DOCUMENT RESUME

ED 148 889 TM 006 805

TITLE	On Further Examination of "On Further Examination." Quick Reference: Testing/SAT.
INSTITUTION	National Education Association, Washington, D.C.
PUB DATE	77
NOTE	32p.; For related documents see TM 006 750-776; For complete report, see TM 006 516
AVAILABLE FROM	Instruction and Professional Development Information Center, National Education Association, 1201 16th Street N.W., Washington, D.C. 20036 ($0.50)
EDRS PRICE	MF-$0.83 Plus Postage. HC Not Available from EDRS.
DESCRIPTORS	Achievement Tests; College Bound Students; *College Entrance Examinations; Educational Problems; Educational Trends; Performance Factors; Predictive Ability (Testing); *Scores; Secondary Education; *Teacher Associations; *Teacher Attitudes; Test Bias; Testing Problems; Test Reliability; *Test Validity
IDENTIFIERS	National Education Association; *Scholastic Aptitude Test; *Test Score Decline

ABSTRACT

Discussions of any test must focus on reliability, content validity, and interpretation of results. In this critique of On Further Examination--the report of the Advisory Panel on the Scholastic Aptitude Test (SAT) Score Decline--the National Education Association (NEA) finds no fault with the reliability of the SAT. Its content validity, however, is suspect. The NEA maintains that the SAT is really an achievement test, and questions its relevance to high school curricula. The SAT is constructed primarily by psychometricians and psychologists who may not be sufficiently familiar with secondary curricula to develop a test with content validity. The College Entrance Examination Board Achievement Tests, which draw heavily on the expertise of teachers for content, do not parallel the SAT score decline. The Panel is criticized for dismissing the issue of SAT bias by noting that racial differences occur on most standardized tests. Questions are raised about the desirability of an unchanging standard and about what a predictive test should measure. From the teachers' perspective, the SAT score decline raises important issues concerning: the quality of instruction, innovative teaching methods, the back-to-basics movement, literacy, automatic promotion, exit examinations at the end of high school, reduction in homework, individual differences in learning and maturation, standardized teacher examinations, and teacher organizations. (EVH)

Quick Reference: TESTING/SAT

On Further Examination of "On Further Examination"

A National Education Association Publication

On Further Examination of "On Further Examination"

This paper was written as one NEA response to the Report of the Advisory Panel on the Scholastic Aptitude Test Score Decline. We extend our appreciation to Roger W. Shuy and Nancy M. Yanofsky, of the Center for Applied Linguistics, who were consultants to the staff on its preparation. The Association's initial response to the panel report came from President John Ryor, whose editorial from the November-December 1977 issue of Today's Education *is reprinted in this booklet.*

The report of the Advisory Panel on the Scholastic Aptitude Test Score Decline has been entitled *On Further Examination.* This title is appropriate in two ways. First, the panel has reported to the public the results of its extensive "further examination" of the problem. Second, the examination is by no means finished. Many questions remain unanswered, questions that warrant—and demand—"further examination."

The panel is to be commended for its lack of indictment. Many questions were raised, many analyses undertaken, many theories entertained, many discussions held, some experts consulted; but even with all of the above, no single reason emerged as THE REASON for the test score decline. The report reflects on *why* the scores have declined but by no means explains all the reasons for this decline.

The College Entrance Examination Board (CEEB) and the Educational Testing Service (ETS) convened the advisory panel two years ago. Its charge was to examine the Scholastic Aptitude Test (SAT) score decline over the last 14 years. More specifically, the panel was charged with determining the reason(s) for the 49-point drop in the Verbal SAT scores and the 32-point drop in the Mathematical SAT scores.

As the panel's inquiry progressed, it became quite clear that it would not be able to identify or explain all the reasons for the declining test scores. It has done, however, a reasonable job of presenting its findings about what has happened over the last 14 years. It has reflected on the SAT itself, other standardized tests, the populations taking the tests, educational changes in school practices and staffing patterns, and a variety of social forces.

