

UTAH FOUNDATION

Research Report

Report Number 629

December 1999

Accountability in Public Education: An Overview and Analysis

Highlights

- Since a Nation at Risk was published in 1983, U. S. presidents, congress, governors, legislators, business people, parents and educators have been debating, passing laws, revising curriculum, raising graduation requirements, and testing students in hopes of improving the performance of the nation's public school children.
- During this period of reform, school *accountability* became a popular term. Many reformers argue that improvement in public education cannot be expected unless schools and students are held accountable.
- Over the years, several ways of making schools more accountable have been tried. These include accountability through: performance rating; monitoring and compliance with standards or regulations; incentive systems; reliance on the market; changing the locus of authority or control of schools; and changing professional roles.
- Recent studies indicate that the important components of successful accountability systems include: clear content and achievement standards; valid testing of students to determine if they meet those standards; reporting of student test results and other characteristics of individual schools which impact student learning and achievement; a system of rewards for schools which meet the standards and interventions for schools which do not; and adequate funding of all elements of the system.
- Utah's 1999 Legislature created a thirteen member Task Force on Learning Standards and Accountability in Public Education. The task force was given the two-year assignment of studying student performance standards and accountability programs and recommending an accountability program to the Legislature. The task force has proposed legislation for 2000 which will create a five-part assessment program to be known as UPASS (Utah Performance Assessment System for Students). The five components of UPASS are:
 1. Systematic norm-referenced achievement testing of students in grades 3, 5, 8 and 11.
 2. Criterion-referenced achievement testing (CRT) at the end of every grade and basic skills course. Constructed response sections are to be added to the CRTs.
 3. A tenth grade basic skills competency test.
 4. Writing assessments for students in grades 6 and 9.
 5. Use of student behavior indicators in assessing student performance.
- The legislation also mandates that the State Board of Education develop a school performance report. Next year the task force will face its greatest challenge - setting a performance standard and establishing a system of rewards and interventions.
- A critical challenge for the legislature will be to find adequate funding for UPASS, the improved school report cards, and, in the future, a system of rewards and interventions.
- A well-designed accountability program could be instrumental in ensuring that students leave Utah's public education system with the basic skills necessary to participate successfully in society. If an accountability program is to be successful in Utah, all participants in the program and the public at large must understand the goals and objectives of the program and be confident that the costs (both fiscal and educational) are acceptable.

UTAH FOUNDATION is a private, nonprofit public service agency established to study and encourage the study of state and local government in Utah, and the relation of taxes and public expenditures to the Utah economy.

Accountability in Public Education: An Overview and Analysis

Since a *Nation at Risk* was published in 1983, U.S. presidents, congress, governors, legislators, business people, parents and educators have been debating, passing laws, revising curriculum, raising graduation requirements, and testing students in hopes of improving the performance of the nation's public school children.

At the national level, President Reagan made the Department of Education a cabinet level post, President Bush established educational goals for the year 2000, and today President Clinton is pushing congress to provide funding to reduce class sizes.

States across the country have raised graduation requirements and required students to pass basic competency tests in order to graduate. Thirty-five states have passed charter school programs, 18 states have public school choice (open enrollment) statewide, a few states have a voucher program, three states provide education tax deductions or credits for private education.

During this period of reform, school *accountability* became a popular term. Advocates of reform argued that improvement in public education cannot be expected until schools are held accountable for educating all students. To address this issue, states mandated school report cards, publication of school test scores, drop-out and graduation rates, among other things.

Utah has not been sitting on the sidelines during this era of reform. Below is a selected list of reforms implemented since the publication of *A Nation at Risk*.

- ! 1984 - State Board of Education asks State Office of Education to identify K-12 specific core curriculum standards necessary for graduation from secondary school.
- ! 1987 - State Board of Education adopts a statewide core-curriculum for K-12.

- ! 1988 - Increased graduation requirements.
- ! 1992 - Legislature mandates statewide open-enrollment effective September 1993.
- ! 1990 - Began annual statewide testing of all students in grades 5, 8, and 11 on the norm-referenced Stanford Achievement Test. In 1999, the third grade was added to the annual testing.
- ! 1992 - Strategic Planning Act for Educational Excellence passed by legislature.
- ! 1993 - Centennial Schools Program begins with 98 schools.
- ! 1994 - Utah's State Strategic Plan approved by U.S. Department of Education as Utah's preexisting Goals 2000 Plan.
- ! 1995 - Legislature passes Highly Impacted School Program.
- ! 1996 - Legislature passes Modified Centennial School Program.
- ! 1998 - 21st Century School Program begins. State School Board approves eight Charter Schools.
- ! 1999 - Annual criterion-referenced testing begins at all grade levels.
- ! 1999 - Legislature creates a task force on Learning Standards and Accountability in Public Education with the responsibility to develop a statewide accountability program. As can be seen by this partial overview, some

significant efforts have been made and continue to be made to improve the state's public school system. Higher graduation requirements, statewide testing on norm-referenced and criteria-referenced tests and definition of a core curriculum are generally recognized as important steps forward. The development of the Highly Impacted School Program shows appreciation of the special challenges faced when educating at-risk students. Charter schools and 21st Century Schools encourage excellence and innovation in schools.

