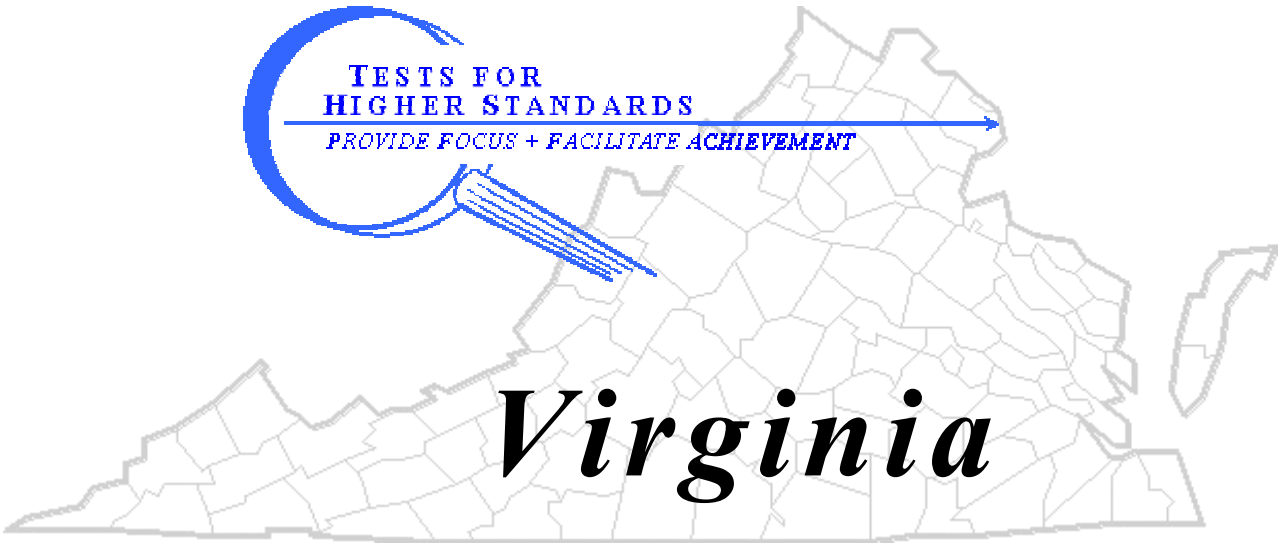


# User Guide

## Mathematics, Science, English, and History and Social Science

# TABLE OF CONTENTS

# OVERVIEW

In 1997, the Virginia Board of Education directed that all public school students be tested on the Virginia Standards of Learning (SOL). ***Tests for Higher Standards*** (*TfHS*) has been developing SOL diagnostic tests since that time. The grades and courses tested by the state have changed over the years. We changed our tests in response. ***TfHS*** has assessment materials from K to grade 11. We have ***Item Banks*** that cover all the core subjects, grades, and courses. We also have survey tests that cover a year's worth of SOL material — ***Grade Level Tests*** (*GL*s). In every grade where there is an SOL test, we have a ***Simulation Test*** (*SIM*) that mirrors it. The *SIM*s conform to the form and substance of the SOL tests as closely as possible.
We change with the times. Several years ago, we changed our fonts to *Tahoma* and the ***SIM*** tests from two columns to one, as Virginia made changes. When revised SOL standards are introduced, ***TfHS*** revises our assessment materials to conform to the changes. **Recently we added TEI items to our *Item Banks***.

The initial test results from the first 1998 testing — in every school division — indicated the need for materials designed to help students learn the content included in Virginia's Standards of Learning. Only two percent of Virginia's schools received passing scores. The Standards were written to provide a challenge for Virginia's students and SOL tests proved them so. Moreover, Virginia students had to learn to respond to the kinds of multiple-choice questions found on the SOL tests. In the early years, the ***Tests for Higher Standards*** products provided opportunities for both classroom instruction and assessment. Times have changed. The students are now achieving at much higher levels. Still, SOL results tell us that many students are *still* at risk. The ***TfHS*** materials help identify which ones who are and where each student's strengths and weaknesses lie. Then, assessments can detect whether remedial actions have been successful.

The ***Tests for Higher Standards*** products were modeled on the successful *Literacy Passport Test* materials created by Dr. Stuart Flanagan, Professor Emeritus of the College of William and Mary. These materials were used by **eight of the ten top scoring school divisions in the state**. In one particular case, King and Queen County students achieved the highest scores in the state during two of the last three Literacy Passport Test years, in spite of the fact that 85% of the students receive free or reduced-price lunches. ***Tests for Higher Standards*** products, like the *Literacy Passport Test* materials that preceded them, enable teachers to focus instruction on specific standards. Our success continues: at a recent SOL Expo sponsored by the Virginia Association of School Superintendents, 12 of 16 presenting schools used ***TfHS***. See our Success Stories at: http://www.tfhs.net/success_va.pdf.

Again, there are three different products for each content area: the ***Grade Level Tests***, the ***Simulation Tests***, and the ***Item Banks***. These three products provide an excellent set of flexible assessment tools for an ongoing diagnostic/remediation approach to instruction. The NCLB agenda requires such an approach. Please request information about our *Diagnosis-Remediation Approach to Instruction*.

The No Child Left Behind Act requires, as a minimum, statewide annual testing in reading and mathematics for grades 3-8 and testing science in one grade in elementary, in middle, and in high school. From this assessment, it is determined if **Adequate Yearly Progress** (AYP) is being made. ***Tests for Higher Standards*** can be an integral part of meeting certain requirements of the Act and ultimately assisting you in meeting your AYP goals. Instructionally, diagnosis and remediation will be the cornerstone of meeting the challenge. ***Tests for Higher Standards*** materials are ideally suited for this. We have a report that is available to our clients that demonstrates the research base of ***TfHS***. This information is intended to satisfy the NCLB research requirement.

Users of the *Grade Level Tests* can point to specific standards in need of remediation. *Item Banks* can specify which standards have been mastered ongoing, including in after-school or summer school programs. Additionally, this item bank can easily be a means to evaluate your programs themselves by **developing interim benchmark/nine-weeks tests, along with snapshot tests**. ***TfHS*** would be delighted to assist in developing a diagnostic/remediation approach through our materials. You can determine the progress of each student, class, school, and then the district. Our reporting forms allow for a diagnostic approach, critical in meeting the demands of **AYP.**

These products are offered in four instructional areas — Mathematics, Science, English, and History and Social Science — for students in grades Kindergarten through eleven (end-of-course). The products are Standards-based (formerly called criterion-referenced). They are aligned specifically and uniquely to the Virginia SOLs.

Most importantly, the ***TfHS*** products provide a way for students and teachers to focus on specific SOL content at each grade level. The material included on the tests is challenging, often thought provoking, and always useful to students who must prepare to do well on the SOL tests. On balance, our tests are slightly more difficult than the State SOL tests.

# Test Publishers - Tests for Higher Standards

Dr. S. Stuart Flanagan, Professor Emeritus
College of William and Mary
Box 276
North, VA 23138
804 / 725-7997
804 / 725-5541 fax
stuflanagan@aol.com

Dr. David E. W. Mott
P.O. Box 7417
Richmond, VA 23230
804 / 282-3111
804 / 282-4126 fax
866 / 724-9722 toll free
dem@rosworks.com

Visit *Tests for Higher Standards* online at http://www.tfhs.net.

# About the Publishers

**S. Stuart Flanagan, Ed. D. –** is a mathematics educator with substantial experience in research and student testing. Education: Washington and Lee University – BS, University of Virginia – M.Ed., and Ed.D. He served as chair of the math department at St. Christopher's School (grades K-12) in Richmond, Virginia. Thereafter, he was a professor at the College of William and Mary for 30 years teaching mathematics and mathematics education-related courses including research, testing, and curriculum development. He is now Professor Emeritus.

During his tenure, he developed test items for several projects at the local and state level, directed several NSF grants, consulted with school divisions statewide, authored numerous articles, and was a senior consultant to CBS Publications for the mathematics series Mathematics Unlimited. Additionally, he was a grader for the ETS Advanced Placement Program. During the late '80s and early '90s, he was known throughout Virginia for his highly successful Literacy Passport Test materials, used extensively throughout the Commonwealth of Virginia. These materials, by providing data for each child on every standard, enabled a number of districts to dramatically increase their student-passing rate. The conceptual model for those materials is the basis for the Tests for Higher Standards.

Special honors include being the first president of both the Virginia Council of Teachers of Mathematics and the Peninsula Council of Teachers of Mathematics. He also served as president of the Greater Richmond Council of Teachers of Mathematics and was presented the William C. Lowry Outstanding Mathematics Teacher award. Additionally, he was an E. I. DuPont Fellow at the University of Virginia and was awarded a Shell Merit Scholarship and NSF grant for study and independent research at the University of Virginia. He now works with Dr. Mott in developing high stakes testing and instructional materials.

**David E. W. Mott, Ph. D. –** by training and work is a psychologist and psychometrician. Education: Vanderbilt University – BA, American University – MA, Virginia Commonwealth University – Ph.D. He worked at the Virginia Department of Education for nearly twenty years within the student testing division. While there, he was Supervisor of Test Development and Virginia State Assessment Program (VSAP) Administrator, among other duties and positions. His activities covered a full range of testing activities: test construction, validation, equating, and utilization. He was involved in test adoption procedures and worked with curriculum development. He was involved with state-wide tests such as: the Basic Learning Skills Tests, the Graduation Competency Tests, the Standards of Learning – Teacher Resource Materials (1st and 2nd editions),

the Literacy Passport Tests, the National Assessment of Educational Progress (NAEP), and the Scholastic Aptitude Test (SAT), as well as with VSAP. He was the testing unit's liaison with the computer support group. He is past-president of the Virginia Research Educational Association, and a member of the American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education. He also belongs to Phi Delta Kappa and to ASCD.

