

PRIMARILY FOR PARENTS:

TEN “MUST-KNOW” FACTS ABOUT EDUCATIONAL TESTS

W. James Popham
University of California, Los Angeles

All parents, especially those who have children in school, need to know some basic things about the educational tests their children take these days with ever-increasing frequency. Typically, those tests will not only have a profound impact on a child's progress in school, but will also make a big difference in what happens to a child when school is finished. Given the heightened significance of today's educational tests, parents surely need to be more aware of what makes such tests tick.

Fortunately, the facts that parents need to know about tests are not that numerous, nor are they that complicated. Presented below are ten key things about educational tests that all parents need to know. Each of these “must-know” facts is followed by (1) a brief explanation and (2) a potential action implication for parents.

Fact 1. Educational tests are much less accurate than most parents believe.

Students' scores on educational tests are usually reported in numbers, sometimes numbers that even include a decimal point. That's true both for teacher-built classroom tests as well as for statewide tests or nationally standardized tests. Such numbers typically convey to parents a picture of measurement precision that is simply unwarranted. A child's test-performance on a given day can be influenced by physical factors (such as the child's having a cold) as well as by emotional factors (such as the child's having had an argument with family or friends). In addition, the tests themselves only *sample* a child's skills and knowledge, and this sampling is often far from sufficient.

Action Implication. *Never place all that much confidence in the accuracy of your child's performance on a single test. Only when the results of several different tests are reasonably consistent should you attach substantial significance to your child's test performances.*

Fact 2. Educational tests allow teachers to make inferences about a student's unseen skills or knowledge, but these inferences may or may not be valid.

The kinds of skills and knowledge that children possess can't be seen. They are, quite literally, *covert*, that is, hidden. When Sally correctly spells all 25 words on her teacher's spelling test, the teacher *infers* that Sally is a good speller. Similarly, when Jane misspells more than half of the test's 25 words, the teacher *infers* that Jane's spelling skills are weak, so they need to be improved. If, on the basis of a student's test performance, an accurate inference is made about the student's unseen abilities, measurement experts call this a *valid* inference. But *people, not tests* make these

* A somewhat abbreviated version of this essay appeared in the November/December 2002 issue of *Our Children*, Vol. 28, No. 3, pp. 4-6.

inferences. And people have been known to make mistakes. Some test-based inferences will be valid; some won't.

Action Implication: *When your child's teacher makes a test-based inference about what your child's covert skills and/or knowledge actually are, do not automatically assume that the teacher's inference is valid. It probably will be, but it may not be.*

Fact 3. High-stakes tests, depending on the particular way these tests are constructed, can have a decisively *positive or negative* effect on the education that children receive.

When students' performances on a test have important consequences for the students (such as diploma-denial) or have important consequences for the educators who taught those students (such as test-based evaluations of a school staff's effectiveness), the test is referred to as a "high-stakes test." There is ample evidence now on hand that if an *inappropriate* high-stakes test is used, its impact on schooling can be quite harmful. For one thing, important but untested content will be eliminated from the curriculum. Second, students will often be forced to take part in dull, excessive test-preparation drills. Finally, some teachers, because of pressures to raise students' test scores, may engage in dishonest test-preparation or test-administration practices.

On the other hand, if an *appropriate* high-stakes test is employed, it can not only supply accurate evidence for the evaluation of schools, but such a test can also help improve the quality of teachers' *instruction*. Clearly, everything hinges on the appropriateness of the particular high-stakes test that is used. Properly constructed high-stakes tests will make instruction better; improperly constructed high-stakes tests will make instruction worse.

Action Implication: *If your child attends a school in which a high-stakes test is being used, try to discover (by speaking with teachers, administrators, and other parents) what effect, if any, that test is having on instruction. If inappropriate high-stakes assessments are having a negative impact on your child's schooling, join forces with other parents to demand more instructionally beneficial high-stakes tests.*

Fact 4. Although *achievement* tests are designed to assess the skills and knowledge a student possesses, and *aptitude* tests are designed to predict a student's success in a subsequent academic setting, it is sometimes difficult to tell them apart.

The *Iowa Tests of Basic Skills*, the *Stanford Achievement Tests*, and the *California Achievement Tests* are examples of widely used nationally standardized *achievement* tests. These *achievement* tests are intended to measure significant skill or bodies of knowledge that children should learn. A standardized *aptitude* test, on the other hand, is intended to predict a student's subsequent academic performance. The SAT and ACT, for example, are academic *aptitude* tests administered to high school

students in order to predict what kinds of grades those students will earn when they get to college. The ACT and SAT are widely used as college admissions exams.

Although the purposes of aptitude and achievement tests are different, an inspection of the actual items used on these two types of tests often reveals only modest differences between achievement tests and aptitude tests.

