Agenda Topic: State Assessment System Update

Materials:
An Overview of the Development of the State Assessment System (under separate cover)
Assessing Students with Disabilities: Questions and Answers (under separate cover)

Staff Contacts: Michael Dunn
Carmen Chapman Pfeiffer
Chris Koch
Merv Brennan
Sue Easton

Purpose of Agenda Item:
• To provide an overview of the State Assessment System
• To update members of the State Board on the progress of improving the State Assessment System, including the Illinois Standards Achievement Test (ISAT), the mathematics pilot that will become part of the Illinois Measure of Annual Growth in English (IMAGE), and the implementation of the Prairie State Achievement Examination (PSAE) and the Illinois Alternate Assessment (IAA)

Expected Outcome:
Board members will be aware of the status of the State Assessment System.

Background Information:

Illinois Learning Standards and Assessment for All Children: the State Assessment System

In 1997 the Illinois State Board of Education (ISBE) adopted the Illinois Learning Standards, a compilation of specific knowledge and skills that all Illinois public school students are expected to master. Since that time, ISBE has taken steps to encourage schools to improve instruction so that all students will have the opportunity to meet the Standards. ISBE uses four major state tests to measure the progress that schools and students have made toward meeting the Standards: ISAT for elementary and middle school students (and for high school
students, on a voluntary basis for certain subjects); the PSAE for high school students; IMAGE for students with limited English proficiency; and the IAA for students with disabilities for whom ISAT or the PSAE is inappropriate.

**Beginning of the State Testing Program: The Illinois Goal Assessment Program (IGAP)**

IGAP began in 1988. Legislation mandated testing seven learning areas (reading, mathematics, writing, science, social science, physical development and health, and fine arts) at grades 3, 6, 8, and 11. IGAP began with reading at grades 3, 6, and 8 in 1988. IGAP mathematics was added in 1989, IGAP writing in 1990, and IGAP science in 1991. The Illinois State Board of Education (ISBE) reported only school and district data.

Before all subjects were folded into the state assessment system, however, the Superintendent’s Committee on Testing (chaired by a district superintendent and representing many educational partners) recommended major changes. As a result, the General Assembly passed legislation that modified IGAP to assess reading, mathematics, and writing at grades 3, 6, 8, and 10 and science and social science at grades 4, 7, and 11 and to conduct special studies in physical development and health and in fine arts. The General Assembly also limited state testing during the school year to 25 hours and required ISBE to report individual scores, beginning in 1994.

**Developing and Assessing the Illinois Learning Standards: the Illinois Standards Achievement Test (ISAT)**

Following adoption of the Illinois Learning Standards in 1997, ISBE developed the Illinois Standards Achievement Test (ISAT), and schools administered the first ISAT reading, mathematics, and writing assessments in 1999. The change from IGAP to ISAT testing took place with only one year of development and was accomplished without a break in testing. Additionally, the General Assembly mandated that reading, writing, and mathematics be tested at grade 5 under ISAT, rather than at grade 6 as for IGAP. Science and social science were added to ISAT in 2000. The State Board decided not to chart trends from IGAP to ISAT, so the 1999 ISAT established first-year baselines.

At the time that the switch from IGAP to ISAT took place, the General Assembly passed legislation that resulted in an additional grade 12 assessment, the Prairie State Achievement Examination, to be administered in January to grade 12 students and again to those students, if necessary, in March. Following the appointment of Superintendent McGee, new legislation was successfully obtained to consolidate the grades 10 and 11 ISATs with the grade 12 PSAE into one grade 11 PSAE.

Schools administered ISAT for the third time in April 2001. Every major program in ISBE uses ISAT data. All grants and other applications for dollars require ISAT data. Title 1 funding is contingent on yearly progress demonstrated by ISAT.
scores. There are 31 test sessions, 14 of which include extended response, for more than 500,000 students in the state. Reports go back to schools showing how students, school, districts, and the state as a whole have performed with respect to the Illinois Learning Standards.

**Report on the First Administration of the Prairie State Achievement Examination (PSAE)**

On April 25 and 26, 2001, grade 11 students took the Prairie State Achievement Examination for the first time. The initial reports have been quite positive: Attendance was high, and students approached the test with a positive, serious attitude. School administrators put in hours of careful planning that resulted in a well-organized test administration. We will be looking at survey data to analyze improvements for next year’s administration.