In attempting to discuss *why* the test scores have declined, the panel states, "Fourteen years of uninterrupted decline in the SAT scores create the illusion that there is some single force or closely related set of forces at work here. This isn't the case. The decline has developed in two distinct stages..." (p. 13).

The first stage of the decline took place between 1963 and 1970 and is described as the period of "compositional change." During this time an increasing number of students became test takers and thus potential college students. This larger group included more women, more minorities, more representatives of all socioeconomic levels, and high school students with lower academic ranking. This period

also marked the beginning of many more opportunities for students to go to college, which can be attributed to factors such as the advent of open admissions, the establishment of two-year colleges, and increased scholarship support for minorities.

The test score decline cannot be attributed solely to this compositional change in the test-taking group, however. The panel has emphasized that other factors in schools and in society must also be considered in order to understand the decline. Hence, the second stage, from 1970 to 1977, is regarded as the period of "pervasive change," a period characterized by uncertainty and unrest in our society. The remainder of this statement will address itself to the panel's analysis of what has contributed to "pervasive change." The panel report notes that this change has a tendency to become so elusive that a "now you see it, now you don't" phenomenon often occurs. "So much has happened that may have affected this record that there is no way of telling what did; the only evidence is circumstantial, leaving it hard to distinguish cause from coincidence" (p. 25).

In contrast with these carefully qualified findings, the panel report was interpreted rather loosely by the press. A representative sample of news articles, editorials, and letters to the editor reveals that the public has not been given an accurate picture of the report. The *Washington Star,* for example, editorialized that the panel was "almost too judicious in its nuances, distinctions and imponderables to account for the falling off of scores." Perhaps the press is the body referred to in the report itself when the panel notes that three false as-

sumptions underline most theories for SAT score declines, the first of which is that there is a simple, clear answer to the problem. The *Star* takes the NEA President to task for his criticisms of the SAT, observing that if teachers develop their own tests, it would be "a classic case of the inmates running the asylum." In that the panel report does not implicate teachers in any major way, such flamboyant verbiage by the *Star* is completely uncalled for.

Education, as every teacher knows, suffers today from the attitudes of the general public, which thinks it knows all there is to know about education, that special thought, study, or experience with education is not needed to qualify one to make grandiose statements about education. Since educators appear to be under siege, careless interpretation of an important report such as *On Further Examination* can be damaging to the profession. It is for this reason that this paper has been prepared.

On Further Examination concludes with the observation that readers who turn to the summary of the report for a quick and easy picture of the panel's view are themselves guilty of the sort of practice which may have contributed to SAT score declines. This paper is not a substitute for the report itself, which is a well-written and important document.* Rather, it will analyze, comment on, and set the report in the context of current American educational practice. It will note in particular what the panel has to say about the SAT, about teaching, and about selected aspects of our society. It will

On Further Examination, Report of the Advisory Panel on the Scholastic Aptitude Test Score Decline, may be ordered from College Board Publications Orders, Box 2815, Princeton, NJ 06540. The price is $4.

conclude with a brief discussion about the future, including a set of questions for even "further examination."

About the SAT

In order to adequately discuss any test, it is important to focus on the following three issues: reliability of measurement, content validity, and interpretation of results. Stated another way, the construction of any testing instrument must consist of three components: what is measured, how it is measured, and what the results mean.

1. *Reliability*

The panel has great admiration for the reliability of the SAT. Reliability in testing is the degree of consistency between two measures of the same thing. Assuring reliability is probably what the CEEB and the ETS do best. For example, the SAT tests compare a test taker's capacities with those of his peers and with students who took the test before. The test is also correlated with the student's grades in high school and, it is claimed, is actually improving as a predictor of success in the first year of college. Likewise, different forms of the test can be measured against each other for reliability. From this, the panel concludes that the SAT is not out of line with teaching practice in the secondary schools.

The NEA has no question about the efficiency and integrity of the ETS to do what it does best. And the panel is no doubt correct in its admiration for the attention to reliability which the SAT contains. But this is not really saying very much because any test can be technically reliable and still totally fail to measure what it intends to measure. The attractiveness of reliability is attached to the fact that it can

be quantified, thus giving the air of scientific respectability.