Action Implication: *Both aptitude tests and achievement tests can give you a relative picture of your child's current performance levels. But because standardized aptitude tests play a significant role these days in determining which colleges your child will be able to attend, you need to give careful attention to such aptitude tests. (Please see Must-Know Fact Number 7.)*

Fact 5. Traditionally constructed standardized achievement tests can provide both parents and teachers with useful information about a student's relative performance levels.

A *standardized* test is any test that is administered and scored in a standard, predetermined manner. For over three-quarters of a century, almost all of America's standardized achievement tests have been based on a *comparative* measurement model in which one student's score is interpreted according to the scores of a representative collection of other students who, having taken the test previously, are called the test's *norm group*. In fact, traditional standardized tests are often described as "norm-referenced tests" because a particular test-taker's score is interpreted by *referencing* it back to the norm group.

Relative comparisons can be very illuminating. For instance, when parents learn that their son has scored at the 87th percentile in mathematics, but only at the 23rd percentile in reading, those parents can work with their son's teachers to improve his relatively weak reading skills.

But, because a student's relative standing typically doesn't change all that rapidly, traditionally constructed standardized achievement tests need not be administered every single year in order to supply parents with this kind of comparative information. A statewide or national standardized achievement test, if administered every three years or so, will usually supply teachers and parents with sufficient comparative data.

Action Implication *If there are no state or national standardized achievement tests administered in your child's school, you should urge your district's school board to authorize the occasional administration of such tests. On the other hand, if your district's policy is to administer traditionally constructed standardized achievement tests at every grade level, this constitutes assessment overkill. Try to get the district's educational policymakers to reduce the frequency of such excessive testing.*

Fact 6. Traditionally constructed standardized achievement tests, that is, tests dominantly designed to permit a student's test-score to be compared with the test-scores of other students, should not be used to evaluate school quality.

Because, when using a comparative approach to measurement it is *essential* that students' test scores be well *spread out* by a test, the developers of traditionally constructed standardized achievement tests must include items on their tests that do a good job in producing the necessary score-spread. Examples of such items are those more likely to be answered correctly by (1) students from economically advantaged backgrounds or (2) students who possess higher *in-born* academic potentials. Unfortunately, these items actually turn out to *measure what students bring to school, not what students are taught in school*. Such tests, of course, should not be used to evaluate a school's success. A school should be judged primarily according to what students have learned there.

Action Implication: *Find out if students' scores on a traditionally constructed standardized achievement test are playing an important role in any evaluation of your child's school. If so, work with other parents to try to get more suitable, non-traditional standardized tests used for that purpose. Standardized achievement tests can, and should, contribute to the evaluation of your child's school—but only if those tests are appropriate for that purpose.*

Fact 7. Even though only about 25 percent of a student's success in college is related to the student's score on aptitude tests such as the ACT and SAT, parents should (1) still help their children prepare for those tests but (2) avoid conveying a negative impression to a child whose test scores are not particularly high.

The research evidence on this point is quite conclusive. There are many factors far more influential than aptitude-test scores in determining how well a student will perform in college. Such things as a student's motivation, study habits, and interpersonal skills play powerful roles in shaping a student's collegiate success. Thus, parents should not think that a child who doesn't earn super-scores on a college-admissions test is destined for certain failure—in college or beyond! Numerous students who earn lofty scores on the SAT or ACT actually take an academic tumble when they go to college. There are many important kinds of intelligence, and the "academic" intelligence measured by most standardized aptitude tests is only one.

Action Implication: *First, if your child doesn't score all that well on the ACT or SAT, do not conclude that your child is "not bright." And definitely do not convey any such negative impression to your child. Those sorts of parental judgments are not only incorrect—they can have a harmful, and sometimes lasting, impact on a child's self-esteem. Second, because ACT and SAT scores do, in fact, play a significant part in current college-admissions decisions, be sure to provide your child with at least some preparation for those tests. For example, purchase one or two SAT or ACT test-preparation booklets. If you can afford it, enroll your child in an SAT or ACT test-preparation course. The predictive power of aptitude tests may not be all that accurate,*

but those tests continue to be a key factor in determining where your child can go to college.

Fact 8. Performance tests, although both difficult and costly to score, often yield the most valid inferences about a student's mastery of the most significant skills.

Some of the most important skills that we want students to master are difficult to measure. For instance, if educators wish to see whether students can write an effective essay, the best way to do so is surely to have students sit down and write such an essay (rather than answering a flock of multiple-choice items.) When students generate an original response to an essay-writing task, we call this a *performance test*. Performance tests require students to *construct* their responses "from scratch" rather than merely to *select* their responses from already-presented options such as those found in multiple-choice items.

