational
Paper

32

Wisconsin Legislative Fiscal Bureau
January, 2005

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 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 of persistent gaps in performance between whites and minorities, economically advantaged and disadvantaged pupils, and males and females. 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 improve 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.

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. Tests 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 if the original norm group's scores are dated. 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 assessments 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. However, 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. Be-

cause 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 assessments 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 assessments 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.

Current Wisconsin Assessment Programs

In 1991 Act 269, a requirement was established that school districts, beginning in 1993-94, administer "knowledge and concepts" examinations in the 8th, and 10th grades and, beginning in 1996-97, administer a 4th grade knowledge and concepts examination. The tests are designed to measure a pupil's knowledge in the subject areas of mathemat-

ics, 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 3rd grade and the 4th, 8th, and 10th 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 3rd grade pupils. The WRCT is 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 3rd grade pupils identified as limited-English proficient (LEP) or children with disabilities may be excluded from taking the WRCT, in which case, these pupils must be evaluated using the WSAS alternate assessment. Pupils must be unable to meaningfully participate in the standard WRCT in order to qualify for the alternate assessment. Under federal law, alternate assessments must be based on state academic standards.

In 2003-04, 56,712 pupils took the WRCT exam. The test is currently comprised of two reading passages and three types of questions: (a) reading comprehension; (b) prior knowledge; and (c)

reading strategies. The exam, which includes approximately 100 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 in-depth understanding of the academic knowledge and skills tested; (b) proficient, meaning competent in the content area; (c) basic, defined as some knowledge in the content area; and (d) minimal, meaning very limited academic knowledge and skills 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.

Table 1 shows the statewide totals of pupil scores on the WRCT for 2003-04 according to each proficiency level.

The 2003-04 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

Table 1: 2003-04 Statewide Reading Comprehension Test Results

Proficiency Level	Number Tested	% of Total
Advanced	22,919	38.8%
Proficient	27,256	46.2
Basic	5,757	9.7
Minimal	<u>780</u>	<u>1.3</u>
Subtotal	56,712	96.0%
Excluded	<u>2,353</u>	<u>4.0</u>
Total	59,065	100.0%

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.

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 8th and 10th 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 4th grade exam was voluntary in 1995-96 and required beginning in the 1996-97 school year.

Currently, the 4th, 8th, and 10th 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 2003-04, the 4th, 8th, and 10th grade tests consisted of multiple choice and short-answer questions in language applications, reading, mathematics, science, and social studies and a writing test related to an assigned reading passage. To familiarize 4th grade pupils with test content and format, school districts are required to administer a practice activities test before the pupils take the 4th grade knowledge and concepts examination.

School boards can decide to exclude from testing limited English-proficient pupils, provided they meet certain criteria or have been enrolled in U.S. schools for less than three years. These pupils may be given an alternate assessment. Districts must include children with disabilities in the tests, with appropriate modifications where necessary or alternate assessments for those children who cannot meaningfully participate in the standard 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. Under state law, any 4th, 8th, or 10th 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 2003-04, approximately 60,349 4th grade pupils (98% of the total enrollment), 67,725 8th grade pupils (99%) and 69,132 10th grade pupils (98%) completed each subject area test. A three-week testing window is provided to allow local flexibility in scheduling for make-up testing.

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 in the same way as for 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.

Wisconsin's statewide test results for each subject area of the 4th, 8th, and 10th grade tests for 2003-04 are provided in Table 2. 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 the past, DPI has also provided national percentile rankings for each content area and grade level. National percentile ranks indicate the relative standing of a student compared with other students in the same grade in the nationwide sample. Beginning in 2002-03, Wisconsin began using a combination of off-the-shelf national test items and customized test items to improve alignment between the knowledge and concepts examinations and the state's model academic standards. This change was required by the federal No Child Left Behind Act of 2001. Customized items are not nationally normed, and therefore national percentile ranks are not available.

Federal law requires that the results must also be reported by gender, race/ethnicity, 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. Federal law requires that at least 95% of the pupils in each subgroup, and overall, participate in testing each year. The 2003-04 results of the 4th, 8th, and 10th grade exams for all pupils, by

Table 2: 2003-04 Statewide Knowledge and Concepts Exam Results (Percent of Pupils in each Proficiency Level)

	Reading	Language	Mathematics	Science	Social Studies
4th Grade					
Advanced	43%	36%	29%	21%	63%
Proficient	39%	43%	45%	59%	27%
Basic	12%	14%	10%	14%	6%
Minimal	4%	4%	14%	3%	2%
Alternate Test/Not Tested	2%	3%	2%	3%	2%
8th Grade					
Advanced	32%	27%	21%	23%	49%
Proficient	47%	40%	44%	46%	35%
Basic	11%	19%	16%	16%	10%
Minimal	8%	11%	17%	12%	5%
Alternate Test/Not Tested	2%	3%	2%	3%	1%
10th Grade					
Advanced	40%	17%	25%	34%	41%
Proficient	19%	50%	44%	36%	32%
Basic	15%	20%	13%	11%	7%
Minimal	13%	11%	15%	16%	17%
Alternate Test/Not Tested	3%	2%	3%	3%	3%

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/spr_kce.html]. Also, test results from 2003-04 are not comparable to scores prior to 2002-03, due to changes in the tests and score thresholds for proficiency that became effective in 2002-03.

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. State 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.