2. *Validity*

The panel report deals very little with the issue of content validity despite the fact that this is the central issue to those involved in standardized testing today. Recent federal hearings related to the language assessment of bilingual children in New York, Chicago, Philadelphia, and many other school systems have stressed this basic issue: "Does the test measure what it says it measures?" Since the SAT claims not to be an achievement test but rather a predictive measure-for-success-in-college instrument, content validity becomes a cloudy notion. The SAT is called an intelligence test by some, since the difference between measurements of intelligence and measurements of aptitude are slight at best. The basic (and relatively minor) difference between intelligence and aptitude tests is that aptitude tests measure specific factors rather than general ones while intelligence tests measure general factors rather than specific ones.

Thus the SAT measures verbal and mathematical factors rather than all the things one might want to call intelligence (provided we could ever agree on what they are). During the history of the development of intelligence tests, the term *intelligence* was thought to be unchanging and innate. To avoid the implications of innateness, test developers have used the term *aptitude*. Since these tests are used primarily to predict school success, the term *scholastic* was added. One could make the case that the SAT is really an intelligence test and that, from what little is known about intelligence

or how to measure it, the SAT is invalid. If one takes the position that the SAT is an aptitude test, however, the test is subject to analysis of *why* such group aptitude tests predict school achievement.

In theory it is possible to distinguish between aptitude and achievement. One is an innate property and the other is accomplishment. Once content validity has been determined, it is possible to measure achievement. How to get at aptitude, however, is not so clear. The SAT selected the mathematical and verbal areas as a locus for measurement. Exactly how this decision separates achievement possibilities from aptitude is still unclear. All children have knowledge, memories, and feelings about the subjects being measured. Furthermore, distinguishing among SAT verbal aptitude items and those items in verbal abilities on achievement tests is virtually impossible. Distinguishing between aptitude and achievement, though possible in theory, is impossible in practice. What is more, ability appears to increase as achievement level rises. Why is it, then, that group aptitude tests can predict achievement? *It is highly probable that they are really achievement tests.*

It has been necessary to follow this reasoning in order to address the issue of content validity. Once verbal achievement and mathematical achievement are recognized as integral to any assessment of the SAT, we must ask the question, How do we know that what we are measuring really matters? The panel did *not* address this question, although to its credit it did recognize it as a legitimate one: "We have accepted, for purposes of this inquiry

and report, the traditional value base of the SAT—its validity..." (p. 11).

Perhaps one of the most significant statements of the panel report grew out of a discussion of why scores on the Achievement Tests (which may also be elected by SAT takers) did not parallel SAT score declines. Scores on tests in English Composition, French, Spanish, Biology, Chemistry, and Physics have increased while other subjects reveal only small declines. Students whose scores went up in these areas showed declines in SAT verbal scores and increases in SAT mathematics scores. The panel lamented that it had not been able to analyze this phenomenon fully and concluded, "It is conceivably important that the College Board and ETS make much larger use of outside committees in connection with the Achievement Tests than with the SAT; the counsel sought is in the one case from experts in the particular disciplines, in the other more from *psychometricians and psychologists*" (p. 23; italics added).

This criticism of the SAT, though cautious and indirect, appears to be one of the central issues in accounting for the score decline. The SAT is actually a disguised achievement test which, in order to be valid, must better address what really matters in the two content fields which it measures. Like most objective tests, it is a discrete point measure. That is, certain items are selected from an inventory of all possible knowledge in order to represent that larger knowledge in as few questions as possible. Only those who know the content of the fields being measured have the key to content validity. Test makers create items to test only after content validity has been established.