Despite these efforts, both the governor and the Legislature continue to keep education issues on the front burner. Governor Leavitt recently proposed a 6.0 percent increase (for FY 2001) in the Weighted Pupil Unit, the basic funding formula for public education. This is the largest increase in several years. The 2000 Legislature will have several education bills to consider, one of which is the result of the work of the Task Force on Learning Standards and Accountability in Public Education.

The purpose of this report is to provide an overview of the issue of school accountability, outline what seems to be a developing consensus about what a good accountability system looks like, and look at the recommendations of the Task Force. However, first it will be helpful to have a brief discussion of what we know about how Utah's public education system is currently performing.

What do Utah Test Scores Tell Us About Utah Students?

Each year Utah Foundation produces a report entitled *Utah Statewide Testing Results*. The main focus of the report is on the annual statewide testing of public school students in the fifth, eighth and eleventh grade. This testing program has been conducted since 1990 as mandated by the 1989 Legislature. The State Office of Education has

used a norm-referenced Stanford Achievement Test for the entire time. In addition to analyzing this test, the Utah Foundation report has looked at other tests such as the ACT and SAT, which college bound students take generally in their senior year, and the National Assessment for Educational Progress (NAEP) test. For a complete analysis of these test results, the reader is referred to the Foundation reports.¹ Following is a brief discussion of the annual Stanford Achievement test and the ACT and SAT test taken by Utah's college bound students as these tests provide sufficient data to form an opinion on how Utah's public schools are doing.

Stanford Achievement Test - 1998 Results

Each year since 1990, the Utah State Office of Education has tested public school students in grades 5, 8, and 11.² From 1990 through 1996 the Utah State Office of Education used the Eighth Edition of the Stanford Achievement Test and from 1997 through 1999 they used the Ninth Edition. Both are norm-referenced tests.³ During the seven years of using the Eighth Edition, the median composite battery test score (reported as

¹ Utah Foundation, *Utah Statewide Testing Results 1998*, (Report 623, April 1999).

² Utah Foundation has published the results of these tests each year. Copies of the annual reports are available at Utah Foundation.

³ A norm-referenced achievement test is designed so the results of the test can be readily compared to students nationally. When the test is designed, it is calibrated to a representative national sample of students (the norm group) for each grade and subject. School results of the test are reported as "median national percentile ranks." The percentile rank refers to what percentage of the national norm group of students achieved results below that of the median student in the schools. Therefore, scores above 50 are considered above the national average and scores below 50 are below the national average.

percentiles as explained in footnote 3) for Utah's fifth grade students fluctuated narrowly between 53 and 55, with the last three years (1994-96) all being 53. In the three years of using the Ninth Edition, the complete battery test score median has fallen to 50 and stayed there.

From 1990 to 1996, Utah eighth grade students had median composite battery scores narrowly ranging between 50 and 53. In the Ninth Edition, Utah's eighth grade students scored a 54 in 1997 and 1998 and then fell to 53 in 1999. Eleventh grade student complete battery test scores on the Eighth Edition were consistently higher than either the fifth or eighth grade scores. Eleventh graders scored between 53 and 56, with 1995 and 1996 being at 56. On the Ninth Edition, Utah's eleventh graders have scored higher. In 1997, they record a median score of 60. Unfortunately eleventh grade scores have fallen for the last two years. In 1998 the median score was 58 and in 1999 a 57. Though declining, eleventh grade complete battery scores are the state's highest with the typical Utah 11th grade student scoring better than 57 percent of students in the national norm group.

While looking at the statewide results is informative, it is also important to look at the results for individual schools and districts. This closer look shows what the statewide averages do not -- that there are significant differences in test scores between schools in different districts and among schools within districts. This is especially true when looking at elementary schools. Elementary school boundaries are small and therefore the socioeconomic characteristics of the students are less diverse than in junior or senior high schools. National studies show that standardized test scores are highly correlated to socio-economic indicators. Looking at elementary school scores provides the opportunity to see how students of similar socio-economic background are doing.