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 4th and 8th 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 4th and 8th grade knowledge and concepts examinations adopted or approved by the State Superintendent and the federal Department of Education must approve the examination.

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.

4th, 8th, and 10th Grade Knowledge and Concepts Examinations. Under 1999 Act 9, beginning with the 2002-03 school year, school districts must administer the state's 4th, 8th, or 10th grade examination or develop and administer its own examinations to measure pupil attainment of knowledge and concepts in the respective grades.

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 4th or 8th grade pupil unless the pupil satisfies the board's criteria for promotion.

Federal Assessment Programs and Requirements

This section provides a discussion of a national assessment program in which Wisconsin 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. The federal government covers all costs associated with administering this exam.

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.

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.

In 2003, Wisconsin participated in the reading and mathematics assessments in grades four and eight. The reading assessment used a sample of 3,048 fourth grade pupils and 2,566 eighth grade pupils. The average score for fourth graders was

221 compared to a national average of 216 on a scale of 500. For eighth graders, the average score in the state was 266, compared to a national average of 261, on a scale of 500. The mathematics assessment used a sample of 3,136 fourth graders and 2,591 eighth graders. The average fourth grade score was 237 compared to a national average of 234, while eighth graders scored 284 on average, compared to a national average of 276, on a scale of 500. Among Midwestern states, the average fourth grade score was 220 in reading, and 237 in math. Among Midwestern eighth graders, the average reading score was 266 and the average math score was 281.

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 all 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 4th and 8th 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. Previously, participation in NAEP was voluntary for states.

Additionally, under NCLB states are required to report the performance of schools and districts in making "adequate yearly progress" (AYP), as defined under Title I and measured by pupil assessments. AYP must be reported 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 in the 2001-02 school year. At

least 95 percent of each subgroup must take the assessments in order for the school to make AYP. No more than 3% of pupils statewide may take alternate assessments for pupils with disabilities or limited English proficiency. Districts receiving Title I funds must identify and sanction schools that consistently fail to make AYP for any subgroup.

If a school fails to make AYP for two consecutive years, then it is identified for improvement. The school district and DPI 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 at least 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, non-profit, 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 at least 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 implements similar oversight requirements for

states over districts as a whole.

In 2004-05, approximately 1,100 schools in 390 districts in Wisconsin will receive Title I funding totaling approximately \$160.65 million. A total of 54 Title I schools were identified for improvement (failed to make AYP for at least two consecutive years for at least one subgroup) in 2003-04 based on fall, 2003, assessment results.

Some changes to the Wisconsin student assessment system were necessary to fully comply with the NCLB. DPI indicates that additional test items were added to the knowledge and concepts examinations in all subject areas to more fully assess state model academic standards. Some standards not assessed by these examinations are instead measured and reported at the local district level for Title I accountability purposes. In addition, the knowledge and concepts examinations are now 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.

Table 3 lists the starting points and intermediate goals of Wisconsin's state accountability plan submitted to the Department of Education, as required under NCLB.

Table 3: Percent of Wisconsin Students Who Need to Score at Proficient/Advanced Annual Objectives

		Reading	Math
Starting Point	2001-02	61%	37%
	2002-03	61	37
	2003-04	61	37
Intermediate Goal (Begin new 3-8 tests)	2004-05	67.5	47.5
	2005-06	67.5	47.5
	2006-07	67.5	47.5
Intermediate Goal	2007-08	74	58
	2008-09	74	58
	2009-10	74	58
Intermediate Goal	2010-11	80.5	68.5
Intermediate Goal	2011-12	87	79
Intermediate Goal	2012-13	93.5	89.5
Goal: All Proficient	2013-14	100	100

Funding for Pupil Assessment

Table 4 provides a breakdown of total funding provided to DPI for pupil assessment programs from 2001-02 to 2004-05. The table identifies costs in three areas:

1. Printing, scoring and reporting costs. This includes the cost of the contracts with 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 2004-05, the Office of Educational Accountability within DPI consists of 11.5 authorized positions, which are directly responsible for assessment-related activities. Federal funds support 8.4 of these positions. The supplies and services budget includes items such as data processing, printing, travel, space rental, postage, conferences and consultant expenses.

Table 4: Estimated Expenditures for DPI Pupil Assessment Programs

	2001-02	2002-03	2003-04	2004-05	
Printing, Scoring, and Reporting	\$2,139,300 0	\$3,080,600 2,160,600	\$2,686,200 225,000	\$1,380,200 1,090,000	GPR FED
Development	1,381,500 0	282,300 384,100	262,700 4,737,200	1,623,900 4,428,300	GPR FED
Program Operations					
Salaries and Fringes	507,600 273,900 64,800	262,700 354,100 64,700	232,300 427,700 20,400	212,900 521,400 0	GPR FED PR
Supplies and Services	64,100 0 0	70,200 0 4,600	46,200 87,100 0	43,500 108,800 0	GPR FED PR
TOTAL	\$4,431,200	\$6,663,900	\$8,724,800	\$9,409,000	
Permanent Positions (FTE)	5.60 6.65 0.75	4.10 5.15 0.75	3.10 4.90 0.00	3.10 4.90 0.00	GPR FED PR
Project Positions (FTE)	4.00 <u>1.00</u>	0.00 <u>3.50</u>	0.00 <u>4.50</u>	0.00 <u>3.50</u>	GPR FED
TOTAL	18.00	13.50	12.50	11.50	