3. *Cultural Bias*

The panel dismisses the charges of cultural bias by observing that a definitive analysis of cultural bias is virtually impossible. It appears to the NEA that definitive analysis of cultural bias is no more difficult than definitive analysis of many other educational concepts. Test questions of any sort involve matching the assumptions, values, and presuppositions of the test maker against those of the test taker. An understanding of basic principles of language and culture make the identification of potential cultural bias available to those who care to find it. However, even publishers of standardized tests have publicly admitted cultural bias and some have made serious efforts to combat it. By this we do not accuse the SAT of such bias but, rather, disagree with the panel about its possible identification. Nor do we understand what the panel might mean when it observes, "These same differences show up on most other standardized tests, and yet this proves nothing" (p. 16). Such an observation follows the logic of saying that because everyone steals, stealing is all right. The panel is quite right, however, when it indicts society as a whole for such bias. Having said this, it is odd to conclude that only the SAT has been preserved from such bias.

About Teaching

In an attempt to further examine the "pervasive change" contributing to test score decline, the panel report discusses ten issues which directly concern schools and teachers. From the teacher's perspective, the reported decline in SAT scores reveals a number of important issues.

1. *Quality of Instruction*

The panel clearly recognizes that the primary intent of the SAT is to predict success in the first year of college. Furthermore, the panel explicitly notes that the SAT is *not* intended to reflect the adequacy of the student's training. Thus the report does not authorize or condone criticism of schools or individual teachers on the basis of SAT scores (p. 5). No evidence has been presented that would lead to the conclusion that the quality of school instruction is what is being measured on the SAT (p. 11). The report does not deal directly with the issue of quality of instruction.

2. *Innovative Teaching*

The panel does not blame innovative teaching for the decline in SAT scores. It observes, in fact, "We find no evidence of any causal relationship between what are commonly referred to as 'experimental teaching methods' and the SAT score decline" (p. 41). The panel carefully examined the results of two different investigations into this topic and notes that there is neither consistent nor substantial association between student achievement and overall level of innovation across the grades. It actually advocates the search for new ways of addressing the learning needs of children: "We do not read the SAT score decline as an instruction that education in this country must or should be more rigid, more selective, more rejective, more uniform. Instead, the instruction is that education, especially secondary education, must become still more diversified, more varied—but without being watered down" (p. 31). The panel's insistence that innovation not

come to a grinding halt is encouraging and highly desirable.

3. *Back to Basics*

The panel does not wish their report or the SAT score declines to be thought to suggest a back-to-basics movement: "In our view, 'returning to the basics' would be wrong unless it included full reappraisal of what the right basics are—taking account of children's different rates and modes of learning and their different interests and plans for the future. The need is not to revert to uniform drills and exercises commended only by a traditional pedagogy, but to move ahead to a larger emphasis on the fundamentals of learning that can be identified as strengthening the base on which all students can build" (pp. 26-27). The report does not suggest a return to the "good old ways," but it does encourage a thorough reexamination of what the basics really are.

4. *Literacy*

The closest the panel comes to criticizing the schools occurs in what the panel calls its "firmest conclusion": that thoughtful and critical reading is not being demanded and carried out, and that careful writing apparently has about gone out of style. Although the panel makes this observation rather forcefully, the fault for the situation is not laid solely at the feet of teachers. The panel argues, in contrast, that the testing industry itself, which introduced quantifiable data, even in the liberal arts, has contributed to the situation which the declining SAT scores illustrate. The NEA is in general agreement with the panel that reading and writing are critical skills which

educators should address with added zest and knowledge.

5. *Automatic Promotion*

Although automatic promotion is thought to be a contributing factor in SAT score declines, teachers are not held totally responsible for this in the panel report. The report says, in fact, that teachers do not like this idea either. Thus the panel disagrees with those critics of the schools who observe that automatic promotion is the invention and fault of teachers. The many factors which have conspired to lead to automatic promotion (many of which reflect the goals of society as a whole) must be equally addressed. It is interesting to note that the panel did not take into consideration the available research on late retention, much of which indicates that retaining a child can also have negative effects on learning.

6. *Exit Examinations*

The panel flatly denies that evidence from SAT score declines suggests the need for the development of so-called exit examinations at or near the end of high school: "...we question an equally 'automatic' answer that such promotion should depend entirely on scores on still other standardized tests" (p. 30). The underlying assumptions of either minimal competency examinations or exit examinations are questionable at best. Once we establish minimal competencies we tend to get just that—minimal competence. One would hope for considerably more than this. To adopt yet another single measure of "competence" appears both unfounded and unnecessary.