For ISBE staff, work on the Prairie State Achievement Examination (PSAE) began with the development and release of a Request for Sealed Proposals (RFSP) on April 4, 2000, to solicit bids for the services necessary to develop a PSAE system. A review of responses received by the May 4, 2000, deadline resulted in a recommendation to accept the proposal from ACT, Inc. Beginning early in June, ISBE and ACT staff engaged in a series of discussions and meetings to clarify and resolve the many outstanding issues that had been identified and put in writing by ISBE staff. On June 19, 2000, a final contract was awarded and signed.

During the next several months, ISBE staff completed the following tasks:

- Met with ACT and NCS Pearson (formerly National Computer Systems) staff to plan and work on details related to implementing the PSAE in 2001;
- Negotiated and finalized the purchase of ACT Assessment and WorkKeys test-preparation materials, which were shipped to schools in September 2000;
- Developed the 2001 PSAE Teachers’ Handbook and 2001 PSAE student-test preparation materials for distribution to schools;
- Developed materials for district and school administrator overview workshops, including a question-and-answer document and PowerPoint presentation;
- Arranged and conducted half-day overview workshops for district and school administrators from September 18 – 28, 2000, at 11 locations around the state;
- Met with the PSAE Advisory Committee to seek input on and discuss outstanding issues related to the first administration in 2001;
- Developed and disseminated a School Schedule and Site Options document to assist schools in meeting the PSAE requirements for standardized testing conditions; and
• Worked with ACT staff to finalize the procedures for schools to request test accommodations for students with disabilities on the ACT Assessment portion of the PSAE.

Later in the year, ISBE staff developed the 2001 PSAE Day 2 Supervisor’s Manual of Instructions and a question-and-answer brochure targeted to parents of Illinois students. Staff also collaborated with ACT to plan and conduct a series of 16 half-day PSAE Test Supervisor Training Workshops held throughout the state from February 26 – March 9, 2001. An audience of more than 1,400 test supervisors, back-up test supervisors, and ROE staff attended these trainings. In addition, staff responded daily to phone calls, faxes, e-mails, and letters asking all manner of questions about the PSAE.

These efforts contributed to a relatively smooth first administration of the PSAE in approximately 650 high schools on April 25 and 26, 2001.

**Update on the Illinois Measure of Annual Growth in English (IMAGE) and LEP Mathematics Assessment**

IMAGE is administered annually to students in grades 3 through 11 who do not take the ISAT or PSAE because of their lack of English-language skills. IMAGE measures reading and writing proficiency in English at three grade-level spans: grades 3 through 5, 6 through 8, and 9 through 11. Both the reading and writing tests are given in two 40-minute sessions, with students taking the test appropriate to their grade level.

IMAGE measures the progress of students with limited English proficiency (LEP) in attaining the English-language reading and writing skills needed to meet and exceed the Illinois Learning Standards. The IMAGE reading and writing tests are administered annually to those students who

• have been enrolled in a state-approved bilingual education program (Transitional Bilingual Education or Transitional Program of Instruction) for less than three years, and

• as an alternative for students to taking either the PSAE or ISAT because of their limited proficiency in English.

In 1999, ISBE staff and the IMAGE Advisory Committee developed a correlation between IMAGE and ISAT (IGAP at that time). The correlation yields predictive information on how students who take IMAGE would do if they were to take ISAT.

Last year 33,897 individual IMAGE student reports, which contain writing and reading scores for the student, school, district, and state, were generated. Students and their parents can see how they performed on the test compared to the school, district, and state. Teachers (for their classes) and administrators (for their schools) can see how students performed compared to each other, the school, the district, and the state. LEP students are permitted to take IMAGE for three years before taking the regular state assessment, so teachers and
administrators can chart students’ progress toward English proficiency by comparing their scores from one year to the next. On average, LEP students improve their scores about 25 scale-score points per year in reading and about 2 points per year in writing.

To have a Title 1 assessment system eligible for U.S. Department of Education approval, ISBE was required to add a mathematics assessment for LEP students who are not yet ready to take ISAT (federal law requires that, at a minimum, states must assess LEP student performance in language arts and mathematics). For this reason, Assessment Division staff developed and pilot-tested a mathematics assessment to measure the Illinois Learning Standards in mathematics that accommodates the language learning needs of students.