He now works with Dr. Flanagan in developing testing and instructional materials. He also has designed, produces, and offers the Reports Online System (ROS) through ROSworks, LLC — an online scoring, testing and reporting system. ROS is an easy-to-use, intuitive, non-resource-intensive, test- scoring system that features powerful, immediate, and detailed online reports: http://rosworks.com.

## Publisher's Acknowledgments

We are extremely grateful for the continued work and support of so many. We are constantly asking teachers, supervisors, directors of instruction, schools, and, at times, entire school districts to provide input and assistance in conceptualizing and developing our assessments. As we are constantly updating and revising to meet changing standards, to introduce fresh content, and to improve the assessment wherever possible, it is essential that we have this ongoing interactive approach with input from all educational directions. We also benefit from unsolicited input and commentary. These approaches have been essential in providing the consistent, quality assessments that we that our clients deserve. Without their varied input, we would not be so successful. Hence, this acknowledgement is most genuine.

We especially appreciate the advice and counsel offered by Dr. Ron Gise of the College of William and Mary who reviewed and edited early copies of the science materials, and likewise to Dr. Richard Weber, Social Studies Supervisor for Newport News Public Schools.

Last, but not least, we are most grateful for all the encouragement, support, and statements of appreciation from so many professionals in the field. We thank you all.

*S. Stuart Flanagan, David E. W. Mott July, 2017*

## Test Authors

| | |
|---|---|
| **Mathematics** | Dr. S. Stuart Flanagan<br>assisted by Amy Lamb, Philip and Susan Sanderson, Ramalingeshwar Mada, Beth Teigen, Rebecca Bienvenue, DeAnna Moreau, Michael Traylor, and Norma England |
| **Science** | David Mott, Andrea Walters, Sharon Spacek<br>assisted by many, many educators who have contributed items and reviewed materials over the years |
| **English** | M. Scott Reynolds, Linda Newman, David Mott<br>and reviewers, too numerous to mention |

**History and Social Science** Dr. S. Stuart Flanagan
assisted by: Dr. Richard Weber, Anne Kennedy, Amy Kovac, Karen McPherson, Linda Olson, Greg Tyson, and Renita Williams (1999-2000).
Contributors: Mayrene Hutchinson, Trish Woller, Allyn Anderson, Gary Carnot, Brian Dunnell, Anissa Frame, Anna Graham, Trey Loker, Tami Ogden, Jane Parrish, Debi Reeves, Amy Smith, Richard Zimermann, Teresa Stallard, Elizabeth Morgan, Jeannine Dearmon, Norma England, Darrell White, Jessica Mitchell, Sherry Elliott, Larry Sabourin, Debbie Manuel, Melinda Sprinkle, Mary Vipperman, Amy Lamb, Tina Pitts, Robyn Lacasse, Debra Noll, Jason Miller

# I. Grade Level Tests — Test Rationale, Description, and Development

## Meeting the SOL Challenge

To meet the continuing SOL challenge, Virginia's schools must use instructional resources in a focused way. Teachers must teach and students must master the content and processes included in the Virginia Standards — not just in some grades — but at all grade levels in each subject. Only an ongoing, systematic approach will ensure that students do well on the state-mandated tests. Along with this challenge, the NCLB requires even greater demands.

In addition to their important role in helping teachers with SOL *content* that challenges their students, the ***TfHS Grade Level Tests*** can also familiarize and empower students to handle the *context* of SOL testing. Though no practice tests can simulate exactly the forms and features of a high-security test, the ***TfHS Grade Level Tests*** provide significant practice in the types of questions that students face on the SOL tests. Creating good practice items is both difficult and time consuming, and few teachers have had training in writing quality test items. Skill in answering multiple-choice questions is the "coin of the realm" for our students; it makes sense to use these available resources.

## Test Coverage

The ***Tests for Higher Standards Grade Level Tests*** in the areas of Mathematics, Science, English (Reading and Writing), and History and Social Science for grades Kindergarten through eleven have been developed to help teachers focus instruction on the content and processes of the Virginia SOLs. Because ***TfHS Grade Level Tests*** are written to closely match the SOLs, they help teachers plan, teach, and assess, providing information about student and program strengths and weaknesses.

## Test Description

***TfHS Grade Level Tests*** are standards-based, primarily four-alternative, multiple-choice tests. Each test contains from forty to eighty-five items — three to six items per Standard — and can be administered in one or two class periods.

Each test item (question) has been designed to measure one or more aspects of *a single Standard.* The resulting ***TfHS Grade Level Tests*** are content-valid, in that test items are clearly referenced to individual Standards and measure them as directly as possible.

Not all facets of all Standards can be measured by any test of reasonable length. When the SOLs call for student performance not measurable in the multiple-choice format, reasonable compromises have

been made. The ***TfHS Grade Level Tests*** measure, directly or indirectly, all SOL in the four core areas, except those excluded from the state-mandated tests.

The tests were developed by experts in each subject area with the help of consultants from Virginia and the nation. Originally, the SOLs and local division curriculum guides provided content objectives and vocabulary. We now focus on the Skills and Essential Knowledge section of the State's Curriculum Framework for content objectives.

Finally, the tests are designed to be pedagogically sound within the broad scope of each of the disciplines. They have been written by educators for educators — *and for students*.

### Testing the Tests

Because it is important that any test be both reliable and valid, teachers, administrators, and students have had numerous opportunities to review and field-test the passages and test items in the ***TfHS Grade Level Tests***.

***TfHS*** products have been student-tested. In addition, teachers have carefully critiqued and edited the tests, evaluating for:

- the content match with Virginia Standards of Learning;
- readability for a particular grade level;
- the correct difficulty of content;
- an appropriate set of formats; and
- cultural sensitivity.

Following these evaluations, authors have revised extensively, adjusting items and answers to address problems noted by teachers across Virginia. In most cases, teachers have again evaluated the tests.

Students have also guided the revisions of the ***TfHS Grade Level Tests***. At the beginning of 1998, some school divisions elected to use the tests in draft form. Schools in those divisions sent feedback about specific test items, and all appropriate changes were made to the tests and the supporting materials.

It is safe to say that the ***TfHS Grade Level Tests*** have undergone the closest scrutiny. Some tests are now in use by more than half of the local school divisions in Virginia. Alternate forms of the tests are also in use in other states whose standards resemble Virginia's. Without doubt, the tests are useful to students. And the tests provide teachers a way to focus instruction so that all Virginia's students master the content and processes included within the Virginia SOLs.

## USING THE TESTS

The ***TfHS Grade Level Tests*** provide a snapshot of student mastery of Virginia SOLs at any time. They are useful in a number of ways. In addition to focusing instruction on the Standards, these tests are designed to allow schools to gauge overall performance on Virginia Standards so that instructional resources can be allocated. An ongoing use of the ***Grade Level Tests*** is diagnosing grade level work. For example, to see the deficiencies for Grade 7 Math, give the Grade 6 test at the start of the year; for Algebra 1, give Grade 8 at the start of the year, etc. Also, the ***Grade Level Tests*** can assist in proper placement for transfer students; extremely helpful if the child is from another state or their record transfer is slow. These tests have more items per SOL than the ***SIM*** tests. The items are different.

## Pre-Post Testing

The ***TfHS Grade Level Tests*** can be used before and after instruction to determine the learning progress. We recommend testing students over the course of a school year, but some circumstances may dictate a shorter pre-post test period. In addition, we suggest these guidelines:

- Students who take a ***TfHS Grade Level Test*** as a pretest at the beginning of a school year should take the test designed for that grade level. If diagnosis of initial weaknesses is the focus, the previous grade level tests might be considered to assist with that.
- In some subjects, teachers could choose to use the ***TfHS Grade Level Test*** given near the end of the previous year. The post test, given near the end of that grade's course of instruction, would be the test for the student's current grade.
- If at least six months intervene between pre- and post testing, it is both practical, and may be preferable, to use the same form of the same ***Grade Level Test*** for both pre- and post-tests.
- The ***TfHS Grade Level Tests*** are necessarily challenging. Teachers who work with students whose achievement is consistently below grade level may find the test for that grade level too difficult. In such cases, instructional leaders may choose to administer tests from earlier levels.

## Special Testing K/1

It has been determined that there needs to be post testing for SOA purposes and *TfHS* has undertaken the task of doing just those assessment materials. We went forward in this development with mixed feelings. Nonetheless, we have tests that can be useful with students. It is quite necessary to read the test to the students and give directions for taking the test. Additionally, this might mean illustrating and clarifying certain aspects of the process with students. It is imperative that the students understand the test format and procedures before we expect them to answer correctly. By giving ongoing instructions to the students, you will pace the students through the test. Depending on the reading ability of the class, these same procedures may apply to second graders. See more details below.

## End-of-Year Testing

The ***TfHS Grade Level Tests*** may also be used as summative, end-of-year tests, without any pretesting. In this case, student progress is measured against the previous year's performance.

The ***Grade Level Test*** may certainly be used *as a portion* of a local division- or teacher-written final examination. To do this successfully:

- the ***Grade Level Tests*** should be supplemented with locally relevant assessment, and
- final grades and promotional decisions should be based on a combination of teacher judgment, cumulative evaluation of student work, as well as other criteria, rather than on the results of any single test.