7. *Reduction in Homework*

The question of reduced homework was not discussed in any depth by the panel but it is interesting to note that the responsibility for the presumed reduction in homework is not placed solely on teachers. Although the panel observes that SAT score declines seem to be related to this factor, it also concludes that it is impossible to tell whether reductions in such assignments are parent or teacher induced. In fact, the panel notes that there is no quantifiable evidence that homework has in fact been reduced.

8. *Individual Differences*

The crucial issue of individual differences was not seriously addressed in the report. Differences in learning curves and maturation were given only cursory attention. One would like to assume that the failure to discuss this critical issue was an oversight, but perhaps it was merely reflecting the state of the art in the field of educational research. From the perspective of the teacher, however, it is apparent that two students of equal aptitude may be in a different learning growth position at the point at which the SAT is taken, grossly differentiating their scores. Such conditions suggest only the obvious: that research be undertaken to see how well the SAT predicts second, third, and final years of college and beyond. The fact that the SAT predicts only the first year seems very suspicious; for that year is a major transition year for young people. It should be noted that the panel did little more than casually mention the important matter of individual differences, an issue that deserved careful attention.

9. *Standardized Teacher Examinations*

The panel argues against reliance on standardized teacher examinations to insure that schools have only qualified teachers. Despite the fact that approximately half of the state legislatures are considering legislation to require such exams as a condition to the hiring of elementary and secondary teachers, the panel concludes, "Our increased understanding of the workings of standardized student examinations in such situations does not, however, commend exclusive reliance on the use of comparable examinations for teachers" (p. 33). Thus, the panel observes, the declining SAT scores should *not* be used as evidence for the movement toward or for legislating teacher examinations.

10. *Teacher Organizations*

The panel in no way accepts the suggestion that teacher organizations such as the NEA or developments such as collective bargaining have contributed to the decline in SAT scores. It states, in fact, that these suggestions "offer little in themselves as far as explaining the decline in the test scores is concerned [and] we imply nothing one way or the other here about the broader values they reflect" (p. 42). The panel in no way casts aspersion on our existence!

About Society

A great deal of what the press has stressed from the report concerns the blame which society as a whole must share for the decline in SAT scores. The home environment comes in for serious questioning along with issues involving equal education, television, and motivation.

As usual, however, since the panel was careful not to assign blame where the evidence was not clear, little more than reflections or hunches are offered: "...the SAT should not become the sole thermometer for measuring the health of schools, family, and student" (p. 40).

The panel expressed belief that the traditional relationship between the school and the family has weakened seriously. Broken homes and changing family relationships, coupled with changes in educational practice, have "contributed to a strain on the parent-teacher relationship" (p. 34). Since no direct or quantifiable evidence is available to suggest that this situation contributes to SAT score declines, the panel expresses this opinion only in the most general terms. Improved home-school relations can contribute substantially to a better environment for learning. However, the profession is not expected to bring about this change single-handedly.

The panel clearly states that there is nothing in the SAT score decline to "warrant generalization about what is happening to the abilities at large of youth as a whole" (p. 21). Thus any portrayal of the "disintegration of youth," so popular in the current literature, cannot be taken as a reflection of or a contribution to such decline. In fact, the panel does not feel that the SAT is a broad measure of the general quality of youth: "It tells us nothing about young people's honesty and integrity, about whether they care about each other, or about a lot of other things that matter more than test scores" (p. 40). This cautious understanding of the weakness of test scores, including the SAT, puts such scores in proper perspective for stu-

dents, parents, administrators, counselors, or whoever else tends to overrely on such measures.

Much has been made of the so-called distractions which characterized American life in the past decade or so. Usually nominated are the war in Southeast Asia, the draft, riots, the corruption in high places, assassinations, etc. Television is also included as a distraction, but the panel has noted that this medium could be used effectively for instructional purposes. There is no rejection of television in the report; only a warning about its misuse.

Like many good things, television can be and often is misused. The panel notes that educators have the responsibility for making effective use of its potential rather than rejecting it out of hand.