A good way to look at the disparity among schools is to see the school test scores of the state's second largest school district. The Granite School District is an inner-urban school district with about 73,000 students. The district has 62 elementary schools, of these, 26 schools (43 percent) had median composite scores of 40 or below, another 26 had scores between 41 and 59. Ten schools had scores of 60 or higher. Clearly, Granite School District has a high percentage of low performing schools, with an equal amount in the middle range.⁴

Of the 26 schools scoring 40 or below, 18 have 25 percent (or higher) of low income students⁵, 12 of the schools have 40 percent or higher. Of the ten schools with test scores of 60 or higher, nine have 14 percent (or less) of students living in low income households. These data show that the Granite School District has some very high performing schools and some very low performing schools. It also seems clear that socio-economic factors play a significant role in test scores in Utah as has been suggested in national studies. Such important data about how Utah schools are performing is hidden in the statewide averages.

American College Testing Program (ACT)

The most widely used national test that Utah high school seniors voluntarily take is the test administered by the American College Testing (ACT) program. About 28 states use the ACT as the standard college entrance exam. In 1999, there were 22,190 Utah students (68 percent) that took the ACT. Similar high percentages of Utah

⁴ For a more complete analysis of this issue see Utah Foundation, *Utah Statewide Testing Results 1998*, (Report 623, April 1999).

⁵ Low income is defined as the percent of the student body who qualify for free or reduced prices meal programs.

seniors take the exam every year. This is a much higher percentage of students than take the ACT nationally. A larger than normal group of test-takers in Utah means that the state's average scores are probably lower than they would be if the same percentage of students were to take the test in Utah as do nationally.

Over the last 11 years, Utah college-bound seniors have consistently scored above the national average. In 1988, Utah's average score was 20.9, while the national average was 20.8. In 1999, Utah senior's average score was 21.4, while the national average was to 21.0. It appears that those Utah students planning on college are improving their position in relation to the college-bound seniors nationwide. According to the State Office of Education, "The large percentage of Utah students taking the test makes the overall high scores even more impressive."⁶

Scholastic Assessment Test (SAT)

Another college entrance exam is the Scholastic Assessment Test (SAT)⁷ administered by the College Entrance Examination Board. It is not required by any of Utah's colleges or universities, but is the primary college entrance examination used by many of the nation's elite colleges and universities. For these reasons, fewer Utah students take the SAT (about 4 percent) than nationally and those that do are generally a select group of students aiming for the better out-of-state colleges, acknowledges the State Office of Education.

Being taken by Utah's top students, it would be expected that Utah's scores should be better

than national scores and they are. Utah's 1999 SAT scores were again well above the national average. Utah students had an average verbal score of 570 and a mathematics score of 568. By comparison, students nationally had a verbal score of 505 and mathematics score of 511.

What do all these measurements mean?

What then does the data show regarding how Utah schools are doing? The annual Stanford Achievement Test indicates that on average Utah fifth grade students performed at the national average. In the eighth grade median scores are slightly above the national average. In the eleventh grade, Utah students perform the best of the three grades and several points above the national average.

As mentioned previously, these statewide averages are important measurements but they do mask the significant range of school scores that exist in the state, especially in inner-city schools. Furthermore, the test scores show no significant improvement over the eight years the Sanford Achievement Test has been administered in how Utah students compare to students nationally.

This does not mean that Utah students in 1998 are not doing better than their predecessors in 1990⁸. It must be remembered these are norm-referenced tests and the Eighth and Ninth Editions of the test are very different. For each edition, the test measures how students taking that particular test compare to the norm group who took the same test. When the average median score in Utah is above 50, it indicates that the median

⁶ Barbara J. Lawrence, "Utah ACT Performance for 1998-99," (Utah State Office of Education, August, 1999, p1).

⁷ This should not be confused with the Standard Achievement Test (SAT) used in the annual statewide testing program in grades 5, 8, 11.

⁸ Norm referenced tests are designed to spread scores over a curve. If a particular test question becomes too easy (too many students are able to answer it correctly), that question is made more difficult on subsequent tests so that the scores will once again be spread across a curve. Norm referenced tests generally become more difficult over time.

student in Utah performed better than half or more of the students in the national norm group.

Given the low levels of per pupil expenditures and large class sizes, on the one hand, and state average test scores generally above national averages, on the other hand, the case can be made that Utah's schools are doing an adequate job given their limited resources. As mentioned, these state averages do mask the fact that some Utah schools are performing significantly below the national average and, in addition, Utah has a relatively small minority population and low levels of poverty compared with the nation as a whole. Since both of these factors tend to be closely related to low test scores, one could argue that Utah's test scores should be higher than the national average given Utah's more homogeneous demographic make-up.

If Utah's demographic make-up has helped Utah teachers succeed in an environment of large class sizes and low per pupil funding, that advantage is beginning to dwindle. In the 1990s, Utah's public school enrollments have become more racially and ethnically diverse. Hispanic, Asian/Pacific Islander, black and American Indian students now make up 11.3 percent of Utah's student population as opposed to 7.4 percent in 1990. The increase in diversity is one of many factors which educators must adjust to as Utah enters the new century.⁹

The current bottom line is that even though some Utah schools are scoring lower than the national average, as a whole the median Utah student performs at or slightly above the national average. Utah college bound students score quite well in relation to other college bound students. Given these facts, what benefits might come from instituting an "accountability system" in Utah?