## Mid-Year Testing

Another way to use the ***Grade Level Tests*** is to administer them during the year, after the students have had substantial instruction on the course material. The results would allow teachers to plan effective remediation on Standards already covered and craft initial instruction on Standards not yet covered. Results from a January or February testing can be useful as a predictor of student success on the spring state-mandated tests. (***TfHS*** has also developed custom semester tests for school divisions based solely on the SOLs covered in each semester.)

## End-of-Unit Testing

Some teachers may choose to use portions of individual ***Grade Level Tests*** as end-of-unit tests. For example, a mathematics teacher might choose to administer only the section of the ***Grade Level Test*** that assessed knowledge of number theory. Or a language arts teacher, following a unit of instruction on persuasive writing, could ask students to answer only the questions about persuasive writing. ***Grade Level Test*** items used in this way would need to be supplemented with additional items or other assessments because it is difficult to establish test reliability if only three or four test items are used.

## Six- and Nine-Weeks / (Interim) Benchmark Testing

One approach to testing that we have seen increasing interest in recently is the use of six- or nine-weeks tests. Teachers have selected items from ***TfHS*** tests indexed to the standards taught in a marking period and used the resulting tests as an ongoing evaluation of student progress throughout the year. This provides a clearer picture of developing achievement, and pinpoints areas of weakness to be remedied as the school year progresses. See **Section III.** ***Item Banks*** for more details**. For your convenience and by popular request, items from the *Grade Level* and *Simulation* tests are now included within the *Item Banks*.**

## Practice Testing and Instruction

Any use of the ***TfHS Grade Level Tests*** will do at least two things:

- provide students with practice for the state-mandated SOL tests, and
- give teachers the chance to see SOL content *embodied in actual test items*.

Some tests, such as the writing and reading ***Grade Level Tests***, will be particularly useful as instructional texts. The writing tests contain writing in passages and answer choices that can become useful subject matter for class discussion. Similarly, the reading tests contain literature, both old and new, that may supplement other class reading.

For many reasons, teachers themselves will want to take the tests prior to administering them to students. Teachers who are familiar with the ***Grade Level Test*** items will be better able to assess:

- personal knowledge of content;
- instructional methodology; and
- student readiness for testing.

## Grade K-1 Grade Level Test Directions and Considerations

Due to the limited reading ability of students in Grades K and 1, we have a special set of instructions for evaluating at this level.

The directions and considerations of Grade K-1 tests are as follows:

1. This test should be given in a group setting. The teacher should determine the size of the group-ranging from individual to large groups including the total class in select situations.
2. The teacher should be in the students' view at all times.
3. The teacher needs to make sure each student is on the proper question.
4. The teacher will read each direction, question, and answer choice.
5. The teacher may use the pictures in the left margin to direct students to the proper question; for example, "Place your finger on the bananas."
6. The teacher may name the picture in the question if they feel the student will not know what the picture is.
7. If a student needs a question repeated, the teacher may do so as many times as needed.

8. The teacher should NOT define or explain concepts being tested; for example, "What is a rhyming word?" or "What does reptile mean?"
9. The student should select the best answer choice for each question and then circle the letter in front of the answer choice.
10. The student will write in the test pages as directed and may otherwise mark on the test.
11. The test may be given in more than one testing period.

### Grade 2 Grade Level Test Directions and Considerations

As we attempt to give suggestions for administering the 2nd grade test, we want you to know there is great variability within and between 2nd grade classes. Hence, you need to use your best judgment in deciding on what and how much to either read the test or pace the class. Here is what has worked:

**Science and History/Social Science tests**: Read the whole test, both pre and post, as reading ability is not being tested here and many of the words may be difficult to read at this age.

**Reading and Writing tests:** Read everything except the stories for the pretest; read the directions to each section and allow the children to read the rest of the posttest themselves. However, make sure they know how to follow the directions given.

**Math test:** Read everything on the pretest; read only the directions on the posttest and have the students work on their own. You may want to do each group of questions together. For example, you might say, "Now do the next three problems on your own. When you are done, stop and wait."

### Math Facts Tests: Grades 2 and 3

We believe that basic mathematical facts are essential for a student to know. This skill requires the ability to recall those facts in some reasonable time frame; hence we provide a systematic assessment of this knowledge through timed ***Facts Tests*** (addition and subtraction in grade 2 and multiplication and division in grade 3) in the *Item Bank*. This skill is basic to doing either mathematical computations and/or estimation. You will be the ultimate judge of the time limits involved, but we will offer a suggestion: allow 3-4 minutes each for section in addition and subtraction. You might consider breaking the tests into four sections and adjust the time for the smaller sections, as you want to see how well the student can either add and/or subtract in both vertical and horizontal formats. The same approach to multiplication and division is suggested: 4-5 minutes each for section in multiplication and division. For young students taking a timed test, we suggest that when time is called, ask the students to mark their last completed item at that time. Then they may be encouraged to finish the test. Having to completely stop the test before the end may be too frustrating for them. For more specific suggestions/strategies, visit the website:http://www2.ups.edu/faculty/woodward/facts%20overview.pdf

## — NOTES FOR TEACHERS —

To teach is to be responsible for student success. This has always been so, but the Virginia Standards of Learning and the SOL state-mandated tests have underscored that responsibility. More than ever, you need the support of good instructional materials, and we feel confident that the ***TfHS Grade Level Tests*** for your subject and your grade level will help *you* help your students succeed. Contact us if we can help you as you use ***Grade Level Tests***, and please know that we welcome your suggestions for new test items, changes in old items, or other materials you would like us to develop, especially semester exams.

## Using the Pre-Post Test Model

In the following discussion, we will assume that you and your colleagues have decided to use the tests as pretests in the fall and post tests in spring.

**Before You Test:** Before the day of pretesting, you'll need to review the tests themselves, as well as all accompanying materials. Make sure that you have a test copy and a scoring sheet for each of your students. A complete package of test materials includes:

- ample copies of the test for your subject and grade level;
- a Classroom Matrix (for viewing class test results);
- a Time/Sequence Planning Chart (for outlining the year's instruction);
- scoring sheets; and
- an answer key.

You'll want to give the test during the first week or two of the school year, so that you can complete the Class Matrix by the second week of school.

**The Classroom Matrix:** The Classroom Matrix will help you focus instruction, guiding the way you spend instructional time, and helping you make decisions about the sequence in which you teach the content and processes for the Standards of Learning for your subject and grade level. The Matrix gives you a picture of your students' strengths and weaknesses, as individuals and as a class. The Matrix also provides data about each SOL.

**Time/Sequence Planning Chart:** The Time/Sequence Planning Chart you have is your trip "routing plan." Use it with your knowledge of the students and a reflection on your own experience, and make a tentative decision on how much time you will devote to each SOL and what sequence you will use. To illustrate briefly, you might combine instruction on fractions with teaching customary measurement and combine decimals with metric measurement. Allow the number of days you feel is necessary. Yes, there are not enough days in the school year for what you need to teach, but you do need a plan to allow students an opportunity to learn the SOLs for their grade level. This can be a real conflict for those who are used to teaching learning for mastery. There is simply not enough time. As you can see, the idea here is for you to follow your travel plan but don't spend too much time in one place or you will not complete the trip. Yet, at the same time, you need to visit all the sites you had in mind. You will make the key decisions that are most important.

**Skill Maintenance:** This is a vital part of any instructional program. Skills, once learned, need to be reinforced from time to time. **There needs to be an especially strong emphasis on maintenance of skills at third-, fifth-, and eighth-grade levels to coincide with the state tests. This is important for other grades as well.** We urge you to develop a plan for doing this systematically, as maintenance is a vital part of instructional planning. Dr. Flanagan found with his *Literacy Passport Test* work that continuing skill maintenance was *extremely important* in having students do well on the Literacy Tests. The coverage of the state SOL tests is generally cumulative! **You have an ideal solution: use our *Item Banks* and the related software to produce snapshot tests as described in the MAKING CUSTOM TEST folder on your CD.**

**And If the Scores are Low:** Don't be surprised if your students' scores on these tests are low. The ***TfHS Grade Level Tests*** are difficult because they are written to a challenging set of Standards. If you teach Mathematics or English Language Arts, and your students answer less than 20-25% of the items correctly, you may want to give the test for the grade level just completed. In these subjects, students

use many of the same processes from year to year; so testing at a lower level will give you good information about what your students can do.

On the other hand, the History and Social Science standards — and to a lesser extent, the Science standards — are based for the most part, on content taught at specific grade levels; thus, it is less practical to administer tests from previous grade levels.

**Planning for Testing:** In most cases, students will complete the ***TfHS Grade Level Tests*** in one or two class periods. The ***TfHS Grade Level Tests*** are — with the exception of the ***Facts Tests*** in Grades 2 and 3 — power tests, not speed tests. Some students are likely to complete tests earlier than others, so you will want to plan for this. Perhaps you can arrange for students who are still testing to use another quiet space; or you might plan an activity away from the testing area for students who complete the test quickly.

Because these tests are relatively difficult, you'll need to watch students carefully. While it's true that the tests are not timed, it is a mistake for students to become frustrated. If a test is *much* too hard for a student, discontinue it and provide him or her with another activity. If the purpose of your testing is diagnosis, you may want to give the student the test from an earlier grade level.