Since much emphasis is placed on changes in society that have taken place in this country, at least since 1970, it is noted that the panel sees these changes reflected in our youth: "...students entering college during that period had gone through five or six years of national disillusionment, especially for young people, virtually unparalled in American history" (p. 37).

This "decade of distraction" should be kept in mind while observing the test score decline. In many ways this puts an additional burden on teachers who are asked to provide stability while other aspects of life are in chaos. The task of the teacher is seen by the panel to be far more complex than it is often assumed to be. Social changes must be considered when discussing any issue related to educational effectiveness.

About the Future

We wish to make an additional comment on the title *On Further Examination.* As noted earlier, the panel investigated past events that led to the SAT score decline and reported its *current* reflections about these events. All that remains is to point to the *future,* and the panel did not neglect this opportunity.

In reviewing the past 14 years of test score decline it became increasingly evident to the panel that no single cause could be identified. Even though the panel acknowledges that the quality of education is not yet where it should be, it does not despair. Rather, it urges educators to move ahead to the future. Perhaps the question of why SAT scores have declined is the wrong question to ask. Perhaps it is impossible to answer adequately. But whatever else the report offers, it urges those concerned with American education to make an even "further examination":

> So there is no *one* cause of the SAT score decline, at least as far as we can discern, and we suspect no single pattern of causes. Learning is too much a part of Life to have expected anything else.
>
> It would be too bad, furthermore, if our concentration on the implications of a decline in the statistical averages on a set of standardized examinations should seem to ignore how incomplete a measure this is of either educational or broader human purpose. While we ask why the scores on college entrance examinations have gone down, T.S. Eliot's probing goes much deeper: "Where is the learning we have lost in information?

Where is the understanding we have lost in knowledge? Where is the life we have lost in living?"

Yet in the panel's view of it all, the fact of the hard asking—of both kinds of questions—offers new promise of new answers. We find nothing in the record we have reviewed to discourage the conviction that learning in America can be made all that is hoped for it. What is clearest is the reflection, in the reactions to these test scores and to the poet's lament alike, of renewed purpose to implement these hopes. The future continues to seem a good idea. *(p. 48)*

If still further examination is given to this subject, some or all of the following questions might be asked:

1. *Is there an unchanging standard?*

It is unlikely that eighth-grade teachers would think it appropriate to give a test to one eighth-grade class in 1970 and to another eighth-grade class seven years later and expect the difference in scores to say anything useful. What would such a difference in scores mean? That the teacher is better or worse? That the students have gotten smarter or dumber? That societal values have changed? That our knowledge base is different? Can we, in fact, compare children of one set of circumstances with those of another? The CEEB expects teachers to believe that there is a single unchanging standard which can be measured and compared across time. Is this a realistic assumption?

2. *What should a predictive test predict?*

Should not the predictive power of

the SAT be measured beyond the first year of college? Why stop there? Why has it not seemed useful for the SAT to predict college graduation or some other measure of life success?

3. *How is validity assured?*

Who should be the arbiters of the validity of test questions relating to verbal and mathematics ability? Why have content experts been used more in the development of achievement tests than the SAT? How can validity be determined for a test which is said to measure aptitude but appears to be measuring achievement?

4. *How is cultural bias determined?*

Why is it that the panel declined to discuss the issue of cultural bias in light of the fact that other test makers not only have admitted such bias but are attempting to correct for it? What might the panel have done to assess bias? How might the SAT avoid it?

5. *What instructional changes are indicated as a result of the panel's assessment of the SAT score decline?*

The panel has strongly recommended that teachers not abandon innovation. They should not go back to "basics." In light of these recommendations, how should instruction change? Should we be concerned with developing minimal competency examinations? Should we increase homework?

6. *Should the SAT attempt to measure thoughtful and critical reading and "careful writing"?*

Can critical reading actually be measured in a timed, pressured environment? Who can adequately determine just what is thoughtful and critical reading? Can it be measured

by readability formulae? How can a test such as the SAT measure "careful writing"? Who should determine the criteria for just what careful writing really is? □

Declining SAT Scores

The recent report of the College Entrance Examination Board (CEEB) on declining Scholastic Aptitude Test (SAT) scores among high school students has done an excellent job in highlighting some of the crushing problems facing students, parents, and teachers—all of whom have a high stake in the future of the nation's public schools. The National Education Association shares concern with CEEB over the seriousness of these problems. No one is more aware of them than teachers themselves.