⁹ For a more detailed discussion of this issue see, "Utah Statewide Testing Results 1998," *Research Report*, 623, Utah Foundation.

What is accountability?

The Issue of Accountability

The idea of schools being accountable (able to prove that they are doing the job of educating students) is not a new idea. In fact it is a very old one. In 19th century England, schools were administered under an incentive system called "payment by results." School inspectors gave standards tests to each student and the schools were paid based on how well the students performed.

Criticism developed quickly over this approach to education accountability because it was claimed that teachers began dropping instruction in disciplines that were not on the test. The state of New York, in 1879, implemented the "Regents exams" with the idea that such testing would improve accountability. In the 1920s, with business at the pinnacle of respect, business cost-accounting techniques began being applied to school systems in order to develop greater accountability for the fiscal resources schools received.

For many years, schools in the United States were held accountable, being judged mainly on the process by which children were educated. The idea was that if good text books were being used, adequate time was allowed for instruction, etc., educators were doing what they should be.

In 1970, with the publication of the book *Every Kid a Winner* by Leon Lessinger, the discussion of accountability in education took on a revised meaning. The author, associate commissioner of the U.S. Office of Education, argued that learning should be measured in quantifiable terms that could be related to cost statements. Lessinger stated that "instead of certifying that a student has spent so much time in school or taken so many courses, the schools should be certifying that he is able to perform specific tasks. Just as a warranty certifies the

quality and performance of a car, a diploma should certify a youngster's performance as a reader, a writer, a driver, and so on..."¹⁰

Lessinger promoted the idea that the outcome of efforts to educate students was a better way to judge the value of the system. He argued that if a good process was not achieving the desired results, the process should be changed.

Following the publication of a *Nation at Risk*, in 1983, a new wave of accountability initiatives began. Since then, several approaches to accountability have developed. They can be summarized as follows¹¹:

- *Accountability through performance rating* - This includes such measures as statewide assessments, school and district report cards, and performance indicators. This accountability component is particularly important because all other accountability approaches rely to some extent on this process of making information readily available to the public.
- *Accountability through monitoring and compliance with standards or regulations* - This is an auditing and budget review approach the focus of which is compliance with set standards of performance which outline what should be learned and how well. This approach relies on performance reporting.
- *Accountability through incentive systems* - The main concept here is reward for

¹⁰ Michael W. Kirst, *Accountability: Implications for State and Local Policymakers*, Office of Educational Research and Improvement, U.S. Department of Education, 1990), p. 3.

¹¹ Michael W. Kirst, *Accountability: Implications for State and Local Policymakers*, p.7-10.

results. Generally specific levels of accomplishment are established and rewarded when they are achieved. Teacher merit pay, bonuses, school awards, and recognition are examples.

- *Accountability through reliance on the market* - This approach relies on competition to create accountability with the options including open enrollment among the public schools, introducing education alternatives such as charter schools and magnet schools, and allowing vouchers and tuition tax credits for students who attend private schools.
- *Accountability through changing the locus of authority or control of schools* - The argument here is that schools will become more accountable when the locus of authority is changed. Parent-advisory councils, community-controlled schools, even state take-overs or privatization are the proposed options.
- *Accountability through changing professional roles* - This more recent approach argues for teachers reviewing each other for tenure or dismissal like universities do. Another approach is for experienced or outstanding teachers to help colleagues who are judged ineffective. Devolution of educational policy decisions are also part of this approach. Here site-based education policy is made with the teachers playing a significant part in the development of the policy.

There are numerous approaches to implementing greater school accountability. Many of the six approaches mentioned can be used in combination. Several recent studies of the various

accountability programs have come to the conclusion that there are certain elements which help make accountability programs successful.

Important Elements of an Accountability System

According to these studies, the important components of successful systems include:

1. Clear content and achievement standards.
2. Valid testing of students to determine if they meet those standards.
3. Reporting of student test results and other characteristics of individual schools which impact student learning and achievement.
4. A system of rewards for schools which meet the standards and interventions for schools which do not.
5. Adequate funding of all elements of the system.

Content & Achievement Standards

Since 1987, the Utah State Office of Education has worked to develop clear content standards, that is to establish a core curriculum. Utah has a core curriculum which outlines what students should learn in each subject at each grade level. This core curriculum has been evaluated by independent researchers and has received passing marks. While there is room for some improvement and updating in the future, Utah's core curriculum outlines the goals of Utah's public schools as they educate students. Having this in place puts Utah one step ahead in the development of an accountability system.