**On the Day You Give the Test:** At the time of the testing, you will issue these materials to your students:

- a test booklet;
- an appropriate amount of scratch paper (for example, two or three sheets for mathematics tests; more for writing tests);
- two or three pencils;
- other items, as needed, such as graph paper for grade eight and Algebra 1 students or calculators for certain ***TfHS Grade Level Tests*** in mathematics;
- an individual response sheet (or other machine-scored answer sheet).

## Scoring and Reporting

### DIRECTIONS FOR HAND-SCORING THE GRADE LEVEL TESTS

You can score the Individual Response Sheet or note the scores on the Classroom Matrix. Using the Answer Key supplied with the test for your subject area and grade level, mark each **incorrect** answer with a check.

#### Individual Response Sheet

If you use the Individual Response Sheet:

- Count the number correct on each Standard and place the number on the sheet.
- Compute number correct and percent scores for the strands and for the test as a whole.
- After you score all individual response sheets, transfer the strand scores to the Classroom Matrix to show the results for your class.

Note: This scoring method gives more individual diagnostic information, but slightly less information about the class as a whole.

#### Classroom Matrix

If you use the Classroom Matrix:

- Place a colored mark or dot in the correct item location for each incorrect response.

- Total the number of incorrect items and write the corresponding number-correct score in the space provided.

This scoring method gives less individual diagnostic information but slightly more for the class as a whole. You may choose to record the information on both forms.

### Time/Sequence Planning Chart

Use the Time/Sequence Planning Chart for lesson planning. In the elementary grades, you may want to combine instruction in one or more subjects, particularly when knowledge in one area — reading or mathematics, for example — enhances learning in another area, such as science or social studies. After you complete the Classroom Matrix and Time/Sequence Planning Chart, you may want to share test results and plans for instruction with other teachers at your grade level, or of your subject area, others, and with your principal or instructional leader.

### Post Testing

Give the test again during the last two weeks of school. Again, record the results on the Classroom Matrix (use a new one, if you used a different level of the test as a pretest). This time, note errors with a red "X". Ideally, this record of pre- and post test results would go with the student to the next grade or to summer school and provide a focus for remediation and instruction.

## Using Scanners and Online Testing, Scoring, and Reporting

In the past, most ***TfHS*** customers have scored our tests and completed reporting documents by hand. We realize that this places a large burden on teacher time, a resource always in short supply. For this reason, we will support systems to relieve this burden by automating much of this work, thereby reducing the demand on time and resources. ***Tests for Higher Standards*** supports a product developed by ***ROSworks, LLC*** in close cooperation with ***TfHS*** called the ***Reports Online System (ROS)***. A brief description of ***ROS*** follows. More information is available in the *CD Info* folder of the ***TfHS*** disk.

The ***Reports Online System (ROS)*** is a powerful, flexible, test scoring and reporting system designed to be used with printed tests and online computer-administered assessments. It provides for plain paper, bubble-sheet scanning and scoring, as well as online test administration and scoring, vertical data aggregation, NCLB disaggregation, and reporting. All operations except local scanning are accessed through the web. The system has been in use and in continued development and refinement for six years. The online testing function is a now an option. An online test construction system is a part of the system. You may contact ***ROSworks*** at 866-724-9722, (804-282-3111), or send an e-mail to salesforce@rosworks.com for more information. The web site is: http://rosworks.com/.

Other automated systems are available that provide a wide variety of features, including online testing, to assist teachers and administrators in the assessment process. Our assessment materials have been delivered by other vendors, including *Linkit*, *Interactive Achievement*, and *Teacher Resources*.

# ADMINISTRATIVE SUPPORT AND INVOLVEMENT

As an instructional leader, you will want to help plan for the use of the ***TfHS Grade Level Tests***. Teachers need your support. Students benefit when administrators and teachers make joint instructional decisions. In most cases, the principal in charge of instruction will do the following:

1. Supply ample copies of the tests to teachers, along with Classroom Matrices and Time/Sequence Planning Charts;
2. Determine a beginning-of-year schedule for test administration and completion of the Matrix and Time/Sequence Planning Charts;
3. Determine suitable times to interact with teachers about initial test results and plans, and about progress throughout the school year;
4. Assist teachers as they develop appropriate class assessment for evaluating student achievement throughout the school year;
5. Develop an end-of-year assessment schedule.
6. Arrange for student pre- and post test results to be forwarded to receiving teachers and/or appropriate summer-school teachers;
7. Assist teachers in getting the support of aides, parents, and others who can tutor students needing remediation.

# GENERAL TESTING CONSIDERATIONS

## Strengths

The ***TfHS Grade Level Tests*** are designed to:

- simulate the experience of taking the state-mandated SOL test;
- provide pre- and post-data about individual students and whole classes;
- provide diagnostic information about individual students on individual standards;
- provide a focus for instruction.

Results indicate that the ***Grade Level Tests*** do all of these things, and in some cases, more. Many students feel more confident about taking the state-mandated tests because they have had practice with the ***TfHS Tests***. And many students have begun to take the SOLs themselves more seriously because they are able to see their own strengths and weaknesses, as measured by the ***Grade Level Tests***.

Perhaps the greatest strength of the ***TfHS Grade Level Tests*** is that they place necessary emphasis on the Virginia Standards of Learning at every grade level, not just at grades three, five, and eight. One grade level at a time, all of Virginia's students can master the Standards of Learning.

## Limitations of These Tests

The ***TfHS Grade Level Test*** items were developed for only those Standards deemed testable in the multiple-choice format; thus, not every Standard is tested. The Standards excluded are the same as those excluded from the state-mandated tests.

## Limitations of All Multiple-Choice Tests

Certain SOLs require students to *create* problems, patterns, or constructions. These skills are important and need to be taught, but it is very difficult to test these skills directly on a multiple-choice test. This limitation also applies to state-mandated tests.

## Working with TfHS

We are aware of the limitations of grade level thinking in regard to individual child development, and we are acutely aware of the limitations of any test that uses only the multiple-choice format. Nonetheless, given the demands of the state Standards and the state-mandated testing program, we believe these tests are helpful to students and to their teachers and administrators.

The ***TfHS*** staff welcomes your suggestions. Please feel free to contact us about items you think we should revise and about materials we could produce to enhance educational processes.

## Grades K and 1 and Possibly 2

For grades K and 1, the teacher will want to give directions to the students on taking the test, by reading the items and pacing class progress. When testing Kindergarten and first grade students, you must read the test and therefore pace the students. To a degree, depending on the nature of the class, you may want to do this for second graders as well. Please see instructions on that test.

## Current Testing Material

All of the tests and other materials published by ***TfHS*** are available electronically in Microsoft Word® format. The documents should be readable in either Word for Windows ’97, 2000, 2003, or Word for Macintosh ’98, 2001. The ancillary materials (answer keys, student response sheets, class matrices, and time-sequence planning charts) are in either Word or in Microsoft Excel®. All of the tests are also available in camera-ready paper format (hard copy).

**Copies printed from the available electronic media should be checked for satisfactory appearance before mass duplication.** Various computer/printer combinations may cause the materials to output differently. Difference in line breaks, page breaks, and fonts, for example, can cause very substantial changes in appearance. Graphics can appear in the wrong position or on the wrong page. Also check the graphics, we have found that with some printer-computer combinations graphics may not print clearly.

## Customizing Tests for Your Division

If you desire, you can modify our tests or use the items to build your own tests. We suggest that you modify ***Grade Level Tests*** carefully to keep the balance ***TfHS*** has design into them. Please see **Making Benchmark and Snapshot Tests** (this document, p.18 and the information about custom tests on your ***Site Agreement*** - Appendix A.

# II. SIMULATION TESTS

## Description

The ***TfHS Simulation Tests*** provide the student with a means to experience a test that resembles the State tests for grades 3- 8, and high school end-of-course tests. The requirements spelled out in the State test's blueprints and the formats presented in the sample items are followed in detail. The State's reporting categories are used for scoring.

When you see the attached chart that comes with the test describing which test items are in which reporting category, you will note one difference from the State's blueprint. Here is the explanation. The State test may have 50 items, total. Of those, just 40 of the test items count towards the student's score and 10 items are included only for field-testing. In order to keep the ***TfHS Simulation Tests*** the same overall length as the State test, the extra 10 items are spread out among the reporting categories in the places where it was believed it would be most helpful instructionally. So, if the blueprints indicate 8 items are in a category, the simulation test might contain 9 or 10 items. However, the total number of items on a simulation test is the same as on the State test. ***Simulation Tests*** are available for all four content areas. ***Simulation Tests*** **are now available for each grade 3-8 in Mathematics and English, conforming to the new state blueprints.** Individual ***Simulation Tests*** are available for US History I, US History II, and Civics and Economics.

We now offer a second form of ***Simulation Tests*** in Mathematics, Science, and English/Language Arts. This form includes a sampling of items from our TEI item bank, to give students the flavor of the State tests. These TEI items match the standards assessed on the traditional all-multiple-choice form.

Our initial field-testing was done in the spring of 1999 with 20 school divisions with 6,500 students participating. The feedback was useful and provided suggestions for correction and modifications. We have found a high correlation with the State tests. See **Section V** for more information on validity and reliability. On your CD you have the cut scores that allow you to compare the score on our tests with the corresponding SOL score. We also found that our tests are generally more difficult than the actual SOL tests. Please request the ***SIMs Study*** for more details. **For your convenience and by popular demand, many items from the *Simulation Tests* have been included within the *Item Banks* for building benchmark and snapshot tests.**

Calculators are not permitted for some sections of the Grades 4-7 Mathematics SOL tests. The ***TfHS Simulation Tests*** **for grades 4-7** have sections at the beginning of the test marked "NO CALCULATORS" at the top of the page. Teachers may wish to give these sections separately from the remainder of the test.