Teachers believe in high standards for their students. We also know that for teachers to teach, for learning to take place, students must be evaluated. But we believe strongly that learning must be evaluated in a variety of ways.

Each year, nevertheless, 1 million graduating seniors—about a quarter of their age cohort—will have prepared for the rite of passage from high school to college by taking the SAT.

For the past 14 years, SAT scores have dropped steadily, after remaining relatively stable during the preceding decade. Verbal scores have dropped 49 points, from 478 in 1963 to 429 in 1977; mathematics scores, 32 points, from 502 to 470. The question is, Why?

This fall the College Entrance Examination Board published the findings of a two-year study commis-

sioned to investigate the causes of the declining scores. The study was conducted by an advisory panel of 21 people chaired by former Secretary of Labor Willard Wirtz. It was sponsored and funded by CEEB and by the Educational Testing Service (ETS) —the General Motors of the testing industry and sole producer of the SAT. (CEEB sponsors the SAT; ETS develops and administers it.)

The final report, entitled *On Further Examination,* is impressive, broad-ranging, and written in immaculate prose. Its findings represent a comprehensive analysis of social and educational change from 1963 to 1977. Many of the conclusions reflect conditions in schools and society that teachers have known firsthand for years.

Despite its studied fairness, the report presents several problems. The panel points with pride to the SAT, which was designed in 1941 as an "unchanging measurement" to predict student performance in college. (Although specific test items change from year to year, the SAT remains essentially the same.) As the report itself documents, society and schools have changed drastically since 1963—and, in my opinion, certainly even more since 1941—but not the SAT. Is it not time to look at the test as well as the test results?

Before discussing the other problems with the report, let's look at the report findings.

The report states that the decline has developed in "two distinct stages, characterized by significantly different balances of...different causal factors."

In the first stage, 1963 to about 1970, the report says, as much as three-quarters of the decline was caused

by the "notable extension and expansion of educational opportunity in the United States" during that period. This statement refers to sharp increases in the proportions of students taking the test who came from groups that have always registered substantially lower-than-average scores on the test—students from families with lower socioeconomic status, minorities, and women (whose average scores on the mathematical portion of the test, but not on the verbal, have been lower than men's).

In the second stage of the decline, after about 1970, there were fewer changes in the demographic distribution of the SAT takers. During this time, the decline in scores swept across the board: scores of traditionally higher-scoring students dropped, and those of traditionally lower-scoring students fell even lower.

The panel attributes three-fourths of the decline in this period to the impact of "pervasive" social forces, whose effects it cannot document precisely. It does, however, cite six sets of developments that may have affected this decline:

- Characterizing the period of sharpest decline in scores (1972-75) and the years immediately before this decline as a "decade of distraction," the panel says that "there is simply no way of knowing how much the trauma between 1967 and 1975 of coincident divisive war,...political assassination,...burning cities, and the corruption of national leadership affected the motivation of the young people... and whether there was consequent effect on their college entrance examination scores." The report conjectures that "this probably made quite a difference."

• The panel also points to changes during this period in the role of the family in the educational process, noting particularly the increase in the number of children living in homes in which two parents are not present. Although lacking definitive evidence on this point, the panel concludes that the effect of these changes is negative.

• Noting that "by age 16 most children have spent between 10,000 and 15,000 hours watching television, more time than they have spent in school," the panel surmises that this activity distracts students from doing homework, competes with schooling in other ways, and has contributed to the decline in SAT score averages.

• "There has been a significant dispersal of learning activities and emphasis in the schools, reflected particularly in the adding of many elective courses and a reduction in the number of courses that all students are required to take." The panel says, however, that a "broadside condemnation" of electives is not the solution; it recommends instead "restoring the tradition of critical reading and careful writing."