This test was developed in close consultation with Illinois educators with expertise in mathematics and language acquisition. The pilot test data will be analyzed this summer with input from our contractor and Illinois educators to determine if this will be a valid and reliable assessment of the mathematics standards for these students.

Update on the First Administration of the Illinois Alternate Assessment (IAA)

The 2000 – 2001 administration of the Illinois Alternate Assessment (IAA) was completed on April 13, 2001, which was the last day on which teachers could enter evidence in a student’s IAA portfolio. Schools mailed completed portfolios to the IAA Scoring Center in Bloomington, Illinois, from May 1 – 15, 2001. The contractor, Measured Progress, will oversee scoring of the IAA, which will take place from June through early August. Recruitment of a cadre of scorers, which will include Illinois special education teachers, is currently underway.

This year’s activities to disseminate information to educators and prepare them to implement the IAA included the following:

- From September 2000 through early April 2001, contractor and ISBE staff provided training on the IAA.
- On March 14, 2001, the Division of Intervention and Assessment conducted a live, interactive satellite teleconference that participants viewed at over 20 downlink sites across the state and via streaming Internet real video on the ISBE Web site.
- The document “Assessing Students with Disabilities: Questions and Answers” was disseminated to the field, along with a cover memo explaining how the 2001 IAA results will be reported.

ISBE staff are currently working with Measured Progress to review the 2000 – 2001 IAA and develop recommendations for refinements to the process. These refinements will be based on input gathered over the past year from the IAA Advisory Task Force and the Ad Hoc Committee, as well as teachers,
administrators, parents, and other stakeholders. Information on the recommended changes will be presented at the June Board meeting.

**National Proposals for Every Year Testing in Grades 3 Through 8**

The position of the Bush administration and current legislative proposals include state testing of all students in grades 3 through 8 in reading and mathematics. At this time, only 13 states have state tests in all of these grades. Testing at these additional grades would require additional funding at both the national and state levels.

In addition to the increased testing as part of the Title 1 requirements – if legislation now pending in Congress is passed – the National Assessment of Educational Progress (NAEP) will be used to validate state assessments. The proposals for NAEP include testing every year at grades 4 and 8 in reading and mathematics. If Congress mandates this testing, the Illinois testing program would have to add at least grade 4 reading and mathematics to make direct comparisons.

**Overview**

1. **Test Development: the Illinois Approach**
   
   The steps in developing ISAT, IMAGE, and the ISBE-developed tests of the PSAE are similar and follow procedures accepted by the industry for use in the development of standardized tests. The following description of test development focuses on ISAT, but the ISBE Assessment Division follows the same steps in all the tests it develops.

   1A. **Relationship of the Illinois Learning Standards to state assessment**

   The Illinois Learning Standards comprise the written expectations that Illinois students are expected to satisfy in the subjects of language arts, mathematics, science, social science, physical development and health, and fine arts. The first step in constructing tests based on the standards is to identify the components of each standard that can be tested in a statewide assessment. For example, Standard 14C, “Understand election processes and responsibilities of citizens,” can be assessed by a multiple-choice, question. Many of the Standards overlap and are necessarily connected. Thus, for assessment individual Standards are clustered into sets in each subject and questions are developed to assess these sets.

   2A. **Creation of the state tests**

   ISBE’s Assessment Division oversees development and implementation of state tests. The task requires the collaborative efforts of Illinois educators and three major contractors.
Test development, which includes the writing of new items and construction of tests, is carried out at this time jointly by MetriTech, Inc. of Champaign, Illinois, and the Assessment Division, which call on classroom teachers, curriculum directors, administrators, editors, and content and education specialists throughout the year for item writing, selection of test items, and review of items, reading passages, writing prompts, and completed tests.

Test production, which includes printing and administering tests, scoring multiple-choice items, and processing and distributing test data, is carried out at this time by NCS Pearson of Iowa City, Iowa.