## Some Recommended Uses

1. Give these tests under the same conditions the State requires:
   - *About 5 days prior to the State tests*. Work through the results with the students and have them psychologically ready for the real test. You will have limited time and information for carrying out a diagnostic remediation strategy, but you can make a start in that direction. Use the testing experience to ease test anxiety and teach productive test-taking strategies. The learning that results will probably not be long term, but the activity will almost certainly improve most students' test scores somewhat — which, after all, is the goal!
   - *Two to four weeks before the State tests*. Use the results to develop a diagnostic-remediation crash program.
   - *Toward the last of the first semester.* Use the period of the winter break to score the test and devise a combination crash program and an augmented, regular instructional process. Remediate the weak spots in already covered content and ensure the coverage of uninstructed areas.
2. The tests can be used in summer school to determine categories of weakness and learn what instructional focus is needed. Then the ***Grade Level Tests*** can be used to refine what specific instructional focus is needed.
3. *Caution:* We warn against using the ***Simulation Tests*** for detailed diagnosis, as they contain only a small sample of all the SOLs being tested.
4. *Note:* Teacher themselves should take this test well before giving it to students. This will make it clear what needs to be covered for the state SOL test.

## AN IMPLEMENTATION PLAN
### USED BY CHESTERFIELD COUNTY PUBLIC SCHOOLS

**Recommendation for Use of *TfHS Simulation Tests* in high school —**

*Goal of Administering the Simulation Tests:*

- The simulation tests are designed to:
  - simulate the experience of taking the State formatted test;
  - provide data about individual students and whole classes;
  - provide information about individual students on individual standards; and
  - provide a focus for review and remediation prior to taking the test.
- The testing scheme is based on the common sense model: first test to find out what students know and don't know (simulation test); second, re-teach what they don't know; third, re-assess for mastery.

*Administering the Test*

- Take the test prior to administration in order to allow teachers time to discover and evaluate:
  - personal knowledge of content;
  - instructional methodology; and
  - student readiness for testing.
- Maintain the daily schedule.
- Allocate no more than 90 minutes or two first period sessions.
- Allow students to write in their test booklet and record answers on the provided key if you can afford it.

*Timeline — Coordination within the School*

- Administer no sooner than Spring Break with a minimum of two weeks before they sit for the State test.
- Coordinate between subject areas so students aren't taking multiple simulation tests in one day.

*What to Do with the Test Results*

- Use as a diagnostic tool, NOT as a grade (extra credit points or participation points may be given), to develop a short-term remediation program.
- Work through the results and help prepare students psychologically for the test.
- Use the testing experience to ease test anxiety and teach productive test taking strategies.
- Divide the test questions for review by strand/category:
  - remediate the weak spots in already covered content and ensure the coverage in uninstructed areas and
  - refer to remediation notebooks for suggested strategies and activities.
- Analyze results using matrix to diagnose student weaknesses.

# III. ITEM BANKS

## Description and Uses

The major portion of this ***User's Guide*** is devoted to describing the various possible uses, rationale, and development of the ***TfHS Grade Level Tests***. With those tests in place, we felt a need to provide teachers and students with a means for ongoing feedback about student achievement through benchmark and snapshot tests. Hence, we produced ***Item Banks***. We believe these *Item Banks* fill a present assessment void in SOL-specific assessment content. Items were generated based on the specific Skills and Essential Knowledge as given in each subject's Curriculum Framework.

The *Item Banks* exist for all content areas, in Grades K-11. Generally, we have provided far more than the number of test items to yield a reliable means of determining either satisfactory or unsatisfactory student achievement on an SOL. In establishing the items to include, we used carefully considered teacher judgments of the number and variety of questions necessary. Our *Item Banks* generally contain some 350 to upwards of 500 items per subject per grade, usually some 25-150 plus items, per SOL. This number of items per SOL is higher in some cases, for example, in Chemistry and most of History and Social Science, because the numbers of standards in those areas is small but the coverage of each standard is quite broad. Due to the nature of these banks, some items may "give away" the answers to other items within the *Item Banks*. Please be aware of this as items are selected for tests and select the items you decide to use accordingly.

Many of the English *Item Banks* are in a format that we refer to as **"Extra Items Passages."** These passages come equipped with up to 20+ items per passage. The large number of items available per passage is a feature designed to make it easier to construct benchmark or snapshot tests. Using these passages and their items would allow a division to follow its pacing guide without need to write new items or to require too much reading per item. The desired passages can be selected; then the items covering unneeded standards can be deleted; and the benchmark is essentially done.

**In the *Item Banks*, the correct answer(s) are colored dark blue. This makes constructing the key for a test very simple — just "follow the blue." After the key is made, the text of the entire test is selected (control-A) and the font color is set to "black." Then the test is ready to go!**

If the SOLs are taught in some combination, several assessments from the booklet can easily be combined into one test. You might shorten the combination by eliminating certain items for given time constraints. Additionally, you might add open-ended or essay questions to the test. We see this as highly desirable and could be a means to increase the evaluations of higher order, analysis-like objectives. Multiple-choice questions can be turned into open-ended assessments by simply eliminating the distracters. However, be aware that this can make the same question either somewhat more difficult or much, much harder. Occasionally it can make a question easier, all depending on the individual question. Also, sometimes the answer choices of the multiple-choice question give a needed context for the type of answer expected. You may have to supply additional context to the stem if the question is to work satisfactorily as an open-ended question.

## Special Needs – Response to Intervention

Response to Intervention (RtI) is a general education initiative written into the special education law IDEA 2004. It offers a framework in which to structure Early Intervening Services. The RtI process involves gathering and examining assessment data. *Test for Higher Standards* provides the data for this process. Scoring services provided by *ROS* can be your solution to this part of the problem. Let us know if more information is needed.

## VMAST

### While VMAST as a program is no longer part of the SOLs K-8, we believe those materials are an excellent resource of open-ended items. Hence, we are including them as a resource.

During the SY 07-08 school year we had several clients ask permission to modify our items so they could be used as VGLA assessments. Given our materials are copyrighted; they were seeking permission to modify selected items. Without any hesitation, we not only said this was permissible but we encouraged them to produce assessments from our items. Since that time, the State has replaced VGLA with VMAST. Our existing VGLA materials were augmented with simple multiple-choice items to adapt to this transition.

We want to encourage you to use our items in the development of VMAST assessments. We require that the pages of these items include a notice of copyright. The footer from the Snapshot Blank file in the MAKING CUSTOM TESTS folder provides this. In addition, we have collected various VMAST assessments made by clients and placed them in a folder marked **VMAST** on the CD. We have not charged for this test bank, with the expectation that clients would share their own VMAST assessments with us. **We are grateful for the extensive assistance provided by Pulaski County Public Schools.** We offer a complete set of VMAST assessments for grades 3-8 Mathematics and ELA.

## Technology Enhanced Items

The recent push by the State for the increased use of Technology Enhanced Items on the SOL tests presents a significant departure from traditional multiple-choice testing. Given the widely varying availability and variety of technology platforms currently available in schools, we have chosen to mimic these advanced items via paper-pencil methods. For grades 3-high school, the folder for each grade/subject on your CD will contain a subfolder labeled **Enhanced Items**. This subfolder contains open-ended and multiple-answer items designed to simulate the types of Technology Enhanced Items found on the State tests. We call them synthetic TEI items, but that will allow for more than one correct answer. The actions used in answering these items will carry over to the items found on the State's TEI items.

## Making Benchmark and Snapshot Tests

Data driven instruction seems to be very effective and there is research to support this assertion. For a more ongoing diagnostic/remediation approach to instruction, we encourage the development of nine-weeks (six-weeks) Benchmark tests with an accompanying Class Matrix. In conjunction with this practice, we also encourage ongoing, systematic Snapshot testing: 3-5 items on a single page. This is a trend among our present users. We have developed software to assist in the development of Benchmark and Snapshot testing. The instructions and software are on your CD within the folder: **MAKING CUSTOM TESTS.**

Without charge, we will be delighted to share samples of these types of tests and related matrices. Within the MAKING CUSTOM TESTS folder we provide instructions on how to select items and build one or two column the tests. Additionally, we provide you with a cover and a blank page with headers and footers to facilitate the necessary cut-and-paste work. Just email your request to stuflanagan@aol.com, if you need assistance. Items from the State Released SOL Tests can also be captured with the snapshot feature in Adobe Acrobat and copied into our Benchmark and Snapshot forms.

**Extract from section VI. Research on Assessment (below):**

Feedback is a critical consideration to improving student achievement. What does research suggest regarding feedback?

- **The most powerful single modification that enhances achievement is feedback. The simplest prescription for improving education must be "dollops of feedback".**

*MAXIMUM POTENTIAL GAIN UP TO 35%*

Hattie, J.A. (1992). Measuring the effects of schooling. *Australian Journal of Education, 36(1)*, 5-13

***TESTS FOR HIGHER STANDARDS*,** through *Simulation Tests*, *Grade Level Tests*, 9-week Benchmark and Snapshot tests, and ongoing assessments such as the *TfHS Item Banks*, provide "dollops of feedback" that is based squarely on the State's standards and only on the State's standards. (For more on this topic, see **Section VI** of this guide.)