However, Utah has not established what an acceptable or passing level of student achievement should be. This is the level of achievement all Utah

students should be able to reach in each grade or before they leave the public school system. The standard must be linked to what is being taught in the classroom, namely, the core curriculum. Establishing the standard is a necessary step Utah will have to take.

Valid Testing

Testing is the traditional way student achievement has been assessed. Accountability programs depend on various assessment instruments (tests) to determine what students know and can do. These tests must be valid and reliable. Validity means that a test measures what it is supposed to measure. Reliability means that the results from the testing could be closely if not exactly reproduced by administering the test a second time under similar circumstances. Proper development and piloting of any assessment instrument used to measure student achievement is critical to the success and credibility of an accountability system.

As mentioned, since 1990, Utah students in fifth, eighth, and eleventh grade have taken the Stanford Achievement Test. Last year, the Legislature added an additional grade (third) to the norm referenced test group, required criterion referenced tests (CRT) to be given in all grades in basic skill courses, and asked the State Office of Education to develop a tenth grade basic skills test. The reasons for these additions will be discussed in detail later in the report. The norm referenced test is considered valid and reliable and has a nine year history. Utah's CRTs have been reviewed by an independent agency and found to be adequate with some minor changes. This is a good beginning as Utah tries to assess just what students know and can do.

Reporting

For schools to be accountable, test results must be made public for each school. Other information about the school which might have an

impact on student performance or which reflects other measures of student achievement must also be given to the public. In Utah, school districts have been releasing an annual performance report since 1990 which contains some of this information. When these reports are coupled with some form of school choice as they are in Utah, parents are empowered to select good schools for their students.

However, district-wide information, while helpful, does not give enough detail to promote individual school accountability or truly informed school choice. Reports which detail performance at individual schools are preferable. Current reporting requirements also lack uniform data collection and reporting specifications.

Rewards & Interventions

A system of rewards and interventions is one of the most important elements of an accountability system. Schools who meet the established standard should be rewarded. These rewards can be symbolic (positive rankings and public recognition), monetary or a combination of both. Various types of monetary rewards have been tried. Generally, either all teachers in successful schools receive bonuses or the school receives a sum to be used to improve facilities and resources. There is currently no evidence that one type of reward is more effective than another in motivating schools to be successful.

Once schools not meeting the standards are identified by the accountability program, the state must decide how to help those schools meet the standards in the future. Many types of interventions have been tried in other states with varying degrees of success. Current research does not show that one system of intervention is best but does show that interventions work best when they are seen as offers of help rather than as punishment.

Outcome-based accountability systems are very new and are based on the premise that all

students will meet higher standards of achievement. Schools must be given the training they need to meet those goals. Interventions must include ample opportunities for teachers and administrators to learn how to use test results to adjust instruction, learning improved methods for working with students who are struggling, and preparing the school to function successfully once the intervention period is completed.

Students who are struggling may need to be taught in different ways, receive more one on one instruction or tutoring. Classes teaching critical subject areas may need to be smaller. Regardless of what the components are of the system of rewards and interventions, they must be fair and consistent and based on clearly understandable rules.

Since Utah has not established a performance standard, to date, no consistent attempt has been made to reward or intervene in schools which are performing well or poorly. Once a performance standard for Utah students is established, a fair and consistent reward/intervention program must be the next step.

Adequate Funding

The reason many states have adopted and other states are considering accountability programs is that citizens want to know that their education tax dollars are being well spent. Accountability programs help citizens see just what student outcomes are being achieved with their education tax dollars.

Proponents of accountability programs would argue that when funds for education are limited (as they are in Utah), it is especially important that those funds are producing the desired outcomes. However, accountability programs cost money. Testing, rewards, and interventions can carry a significant price tag. Each important element of an accountability system must be adequately funded if the system is to be successful.

With a basic understanding of the important

elements that should be included in an accountability program, it is now possible to look more closely at what Utah's task force is recommending.

Task Force on Learning Standards and Accountability in Public Education

The 1999 legislature created the thirteen member Task Force on Learning Standards and Accountability in Public Education. The task force is made up of three members of the Senate, five members of the House of Representatives, the governor or his designee, two members of the State Board of Education and two local school board members. The task force was given the two-year assignment of studying (with input from groups including education professionals, colleges and universities, students and parents) student performance standards and accountability programs already in place and "proven to be successful," measurable objectives of student performance that can be validated and "proven education systems that have successfully incorporated standards, testing, and local autonomy to raise student achievement."

From this research, the task force is to make recommendations to the legislature that will enhance student achievement in every public school in the state by adopting learning standards, implementing assessment methods tied with the state learning standards, outline a program for assisting schools which do not meet the learning standards, and recognize and reward schools which do meet the standards. The goal stated in the legislation was to have "an effective statewide standards and accountability program in public education functioning by July 1, 2000." The task force was asked to issue two reports. The first report to be due no later than November 30, 1999 and the second, due no later than November 30, 2000.