The scoring of writing samples and other extended responses items is carried out at this time by Measurement, Incorporated, of Durham, North Carolina

3A. Cycle of test development

Three years is the ideal schedule for the production of each state census test (referred to as census tests because they are given to each student eligible to take them). Items for the test to be given in 2002 are

- written in 1999,
- tried out in 2000, and

Simultaneously, the preparations for the tests given in other years proceed as well. Thus, in any given year, all phases of test construction are performed. This continual preparation takes places for each grade and subject area for which there is a census test. ISAT includes seven subjects tested at five grade levels.

4A. Scoring the state tests

After schools ship the tests back, the multiple-choice and extended-response items are mechanically separated.

- The multiple-choice answer sheets are machine-scored by NCS Pearson.
- Essays and extended-response items are sent to Measurement Incorporated for individual scoring by trained readers.

Measurement Incorporated hires about 1,300 staff members to read and score approximately 3 million responses, a process that takes at least 6 weeks. Scorers must be knowledgeable in the subject and have at least a bachelor’s degree (about a third have advanced degrees). The scoring process incorporates quality control procedures that result in scores with a high rater agreement (> 0.85).

5A. Data checking

Before scores are reported back to schools, NCS Pearson must process the multiple-choice scores, and Measurement Incorporated must score the
extended-response items and send those scores to NCS Pearson. NCS Pearson then sorts the data and assembles them into school and district scores and individual student reports. If all goes well, the necessary steps from testing to reporting require about 3½ months. This length of time is necessary because of the complexity of the state tests and to ensure accuracy of reported results. ISBE staff and the ISBE contractors check the correctness of the ISAT results in a number of ways.

- All tests are reviewed several times by ISBE staff before being printed to ensure that scoring keys are correct.
- The correctness of the scoring keys is rechecked once the tests are returned to the scoring contractor by reviewing the answers of the highest and lowest scoring students.
- Before any ISAT reports are send out to districts, schools, or students, ISBE staff checks the data of a small number of school and districts. This involves doing by hand the computations that the computer performs, a process that takes several weeks.
- Finally, schools and districts are asked to verify the preliminary data and key pieces of information. Typical problems uncovered are miscoding students by category (ethnic, gender, disability, and other classifications), erroneously assigned grade levels, and missing data.

ISBE and the printing contractor correct any erroneous information that schools make known to them. Finally, ISBE uses information about the number of students in each school by grade (supplied by schools in November of each year) to check each school’s rate of participation in ISAT and to follow up with schools having low participation rates.

2. **Logistics of Statewide Testing**

The logistics of giving a state test on the enormous scale required by Illinois are complex. The final version of the census test is typeset, printed, packaged, and shipped to schools by NCS Pearson.

- The schedule must be set at least six months in advance so that this work can be completed on time.
- ISAT administration is accomplished during a two-week period in early April. It is important that all Illinois students sit for the test during the same limited time period so that statewide scores can be compared.
- The number of test takers ranges from about 140,000 in grade 3 to 110,000 at grade 11. (In 2000, about 550,000 students were tested.)
- Schools must return tests to NCS Pearson by the established deadline so that scoring can be completed within the scheduled time frame and so that test security is not compromised. Extended-response scoring can be seriously delayed if papers are not available as scheduled.

3. **Test Security**
Test irregularities may affect a school’s designation: The proposed school designation model will establish criteria for inclusion of schools and districts on the Academic Early Warning List. Part of the compliance prerequisites for this new model will include sanctions for cheating or other inappropriate practices. For this reason, the Assessment Division has prepared and distributed a document, Professional Testing Practices for Educators, to ensure that administrators and teachers are aware of testing practices that if followed will result in standardized testing conditions. The eight-page document summarizes required, prohibited, and strongly recommended practices.

Only educators who have become familiar with the testing practices that are described in this document will be able to ensure that their students take the ISAT under fair and equitable conditions. To achieve this goal, Illinois public school educators are expected to read the testing practices document carefully and to adhere strictly to the principles and policies it describes. Superintendents and principals are required to sign a test security and confidentiality document to indicate that they have taken sufficient and appropriate steps to ensure that all who handle secure test materials in the school are fully aware of their responsibilities regarding test security and confidentiality and are properly trained to handle these responsibilities.

4. Technical Quality of ISAT: Reliability and Validity

Illinois census tests are equal to or better than commercially produced tests with respect to technical quality. Table 1 shows that the reliability of the ISATs ranges from 0.87 to 0.95, values that are equal to those of commercially available tests. For instance, the reliability of the California Achievement Test ranges from 0.84 to 0.96 for the reading subtests and from 0.84 to 0.96 for mathematics.