### Grades K and 1 and Possibly 2

For Grades K and 1, the teacher will want to give directions to the students on taking the test, by reading the items and pacing class progress. If necessary the teacher can talk the class through the problem solving steps for the first question of each type. Since these *Item Banks* were developed as a means of determining how much a student knows, please feel free to illustrate and clarify questions as appropriate. After all, it is necessary to understand the question in order to answer it.

### The Other Grades

For Grades 2-11, the questions were developed primarily according to the item specifications given in the test blueprints published by the State. In Grades 2 and 3, you may need to assist the individual student by clarifying or explaining the question. (Students who read very poorly could be tested separately in a small group. The questions could be read aloud to this group.) Beyond Grade 3, this should not be necessary, but we recommend giving individual question clarification if that is needed.

The Grade K-1 *Item Banks* were designed to be consumable in some instances, while the Grade 2-11 *Item Banks* were designed to accommodate reuse. One set can serve several classes as long as the order of SOL instruction is different. You may want to develop a bubble sheet for scoring beyond Grade 1.

As many of the state SOL tests cover content over one to three years, we urge you to build continuing skill maintenance in your instruction and assessments. This is especially true for grades where the SOL tests are actually given. Hence, you may want to pull selected items from different grades or earlier-taught SOLs to help students maintain their learning.

# IV. SUBJECT AREA ADDENDA

## Using the *TfHS Grade Level Tests in Mathematics*

### About Calculators

The State SOL test blueprints specify that students are allowed to use a 4-function calculator for a portion of the SOL tests in Grades 4-7, all of the test in Grade 8, and a graphing calculator is permitted for Algebra I, Algebra II, and Geometry. We trust your division or school has a policy in place regarding calculator usage. If not, we urge you to develop such a policy which would provide guidance for your teachers.

### Special Considerations for the Grade Two Mathematics Test

- You may allow students to mark answers in the test booklet rather than on an answer sheet.
- You may choose to read the test and pace the class.

### Facts Tests

The Grades 2 and 3 *Item Banks* contain a separate timed section. Allow 3-4 minutes respectively for each section of the 2nd Grade *Facts Test* (SOL 2.5; + and –) and 4-5 minutes each for each section of the 3rd Grade *Facts Test* (SOL 3.5; × and ÷).

### New 2016 Mathematics SOLs

We are developing a complete set of materials for the revised Mathematics SOLs to be fully implemented for SY 18-19. SOL tests for SY 17-18 will include field test items for the new standards. Thus, you may wish to introduce some new material to students for this year.

## Using the *TfHS Grade Level Tests in Science*

The science assessment materials were constructed with the understanding that process and content are both vital to and largely inseparable in science. Thus, these tests tap student knowledge of science "facts" and the students' understanding of how facts relate to each other and to the processes of scientific inquiry. The universal themes of science — change, equilibrium, interrelationships and the scientific dispositions or habits-of-mind — can also be found throughout the tests.

At all grades, each test question is designed to measure a single SOL (or a substandard). Because of this, specific diagnosis of student achievement on each SOL is possible. However, each test is more than a set of unrelated items. The item balance and coverage provide total test scores which are a meaningful index of how well a student has mastered the *overall* body of scientific knowledge represented by the SOLs at that grade or for that course.

### Suggested Science Test Use

This testing scheme is based on the common-sense model: First, test to find out what the students know and don't know; second, teach them what they don't know, introducing new material along the way; and third, test them again to discover what they have learned.

*Pre-post testing within a school year.* A teacher tests at the beginning and again toward the end of the school year. If the same test is used at the beginning as at the end, gain scores can be directly calculated. If different tests are used, a comparison of the content known at the beginning and at the end can be made.

*Post testing only.* Test scores from the previous year can be used to establish a beginning point for each year's instruction. The value of this approach is limited in grade eight and in the high school courses because the previous year's content is not completely relevant to present course content.

## Suggestions for Different Levels/Courses

In ***elementary grades two through five, and in grade six***, the science SOLs are somewhat general and cumulative. Because of this, it is quite appropriate to administer the science test for the previous level at the beginning of the year as a locator test/pretest. An alternative is to use the current level of the test. In either case, the test given near the beginning of the year is the pretest in a pre-post test scenario.

***Grade seven*** SOLs focus on life science and ***grade eight*** SOLs focus on physical science. The grade-six test is a logical pretest for use in both of these grades. The alternative of using the grade seven and eight tests for both pre- and post testing in grades seven and eight, respectively, is also appropriate.

The ***Earth Science*** course, generally given in the ninth grade, is an introductory compendium of a variety of science contents from the fields of astronomy, physics, chemistry, biology, geology, paleontology, meteorology, oceanography, ecology, and environmental science. All of this is combined with a strong focus on the processes of science. An appropriate pretest would be sections from *both* grade seven *and* grade eight tests, or the Earth Science test itself.

The grade seven test can be used as a pretest for the ***Biology*** course, although it may be too easy for some students. The Biology test itself can also be used as a pretest, because many students have been exposed to a substantial part of its content.

The ***Chemistry*** course is quite different from other science courses, in that a great deal of the material covered is *completely* new to most students. Previous levels of the science tests are of limited use as pre-tests for Chemistry because of this novelty of content. The Earth Science test is a possible pretest, as is the Physical Science test. The Chemistry test itself can be used as a pretest. In this case, the students need to be informed that they are not expected to know all of the material because they have not yet been taught it. However, this use of the Chemistry test as a pretest can serve to give students a general overview of what they are *expected* to learn in the course (and may scare them thoroughly).

*——— The English portion of the Virginia SOLs is covered by two, separate **TfHS Grade Level Tests**: Reading and Writing. ———*

## Using the *TfHS Grade Level Tests in Reading*

### Develop Sound Curriculum

At one level, the test items simply assess a student's ability to know or do what a SOL requires, and thus predict that student's ability to pass the state-mandated SOL tests. A student who has difficulty on ***Grade Level Tests*** will likely have the same difficulties to much the same degree on the state SOL tests. The primary difference is that while the state tests contain items from SOLs at several grades, A ***Grade Level Test*** measures only the SOLs for a single grade. You can give the appropriate ***Grade Level Test*** for each year, and — if you choose — you can give all the grades tested on the state-mandated test during grade three, five, or eight as a way to review.

In either case, it is **particularly important** to have a sound curriculum that correctly distributes content and skills from grade to grade. Otherwise, you will run out of time long before you run out of students' learning problems.

### Understand the Operations

In addition to predictive value, ***TfHS Grade Level Test*** items can help guide planning and instruction. Look at an item and consider what mental operations a student will need to accomplish in order to answer it correctly. Of course, we cannot see a student's reading operations, so teachers must in effect take themselves through the process. Usually, students need to go through several steps, and it is these steps that teachers will need to teach, in the most appropriate ways.

### Lead to the Standard

In almost every case, teachers and curriculum planners will discover that students need to take in information, grasp concepts, and master certain processes – on the way to becoming skillful at whatever a particular SOL requires. Instructional planning that starts with the perceived requirements of a type of test item, articulates these into reasonable instructional steps, and leads a student through these steps will be successful more often than not.

## Using the *TfHS Grade Level Tests in Writing*

***Grade Level Tests*** such as these can never replace the writing your students do, or the classroom discussions you and your students have about their words and works.

### Student Practice

What the ***TfHS Grade Level Test in Writing*** will do, however, is allow your students to practice thinking about writing in the way that will be required when they take the state SOL tests. It will also give them an opportunity to think about the decisions writers make — at a distance, without the anxiety that sometimes comes with deadlines and grades.

As your students work through a practice **scenario**, they will have many opportunities to think about a writer's voice, her tone, her language, her readers, and her purpose for writing. Each **scenario** asks students to consider:

- ways to plan and organize writing;
- the sound (voice and tone) as well as the meaning of words and sentences; and
- revision and editing concerns.

### Class Discussion

Most importantly, students will have examples of writing to discuss which was not written by either a classmate or a professional.

The class discussions that follow your students' individual decisions about the practice items can be invaluable. When students discuss the writing of a professional, they do so with a sense of awe . . . and of impossibility. They can't write like E. B. White or Langston Hughes, so why try? On the other hand, when they discuss the writing of friends and classmates, they are sometimes reluctant to discuss at all.

### Recommendations for Classroom Use

Each test will be useful for instruction, as well as diagnosis and test practice. For example:

- Students can take the practice tests, one scenario (assignment) at a time.
- Teachers can evaluate students' responses to that scenario before a class discussion.
- During class discussion, students can review the scenario, reading answer choices aloud and talking about the correct and incorrect answer choices.

Follow-up activities might include asking students to do the entire scenario assignment or to write pieces of the assignment: new beginnings or endings, or additional paragraphs.

### Editing Practice

Each ***TfHS Grade Level Test in Writing*** includes a large number of editing questions. These questions give students extended practice, using skills required by the Virginia SOLs, and in some cases, reflecting skills noted in many local division curricula. Editing passages can be useful as homework, as springboards to new writing assignments, or as the beginning of a lesson plan for a substitute teacher.

Each writing test also contains several prompts. After grade two, some are typical of those assigned by English teachers. Others are the kinds of writing students can do to learn about other subjects such as social studies and science.