Typically, students' responses to performance tests are evaluated using a scoring guide that's called a *rubric*. Some rubrics, such as those typically employed to evaluate students' written compositions, are not only useful for scoring students' responses, they are also helpful instructionally. This is because they clearly set forth (for teachers, for students, and for students' parents) precisely what factors are to be used in judging a student's performance. These instructionally helpful rubrics are *skill-focused* in the sense that they explain how a student's mastery of an important skill will be evaluated. But not all rubrics are useful, however, either for evaluating students' responses or for supporting instruction. Some rubrics are far too general to be very clarifying. Other rubrics are focused not on the *skill* being measured, but only on the specific *task* that has been used in a particular performance test. Neither these excessively general rubrics nor task-focused rubrics contribute much to improved instruction.

Action Implication: *Find out the extent to which performance tests are being used to assess your child's mastery of significant skills. If not, collaborate with other parents to urge that at least some key skills be assessed via performance tests. If performance tasks are used, however, be sure to review the rubrics that are employed to score students' responses. What is needed are skill-focused rubrics, that is, rubrics whose evaluative criteria help clarify how a student's skill-mastery is to be judged. If you encounter rubrics that are (1) excessively general or (2) focus only on a specific task rather than on a skill, urge that more suitable rubrics be employed.*

Fact 9. Although an *individual* student's attitudes or interests are almost impossible to accurately measure, the assessment of such *affective* outcomes on a group-basis can provide accurate and powerful evidence regarding the quality of a school's instructional program.

When educators talk about the *affective* outcomes of education, they are referring to (1) important attitudes such as students' attitudes toward learning or (2) key interests such as students' attraction to in reading or mathematics. Because these

affective variables are often so significant in determining children's future successes, it is surprising that we rarely see students' affect being assessed in our schools.

The use of *self-report affective inventories*, completed anonymously by students, can supply both educators and parents with important insights about students' affect. Such self-report inventories, however, are *suitable for identifying an individual* student's affective status. Yet, when students' anonymous responses to such inventories are collected and analyzed all together, they yield a reasonably accurate picture of the affective status of a student *group*. Thus, it is possible to obtain, at little cost, satisfactory estimates of affect dealing with such variables as students' (1) perceived safety when at school, (2) interest in various school subjects, (3) attitudes toward learning, and (4) confidence in being able to carry out significant school-stressed academic skills.

Most parents realize that if a child is taught to excel in mathematics, yet learns to hate mathematics in the process, the long-time payoff of such math-mastery will almost certainly be minimal. Insights about students affect, even if available only for groups of students, can help parents judge the quality of education taking place in their child's school.

Action Implication: *Check to see if there is any affective assessment going on in your child's school. If not, urge that some group-focused assessment of key affective variables be initiated.*

Fact 10. Because educational tests (both teacher-made classroom tests as well as commercially-developed standardized tests) vary in their quality, it should not be assumed that every educational test is a good one.

Many parents tend to defer to educational test-makers. That's unwise. Although most teachers try to do the very best job they can when creating their classroom assessments, some of those tests turn out to be less than wonderful. Few parents realize that the most of today's classroom teachers, during their teacher-preparation programs, were never required to complete *even a single course* dealing with test-construction. It is not surprising, therefore, that some teachers will create tests that are unlikely to yield valid score-based inferences about students' skills or knowledge.

But even commercially developed *standardized* tests also vary in their quality. We live in an era when more and more standardized tests are being required in the field of education. Not surprisingly, there is a limited supply of test-construction expertise in our nation. As a result, sometimes today's standardized tests, even those created by reputable test-development firms, will fail to be at the high quality level that those firms would prefer. Put simply, there's too much standardized test-development being demanded these days, and too few talented test-developers to satisfy such demands.

Action Implication: *Do not give automatic deference to any educational test, either one that's created by your child's classroom teacher or one that arrives,*

professionally printed, from a commercial test-development company. Use your own good sense in critically judging whether a given test is likely to yield valid inferences about your child.

Wrap-Up

As you can see, these 10 "must-know" facts aren't all that intimidating. They're simply based on common sense and what's been learned over the years by measurement specialists. Do parents need to tackle all ten action implications suggested here? Of course not. But undertaking one or two isn't a truly terrible idea.

If you want to follow up on any of the topics treated here, you might care to consult either of the following books, both of which were written specifically for parents. Remember, today's educational tests are having an ever-increasing impact on children's lives. Parents owe it to their children to learn more about those tests.

References

Popham, W. James. *Testing! Testing! What Every Parent Should Know About School Tests*. Needham Heights, MA: Allyn and Bacon, 2000.

Stiggins, R., and Tanis Knight. *But Are They Learning?* Portland, OR: Assessment Training Institute, 1997.