Members of the task force met throughout

1999. In addition to their regular meetings the task force held a two-day symposium during which they heard testimony concerning accountability programs in other states. In September and October of 1999, they held a series of public hearings throughout the state where citizens were allowed to comment on a conceptual outline of an accountability program presented by the task force. The task force put in a lot of time reading, studying, and listening and are to be complimented for the thoroughness of their investigations into what should and is being done with accountability programs in other states and what Utah citizens think about accountability in education.

Proposed House Bill 144

At the end of November, the task force presented its first report to the Legislature's Education Interim Committee along with proposed legislation in the form of House Bill 144. The focus of this first report and the accompanying legislation is the assessment portion of the proposed accountability program. The legislation outlines a five-part assessment program to be known as UPASS (Utah Performance Assessment System for Students). The five components of UPASS are:

1. Systematic norm-referenced achievement testing of students in grades 3, 5, 8 and 11 (third grade testing begins in the fall of 2000).
2. Criterion-referenced achievement testing (CRT) at the end of every grade and basic skills course (beginning in the fall of 2001). Basic skills courses include reading, language arts, math and science. Constructed response sections are to be added to the CRTs as a pilot during two school years (2001-03) and to be included on all tests 2003-04 and beyond.

3. A tenth grade basic skills competency test (beginning with the 2002-03 school year).
4. Writing assessments for students in grades 6 and 9 (beginning 2001-02).
5. Use of student behavior indicators in assessing student performance (beginning in 2001-02).

Since 1990, the state has been administering and reporting on the norm-referenced testing of students in grades 5, 8, and 11. In 1999, the legislature also passed House Bill 33 which added the third grade to the norm-referenced testing program, required statewide administration and reporting of CRTs in all grade levels and courses in basic skills areas, and instructed the State Board of Education to develop the tenth grade basic skills competency test. CRTs were added to the state testing program because they are designed to test student mastery of Utah's core curriculum. The tenth grade basic skills test will ensure that Utah students who receive a high school diploma have mastered the desired basic skills and will also motivate students to reach that goal. All of these are multiple choice tests.

One objection to multiple choice tests is that they are, by their nature, not able to measure some important aspects of student performance. Writing ability is one of these. The constructed response sections to be added to the CRT under House Bill 144 and the sixth and ninth grade writing tests would provide valuable information on student writing skills. This is an area where the current norm-referenced testing indicates students need improvement. The constructed response questions also yield more information on comprehension, problem solving, critical thinking ability and other educational goals than do the multiple choice tests. Studies of assessment programs indicate that when multiple measures (types of tests, samples of student work, etc.) are

used, a more accurate picture of what students know and can do can be obtained.

In addition to establishing UPASS, House Bill 144 requires that school districts report their CRT scores on their annual performance reports beginning with the report to be issued in January 2001 for the 1999-2000 school year. The State Board of Education, in collaboration with the school districts, is to develop a school performance report containing information on each school in each school district which will include many of the measures currently included on the school districts annual performance reports plus other measures (some of which are to be determined during the second year of the task force).

The board will design data collection forms and procedures, collect and compile the data. The information will be sent in report form to the districts who are responsible for distributing the information to the residence of each student. The first reports will be issued during October 2002 for the 2001-02 school year. This report will take the place of the school district annual performance report and will provide information on a school by school basis and will include District and State aggregated totals.

The final section of the bill outlines the work that remains for the task force during its second year. Among the items the task force will continue to study are:

1. Recognition and rewards to schools and school districts who display exemplary student performance or show significant improvement or gains in student performance,
2. Interventions, including identification of resources to assist schools whose students are not achieving acceptable levels of performance,

3. Determining what an acceptable level of performance is,
4. How best to in-service teachers and administrators to maximize the usefulness of the system,
5. Discontinuing social promotions,
6. Using parental satisfaction surveys as a component of assessing school performance, and
7. How to implement public awareness programs about the benefits of UPASS.

The task force is to complete its work and report its findings by November 30, 2000.

An Evaluation of House Bill 144

An examination of the contents of the proposed House Bill 144 and the time frame specified in the bill makes it clear that the task force decided the goal of the 1999 legislation to have a statewide standards and accountability program functioning by July 1, 2000 was too ambitious. The proposed legislation deals only with the assessment system (UPASS) and performance reports for schools and school districts. As mentioned, many of the elements of the assessment system were put in place by House Bill 33 in 1999, however, the new elements shed some light on what the task force learned during their year-long studies. The new elements of the assessment system are:

1. The constructed response sections which are to be added to the CRTs,
2. The inclusion of science as a basic skills course,
3. Writing assessments for students in grades

6 and 9 (beginning 2001-02).