### Evaluating Responses to the Writing Prompts

Although teachers can evaluate student writing in any of a number of ways, evaluating responses as they will be evaluated for state SOL tests will help students see the relative importance of particular writing concerns. In order to learn the scoring model used by the Department of Education in scoring

the direct writing assessment, we urge you to make use of the excellent training CD-ROM, *NCS Mentor for Virginia.* This is available to every school division in Virginia free of charge. Another excellent resource is *The Virginia SOL Writing Tests: A Teacher's Resource Notebook for Enhancing Writing Instruction and Improving Scores on the State Assessments*. This document is available on the Virginia DOE's web site at: http://www.pen.k12.va.us/VDOE/Instruction/English/writing/.

## Using the Item Banks for English

These *Item Banks* were developed to assist teachers in evaluating student achievement on a short-term basis. Unlike the *Grade Level Tests*, these tests are not to be taken in one sitting. Rather, they are to be used in evaluating achievement for individual or clusters of SOLs as you normally teach and test – ongoing.

It is important to realize, as you do, that it is not practical to isolate and teach and subsequently test the English SOLs one at a time. You need to do this in clusters. The index page of the *Item Banks* will help you in this regard.

As students in the earlier grades need to have materials read to them, we have paid special attention to this problem and directions are given on the test. You may need to employ small groups for some assessments in combination with using description/checklist type of evaluations. Realize the student needs to clearly understand the question to have an opportunity to answer it correctly. For Grades K and 1, you might find it necessary to read, explain, or clarify beyond what is given on the test. "Use good judgment" is our advice.

## Four Helpful Steps:

1. **Gather the information.**
   - Give either a *Grade Level Test* that covers all the SOLs for a grade, or use an item bank targeted to one or a cluster of SOLs.
   - Score the assessment using the answer key provided.

2. **Sort the information.**
   - *Group answers*. Group the answers by SOLs (defined here as the numbered Standard plus a bullet, indicated by a letter).
   - *Determine proficiency*. Use the percentage of correct answers to determine proficiency.

3. **Record the information.**
   - You may be using an NCR scoring format supplied by the state, or a form developed locally. In either case, the information you have sorted will fit into the form. If your form has only three categories, group your results into top (proficient and above), middle (minimal), and bottom (significant problems).

4. **Analyze the information.**
   - *Look for patterns.* In reading, for example, do scores show a problem in word analysis, information, or inference? Since this is a cognitive sequence — you can't get the literal information without the words, and you can't infer without the literal information — you need to see where the problem starts.

- *Consider the concepts involved.* More often than not, a student who has problems with a SOL never really grasped the essential concept(s) required by it. For example, it's hard to know about free markets if all you understand about money is that it's something in your pocket.
- *Think about the mental operations required.* Go back and look at the questions and the answers connected to a student's score. How did the student have to think to get the correct answer? How might he/she have been thinking to get that incorrect answer? Where did he/she get off track?

Steps 1-4 will help to keep your instructional focus on the SOLs. This is essential.

## Using the *TfHS Grade Level Tests in History and Social Science*

The good news is that the release of the *History/Social Science Curriculum Framework* makes Virginia's History standards clearer and more precise than they once were. The bad news is that Virginia's teachers still have to interpret, and sometimes guess at, the challenges their students will face when they sit down to do the SOL tests. The ***TfHS Grade Level Tests in History and Social Science*** give teachers a powerful tool for helping their students master the SOLs. They enable teachers to gauge how their students are doing on mastery of the SOLs, to focus instruction on critical points, and to understand how their colleagues are interpreting the most challenging parts of the SOLs. When it comes to interpreting the History and Social Science SOLs, having the ideas of other teachers can be extremely helpful. Most importantly, with the *Grade Level Tests*, adjustments can be made in time to help students succeed at each testing interval.

### New 2016 History and Social Science SOLs

We have developed assessments for the new 2015 *History and Social Science* SOLs. The new ***Grade-Level Tests*** and ***Item Banks*** are in a separate folder on your CD. These materials are provided for teachers who wish to introduce new content to their students for SY 17-18. HOWEVER, testing and remediation should be based on the existing 2008 *History and Social Science* materials. We are monitoring the Department of Education for further developments on the future of SOL testing in *History and Social Science*.

# V. VALIDITY AND RELIABILITY STATEMENT

## Tests for Higher Standards

Reliability and validity of tests are twin pillars which support the entire testing enterprise. At its simplest, reliability is the consistency with which a test measures any attribute. In both the state-sponsored SOL tests and the various ***TfHS*** tests, this attribute is some type of academic proficiency. The validity of a test is how well it does indeed measure what it purports to measure. For both SOL tests and the ***TfHS*** tests the goal is to measure student proficiency on the competencies stated in the Virginia Standards of Learning. As the overall function of the ***TfHS*** products is to provide focus and feedback for instruction, it is essential that our tests be both valid and reliable! Below are some of the types of evidence we have collected or produced. We are grateful to the Chesapeake, Essex, Hopewell, and Orange school divisions for sharing some of their internal results with us.

***TfHS* Reliability.** The evidence for the reliability for our tests presented here is based on the customary ***KR-20 internal consistency reliability estimates***. Each formula *KR-20* estimate is simply a statistical estimate of how each test question contributes to the overall score – averaged over all questions. It ranges from *0*, no reliability, to *1*, perfect reliability. Reliability estimates have been calculated for the current edition of our ***Grade Level Tests***. These were given in several grades and cover our five content areas (reading, writing, mathematics, science, and history and social science). Some example figures are presented in *Table 1* alongside some equivalent reliability estimates for the state-sponsored SOL Tests. The data were gathered from about 55 schools in several school divisions and were gathered in early spring of the year 2000.

| *Grade* | ***Reading*** | ***Writing*** | ***Mathematics*** | ***Science*** | ***History & Soc. Sc.*** |
|---|---|---|---|---|---|
| 5 | **.87** (.89) | **.74** (.84) | **.86** (.88) | **.88** (.81) | **.79** (.80) |

*Table 1.* Internal Consistency Reliabilities (KR-20) for ***TfHS Grade Level Tests***. The numbers are adjusted to correspond to the lengths of the respective State SOL Tests using the Spearman/Brown formula (The numbers in parentheses are the equivalent KR-20 figures for the State SOL Tests.)

However, in comparing remember that the two sets of tests have quite different purposes. The state test is designed to make just one primary determination: "Is the student at or above the passing score in a subject area or course?" Any other use of the score, including diagnostic uses, is quite secondary to the primary purpose of the State SOL tests. The ***TfHS Grade Level Tests***, on the other hand, are mainly aimed at diagnosis. We ask, "On which of the SOL skills are the students proficient and where are their weak areas?" This is the information a teacher will need to craft instruction. (Note that ***TfHS*** has never set a "passing score" on these tests. That is not their purpose.) Although the ***TfHS Grade Level Tests*** can determine how well a student or a group of students is doing overall, this use is usually secondary. The ***TfHS Grade Level Tests*** are also longer than the State SOL Tests and each of them covers just one grade of Standards. (However, the ***TfHS Simulation Tests*** are designed to be as much like the State SOL Tests as possible.). As the ***TfHS Grade Level Tests*** set out to measure many different proficiencies, the tendency to measure a single, overall trait may be slightly lower than the State SOL Tests. The reliabilities reported are highly appropriate for diagnostic tests.

***TfHS* Validity.** The primary validity evidence appropriate to the various tests we publish is ***content validity***. That is, "Do the tests adequately reflect the content stated or implied by the Standards of Learning themselves?" By "content" we mean knowledge, understanding, skills, habits of mind, and so forth, contained in the standards. Content validity is established in the beginning by having our authors keep the standards directly in their view as we write, review, and revise test items. Each item is directed at measuring a specific, individual standard. We have always had teachers, administrators, and curriculum specialists carefully review all of our tests for content validity. Any item that appears upon review not to match its stated standard is removed from our test. There have been relatively few of these. Occasionally, we also receive comments from our users concerning the appropriateness of test content. All such comments are seriously reviewed and changes are made to our tests, where appropriate.

When the ***TfHS*** tests were first developed, we worked intensively and extensively with all teachers and with all students of one school division to ensure that each question and the tests as wholes measured the important content of the SOLs. As much as possible, we tried to follow development procedures similar to those being used to develop the State SOL Tests. We also carefully consulted three different curriculum guidelines, representing the work of some 20 school divisions, as we developed and refined our tests. When the SOL *Teacher Resource Guides* were released by the state, we again reviewed our tests with teachers and administrators in the light of those guidelines and made what revisions were called for. In the case of the History and Social Science standards, the changes were substantial. In the other subject areas, fewer updates were needed.

Moreover, ***TfHS*** consults well-respected content specialists as we create and revise our tests. In the case of the reading and writing tests (English), Ms. Elizabeth Bradford and Dr. Kenneth Bradford, two well-known educators in Virginia were the original test authors. Dr. S. Stuart Flanagan, a long-time and respected mathematics educator is the primary author of the ***TfHS*** mathematics tests. He has had extensive experience from being a test grader at the Educational Testing Service (ETS) to writing test questions for the Virginia Department of Education. (He is also Professor Emeritus at the College of William & Mary and one of the ***TfHS*** partners.) In the area of History and Social Science, Dr. Richard Weber, Social Studies Supervisor with the Newport News Public Schools, coordinated the primary development of the ***TfHS*** tests. Finally, Dr. Ron Geise, a distinguished science educator at the College of William & Mary, consulted with us on Science issues.

The content validity of the ***TfHS*** tests is most important for an additional reason. We find that teachers find our tests especially helpful because they show very specifically what a Standard means. Tests are an excellent medium for conveying the meaning of educational objectives or standards, as each question is a concrete instance, rather than an abstract generalization. Students find our tests valuable, for the same reason – the tests help them understand what they must learn to be proficient on the Standards. Since these two audiences do look to us for guidance, we are especially careful not to lead them in the wrong direction. We take this responsibility seriously.