Table 1: ISAT Reliability

<table>
<thead>
<tr>
<th>Grade</th>
<th>Reading</th>
<th>Mathematics</th>
<th>Writing</th>
<th>Science</th>
<th>Social Science</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0.95</td>
<td>0.94</td>
<td>0.95</td>
<td>0.93</td>
<td>0.87</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.92</td>
</tr>
<tr>
<td>5</td>
<td>0.93</td>
<td>0.91</td>
<td>0.95</td>
<td>0.94</td>
<td>0.90</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.91</td>
</tr>
<tr>
<td>8</td>
<td>0.93</td>
<td>0.91</td>
<td>0.95</td>
<td>0.95</td>
<td>0.90</td>
</tr>
</tbody>
</table>

Also, there is solid evidence for the validity of the ISATs relative to commercial tests. The correlation between students’ scores on the ISAT reading and mathematics tests with their scores on the Stanford Achievement Test (SAT9) are shown in Table 2. (Note: The SAT9 is administered to a sample of Illinois
students to determine the percent of students in each national SAT9 performance quarter).

<table>
<thead>
<tr>
<th>Grade</th>
<th>Reading</th>
<th>Mathematics</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0.84</td>
<td>0.86</td>
</tr>
<tr>
<td>5</td>
<td>0.81</td>
<td>0.87</td>
</tr>
<tr>
<td>8</td>
<td>0.80</td>
<td>0.83</td>
</tr>
</tbody>
</table>

Table 2 shows that correlations exceed 0.83, indicating that ISAT and SAT9 scores measure highly similar concepts in reading and mathematics. The ISAT, however, is directly linked to the Illinois Learning Standards, so the ISAT is preferred for state accountability purposes.

5. Getting Scores Back: How Scores Are Used (or Timing of Assessment: the Best Time for Testing)

Schools have two irreconcilable desires: (1) to test students as late as possible in the school year and (2) to have test results before the end of the school year.

The turn-around time for test results is a major problem in statewide testing across the country. Many people demand a quick return of results in the mistaken belief that all tests are the same and that data can be returned equally quickly from any standardized test.

Several factors contribute to the length of the interval between testing and reporting for Illinois state assessments:

- Enormous scale of the testing – more than 700,000 students are tested each year.
- Complexity of the tests, which include extended responses in reading and mathematics as well as writing samples. After NCS Pearson receives completed answer documents, it must separate extended-response pages from multiple-choice pages and send each part to the proper location for scoring.
- Scoring of extended-response items and writing require more than six weeks and take place a multiple centers throughout the country.
- Delay in return of completed tests by some schools. With more than 3.5 million test booklets to score, NCS Pearson cannot complete even the machine scoring in a timely manner unless schools ship the completed tests on the day the last test is taken.
Data compilation, checking, and equating. After scoring is completed, NCS Pearson must match the extended-response scores with the multiple-choice scores for each student. Numerous procedures are in place to check the accuracy of the generated data.

Finally, test scores must be equated so that they are comparable from year to year.

ISAT is only one part of the testing system that schools use for planning their curricula. It is not designed primarily as a diagnostic test for individual students. In fact, a single test cannot be – and should not be – used for diagnostic purposes. True diagnosis, whether it is in education or the medical field, requires many types of assessment to determine what may be needed to improve academic or physical health, respectively. ISAT provides accountability at the school, district, and state levels, and its purpose is to guide the program-development efforts of schools and districts. ISAT provides reliable data that can be compared from year to year, which schools can use to assess the effectiveness of curricular modifications.

**Superintendent’s Recommendations:**

To support the assessment staff and education center in providing and revising the state testing program through a model of continuous improvement.

**Next Steps:**

- Continue to investigate and work toward extending testing in grades 3 through 8
- Continue to collect data to improve ISAT, IMAGE, the IAA, and the PSAE
- Connect the proposed assessment changes to the designation system
- Use the data to improve PSAE administration with minor adjustments rather than major changes to ensure continuity and planning.
- Finalize recommendations for refinements to the IAA and present this information to the Board at the June meeting.