4. Use of student behavior indicators in assessing student performance .

As discussed, the addition of the writing assessments and the constructed response portions of the CRT to the statewide testing program will make it possible to obtain a more accurate estimate of student performance statewide. The down side of using multiple measures is that they add to the cost of the program. The constructed response portions of the CRTs and the writing assessment will be much more expensive to score than a multiple choice test. Such assessments require highly trained readers to evaluate each writing sample. At least two readers must read each sample to be sure it is judged fairly. These tests will provide very valuable information about student progress but they also require a substantial investment.

The addition of science to the list of basic skills courses which will be tested by the CRT is also meaningful. Utah has established a core curriculum which covers many areas including language arts, math, science, social studies, music, art. etc. Just which of these subjects are "basic" has been the subject of many debates over the years.

The task force heard comments that to become citizens of the modern world, it is important for students to have an understanding of the basic principles of science which play such an important part in everyday life. The NAEP and norm-referenced tests cover science, and the CRT test should also assess the state's science curriculum. In the future, the state may want to add other areas such as social science (including history, civics, and geography) to the list of basic skills courses.

The phase in period of the various elements of the assessment system over the next four years is important. Developing tests that meet the requirements of reliability and validity takes time.

The tests must be piloted before they are finalized. Commercial tests such as the norm-referenced test the state has been using (Stanford 9) have been carefully developed and piloted. Tests produced by the state such as the CRTs, the tenth grade basic skills test, etc., must meet the same standards. The four-year phase in of the UPASS system hopefully allows enough time for the important development work and piloting of each of the assessments to be accomplished before the tests are finalized. This work should not be rushed if the assessments are to be valid and reliable measures of student performance. Test results should not be used to evaluate school performance until the tests have been adequately piloted and tested.

Student behavior indicators have not been defined in the proposed legislation. These measures are intended to add some degree of student and parent accountability to the system. During the public hearings, the task force discussed using absenteeism and graduation rates as measurements of student behavior. An exact definition of what is to be included in this element of UPASS must be developed during the next year before it can be adequately evaluated.

Performance Reports

Currently, annual school district performance reports are required to include norm-referenced test scores and trends, ACT (college entrance exam) scores and trends, AP (Advanced Placement) data including number of tests taken and the percentage who passed, enrollment data, attendance data, expenditures, pupil-teacher ratios, education and experience information on staff, some student demographic information, etc. School districts must publish their report by January 15 of each year and may add any additional information beyond what is required that they wish.

Since each school district compiles their own data and prepares the report in whatever format

they choose, it is difficult to compare one district with another. Under the proposal in HB144, the state would assume the responsibility for collecting the data. Uniform collection instruments and systems would be put in place, the state would compile the data and send the report to the school district by October 1 for the previous school year. The district would be responsible for disseminating the information to each student household before November 30.

The information would be reported school by school and totaled by district and at the state level. This would make it easier for citizens to make school to school and district to district comparisons. Providing adequate and understandable information to all the parties involved in education in the state (students, parents, educators, business etc.) that accurately portrays what schools are accomplishing is, as discussed earlier, a key element of a workable accountability program.

While assessments and performance reports are important parts of accountability systems which promote accountability by reporting outcomes, there is much work still to be done before Utah will have a working accountability program. Two of the most important elements of accountability programs the task force will be working on during the next year are establishing state standards and preparing interventions for schools who do not meet those standards.

Utah has a core curriculum and a means of testing that curriculum (the CRT). With the other assessments of the UPASS system, Utah will have a good idea of what students know and can do. The question remaining is, at what level should Utah students perform in order for schools to be considered successful?

Some states have set high standards and expect all students to meet those standards right away. Other states begin with lower standards giving schools a chance to adjust to new programs and ideas and then raise those standards, requiring

improved performance over time. Still, other states take improvement from year to year into consideration in addition to raw test scores. The task force (after receiving input from educators, parents, the business community, and the public) will suggest what standard of performance constitutes a “successful” school during the upcoming year.

The task force will also have the difficult task of recommending an effective and affordable system of rewards and interventions. Some of the possible elements of such a system have been mentioned. However, Utah’s system, to be successful, should also include a “last resort” intervention.

Schools who fail to meet the standard even after a fair intervention period should face stronger sanctions. In many of the states currently using accountability programs, the last resort is to reconstitute a school. This means bringing in a new administration and teaching staff and starting over. There is currently no proof that reconstitution is effective in raising student achievement.

Over the next few years, studies should be able to determine if student achievement does improve in schools where reconstitution has been used. Clearly, interventions should be designed so that such drastic measures would seldom be used. However, schools should not be allowed to fail year after year and, on occasion, drastic measures may be necessary to prevent that from happening.