Another highly relevant type of validity is ***predictive score validity***. That is, "How well do scores on our tests predict scores on the State SOL Tests?" (However, remember that the *primary* purpose of the ***TfHS*** tests is *diagnosis*, not prediction.) We have some evidence of this type from several school divisions. One small school division ran a set of regressions to predict the score on the State SOL Mathematics Test at grade 3 on the basis of the ***TfHS Mathematics Grade Level Test*** given somewhat earlier in the year. In this case, which yielded a correlation of **0.89** between the two tests; it was found that a score of 85% correct on the ***TfHS*** test most closely predicted a scaled score of 500 (passing) on the state test. The ***TfHS Grade Level Test*** scores correlated with scaled scores on the State SOL Test **0.95** in grade 5.

In another division, administrators collected the scores shown in the in the table below. *Table 2* presents correlations between ***TfHS Simulation Tests*** percent-correct scores for individual students given early in spring and their later scaled scores on the State SOL Tests given in April.

| *Grade/Course* | *Correlation* |
|---|---|
| Reading, Gr. 5 | .61 |
| Reading, Gr. 8 | .54 |
| Mathematics, Gr. 5 | .71 |
| Mathematics, Gr. 8 | .64 |
| Algebra 1 | .76 |
| Geometry | .73 |
| Earth Science | .62 |
| Biology | .74 |
| Hist. & Soc. Sc., Gr. 8 | .79 |
| World Hist. 1 | .85 |
| World Hist. 2 | .81 |

*Table 2.* Product-moment correlations between ***TfHS Simulation Tests*** and SOL Scaled Scores for students in two or three classes in a small Virginia school division.

A different variety of ***predictive validity indicator*** is the school-level rank correlation between the average number correct score on the ***TfHS Grade Level Tests*** given in February and the percent passing the State SOL Test for that same school later in the year. We used rank correlations, as the percent-passing scores are not normally distributed. Over 31 schools, the grade-5 reading scores are correlated about **0.91** with percent passing reading. Over 32 schools the grade-5 mathematics was correlated **0.81** with the percent passing mathematics. Figures 1 and 2 are scatter-plots showing these data. The correlations are substantially higher than the correlations reported between the State SOL Tests and the *Stanford 9* tests (Reading, grade 5 was 0.78 and mathematics, grade 5 was 0.74.) Naturally, our correlations should be higher, as both the ***TfHS Grade Level Tests*** and the State SOL Tests were designed to measure the same Virginia SOL Standards!

The schools in this particular sample exhibited a *very* wide range of performance. The highest scoring school in the sample had a passing percent of 93 on the State SOL Reading Test and a passing percent 88 on the State SOL Mathematics Test; whereas, the lowest school's scores were just 17% passing the State SOL Reading Test and 12% passing the State SOL Mathematics Test. That high scoring school scored an average of 71% correct on the ***TfHS Reading Test*** and an average of 74% correct on the ***TfHS Mathematics Test*** in grade 5. The lowest scoring school scored an average of 40% correct on the ***TfHS Reading Test*** and an average of 44% correct on the ***TfHS Mathematics Test***.

Students at both of these schools fell within the range of valid measurement for the ***TfHS*** tests. Thus, these ***TfHS*** tests are suitable for measuring students' achievement over a very wide range.

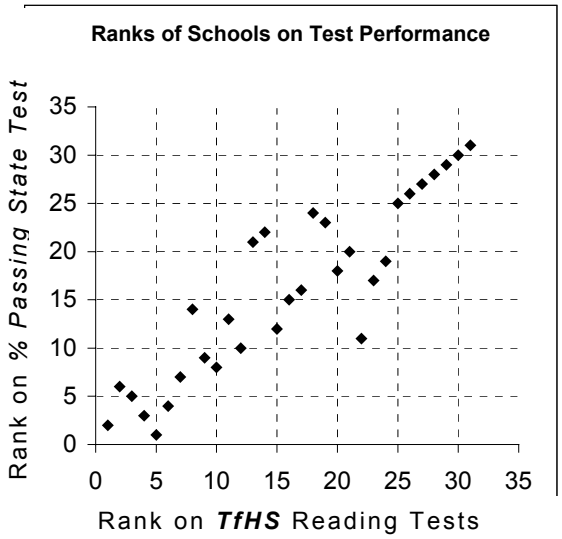


Figure 1. Relationship between percent-correct scores on the ***TfHS Grade Level Tests*** in **Reading** given in February and the percent passing the State SOL Tests given in the late spring of 2000. Plotted are the ranks for 31 schools. Rank 1 is the lowest.

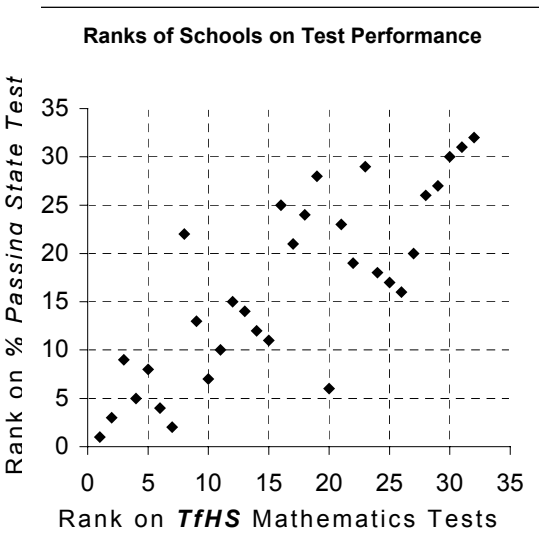


Figure 2. Relationship between percent-correct scores on the ***TfHS Grade Level Tests*** in **Mathematics** given in February and the percent passing the State SOL Tests given in the late spring of 2000. Plotted are the ranks for 32 schools. Rank 1 is the lowest.

Overall, we believe that ***TfHS*** has a set of valid and reliable assessments and would enjoy providing additional evidence to support this assertion.

David E. W. Mott
01/28/2001

# VI. RESEARCH ON ASSESSMENT

- **The most powerful single modification that enhances achievement is feedback. The simplest prescription for improving education must be "dollops of feedback".**

*MAXIMUM POTENTIAL GAIN UP TO 35%*

Hattie, J.A. (1992). Measuring the effects of schooling. *Australian Journal of Education, 36(1)*, 5-13

**TESTS FOR HIGHER STANDARDS,** through *Simulation Tests*, *Grade Level Tests*, 9-week Benchmark and Snapshot tests, and ongoing assessments such as the *TfHS Item Banks*, provide "dollops of feedback" that is based squarely on the State's standards and only on the State's standards.

- **Students need assessments that will tell them what they are doing is correct and what is incorrect. They need to know how to correct the performance and work on it until they do.**

***MAXIMUM POTENTIAL GAIN UP TO 30%***

Lysakowski, R.S. and Walberg, H.J. (1981). Classroom reinforcement in relation to learning: A quantitative analysis. *Journal of Educational Research, 75*, 69-77.

Lysakowski, R.S. and Walberg, H.J. (1982). Instructional effect of cues, participation, and corrective feedback: A quantitative synthesis. *American Educational Research Journal, 19(4)*, 559-578

Bangert-Downs, R.L., Kulik, C.C., Kulik, J.A., & Morgan, M. (1991). The instructional effects of feedback in test-like events. *Review of Educational Research, 61(2)*, 213-238

**TESTS FOR HIGHER STANDARDS** provide diagnostic data so that the student and teacher will know specifically what needs correcting and the related assessment tools to determine the performance in question is correct. The feedback needs to be standards specific to maximize student gains.

- **Assessment results need to be immediate. The greater the delay the less impact there is on achievement.**

***MAXIMUM POTENTIAL GAIN UP TO 20%***

Bangert-Downs, R.L., Kulik, C.C., Kulik, J.A., & Morgan, M. (1991). The instructional effects of feedback in test-like events. *Review of Educational Research, 61(2)*, 213-238.

**TESTS FOR HIGHER STANDARDS** provide assessments that allow for immediate feedback. There are several ways for this to happen depending on the scoring technique the teachers opt to employ: hand scoring, Scantron, online, student scoring, etc.

- **Corrections should be specific to the task/objective.**

Crooks, T.J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research, 58(4)*, 438-481.

**TESTS FOR HIGHER STANDARDS** identify the specific standard that each and every test item measures. The standard is coded next to every item on each and every *TfHS* test. This better enables the teacher to assist the student in making specific corrections of specific, individual standards.

- **Students could use TfHS's Classroom Matrix to know and evaluate their progress.**

Trammel, D.L., Schloss, P.J., Alper, S. (1994). Using self-recording and graphing to increase completion of homework assignments. *Journal of Learning Disabilities, 27(2)*, 75-81.

**TESTS FOR HIGHER STANDARDS** offer the student a means of mapping and/or seeing their own progress. Each and every TfHS test has an individual student response sheet that indicates the level of student achievement. For *Grade Level Tests* and *Simulation Tests*, there is an individual student profile and related **Classroom Matrix** that demonstrates what the student knows and the related ongoing progress or lack thereof.

*For more information on this topic that is in summary form, you might want to obtain the following ASCD publication:*

**CLASSROOM INSTRUCTION THAT WORKS**
***Research-Based Strategies for Increasing Student Achievement***
***Robert J. Marzano, Debra J. Pickering, and Jane E. Pollock***