Adequate Funding

No accountability program will be successful unless it is adequately funded. The anticipated annual costs associated with proposed House Bill 144 are shown in **Table 1** (page 172). Note that some of the expenses related to the assessment program pertain to tests already part of Utah law as contained in House Bill 33 passed in 1999. Other costs are associated with the new assessments and new data collection required to

produce the state’s school by school performance reports. In FY1998, \$1.677 billion dollars of state money was spent on public education in Utah. This does not include local government expenditures which were about \$700 million.

The ongoing annual costs of the UPASS testing program will be almost \$7.5 million in state funds and an additional \$1.4 million in local school district funds. The required state monies represent a 0.5 percent increase over current expenditures and local funding would have to increase by 0.2 percent. In Utah, where such a large percentage of the state’s population is made up of school age children, adequately funding education is a continual problem. Should the legislature consider funding this expensive UPASS system?

Proponents of accountability programs would argue that when funds for education are tight, it is especially important that those funds are producing the desired outcomes. Opponents of the new assessment system argue that the norm-referenced tests used in Utah for years, which are relatively inexpensive, already provide enough information about student outcomes to make comparisons. This is true to some degree. Norm-referenced tests do provide information about the general knowledge of Utah students and how that compares to the knowledge of students in the national norm group. They facilitate some comparison of Utah students to other students. However, what they cannot do, is provide information about how much of Utah’s core curriculum students have mastered and where they still need work.

Criterion referenced tests fill this important need. So why not just use CRTs? CRTs do not provide information comparing Utah’s core curriculum to the curriculums of other states. The best information is obtained by using both the CRTs and norm-referenced tests as outlined in House Bill 33 and UPASS. The valid reasons for adding additional measures of student performance (writing assessments, tenth grade

basic skills etc.) to the testing program were discussed earlier. Norm-referenced or CRT tests alone provide such a limited view of student performance, an accountability program based on that limited view would lack some credibility.

The same is true of the costs associated with the data collection program. Without uniform standards for collecting, compiling, and publishing data no valid comparisons can be made. When school performance cannot be compared, it is difficult to justify any system of rewards, interventions, or sanctions.

If the legislature does decide to fund an accountability program, the key to the program's success will be adequate funding for the intervention programs which will bring low performing schools up to the state standard. The elements and costs of this intervention program have not yet been identified but will likely be significant. An accountability program will only improve student performance if schools have the resources (text books, supplies, adequate facilities, teacher training, time for tutoring, incentives for improved parental involvement, etc.) that will allow every student to meet the standard. The legislature will be accountable for finding the long-term funding necessary to make the program a success.

Conclusion

Utah Foundation believes that a well-designed accountability program could be instrumental in ensuring that students leave Utah's public education system with the basic skills necessary to participate successfully in society. Such a system must include all of the five key elements: content and achievement standards, testing, reporting, a system of rewards and interventions and adequate funding.

It is important to note that accountability programs are quite new. In some cases these programs seem to be successful in improving student performance. These reported

improvements are currently being tested by independent researchers for their validity. While the logic behind accountability programs is impressive, there is no "proof" yet that they are effective in improving student performance in the long term.

When the task force was established, it was charged with studying and implementing "proven" programs. This is not currently possible. Utah will have to work to establish a system based on the best information available to date, carefully monitor the results over time, and be willing to make adjustments, if needed, to meet the desired objectives.

One final note, House Bill 144 mentions that the task force will, in the next year, study how to implement public awareness programs about the benefits of UPASS. This is an important part of developing an accountability system that must not be overlooked.

In other states, some resistance to testing, rewards, and interventions has come from the public who felt they were left out of the decision-making process and did not like the elements of their state's program. Law suits over testing and performance standards have been filed.

If an accountability program is to be successful in Utah, all participants in the program and the public at large must understand all the elements of the program (their goals and objectives) and be confident that the costs (both fiscal and educational) associated with the program are acceptable. With broad support from educators, parents, students, and the public at large, implementing an accountability system in Utah could have a very positive impact on achieving the goal of helping Utah students leave schools with the basic knowledge and skills that will enable them to succeed.

Table 1

**Annual Costs Associated with Proposed
House Bill 144 for Assessing, Evaluating
and Reporting Student Performance**

| | Annual Cost |
|--|--------------------|
| Norm Referenced Tests | \$135,000 |
| CRT With Constructed Response | 5,210,700 |
| CRT Science Assessment | 78,000 |
| 10th Grade Basic Skills Test | 17,500 |
| 6th & 9th Grade Writing Assessment | 830,000 |
| Staff Development & Public Information | 1,000,000 |
| Other Support Costs | 256,200 |
| Subtotal - State Costs | \$7,527,400 |
| District Costs | 1,426,600 |
| Total | \$8,954,000 |

Source: Utah State Office of Education