• The report cites "clearly observable evidence of diminished seriousness of purpose and attention to mastery of skills and knowledge...in the schools, the home, and society generally." It points to the condoning of excessive absenteeism, grade inflation, automatic promotion from grade to grade, less homework, and easier textbooks.

• "There has been an apparent marked diminution," the panel finds, "in young people's learning motivation, at least as it appears to be related, directly and indirectly, to their performance on college entrance examinations."

The report does not fault teachers. In fact, the panel shows some understanding of their difficult situation. At the CEEB press conference to announce the findings, Harold Howe II, panel member and former U.S. Commissioner of Education, said there is evidence that the job of teaching is tougher today than ever before. The report's discussion of teachers is mainly descriptive, noting on the one hand a drop in the average years of experience of elementary school teachers, but pointing on the other to an increase in teachers' average educational levels. It also refers to a decrease in pupil-teacher ratios and cites an increase in teaching salaries.

The panel concludes that teachers' and school administrators' responsibility for what has happened "centers in their having made more concessions because of changing circumstances and demands...than has been good for anybody involved. But this becomes a hard question of how much choice they have had and of how the demands of a changing student clientele are best met."

How are the demands of a changing student clientele best met as they make the transition from high school to college? By a standardized multiple-choice test, designed in 1941 and remaining virtually unchanged, rigorously guarding the gates to the nation's colleges? The NEA does not think so. While the nation endures a "decade of distraction," the SAT sails confidently into the past.

In fairness to the panel, it must be said that they were not unaware of this issue. While they defend the SAT rigorously on its merits as a predictor of academic college performance, they recognize that high school grades are

even more valid predictors. They also acknowledge current broader criticisms of the SAT related to common misuses and abuses of the scores. In January 1977, CEEB issued a set of guidelines for avoiding misuse of the scores as measures of the effectiveness of elementary and secondary education in general. The scores are, nevertheless, often used erroneously as exclusive measurements of individual or institutional quality. The report notes that when this occurs, "a very real 'relevancy' issue arises."

A far more serious consequence of test misuse is that it may elicit the cry for "back to basics." The report itself is provoking this kind of misuse. When the public and state legislatures look at it, they may see only one issue, lower test scores. They may not heed the panel's thoughtful advice to schools not to become "more rigid,... more rejective, more uniform." They may not pay attention to the panel's observation that "the record may suggest as much about youth's inherent resiliency and the resourcefulness of the formal educational process under unusual circumstances as about deterioration in either personal or institutional fibers. We have wondered sometimes in the course of our inquiry why the score declines haven't been larger."

The panel believes that there is "almost certainly" some causal relationship between the shift from traditional courses to electives and the decline in the SAT verbal scores. But despite its warning against "oversimplistic interpretation of this finding," the knee-jerk response to declining test scores may be back to basics.

A most important part is that virtually no one is addressing the question

of the relationship between the SAT test and curriculum as it exists in 16,500 school districts in the United States. The problem does not lie in identifying and explaining causes for the drop in SAT scores; the committee did that and did it well. A fundamental value question must be answered: Should a SAT test which hasn't changed significantly in 36 years be allowed to become a major determinant of school curriculum?

If legislatures and school boards rush to unfounded judgments about their own curriculums (based on SAT scores), they will be giving assent to the SAT as the basis for a national standard.

In the final analysis, it might be in the best interests of public education to have some standards national in scope, but who sets those standards should be a subject of thoughtful decision. As things stand, curriculum is a primary province of thousands of local school boards. Who's to say that a few experts in the testing industry are better able to determine curriculum needs?

And so we come full circle to the SAT test itself. "The panel accordingly commends further inquiry by the Board and ETS into the function of tests at this critical passage point.... The purpose of such inquiry would be to identify and put in appropriate priority whatever can be distilled from current national concerns about the society's educational values, and then to evaluate the traditional tests in the light of that determination."

There is just one more problem. There is no time frame, no target date, no deadline. The need for urgency at the tranquil ETS campus in Princeton, New Jersey, is not the

same as in the high schools of Baltimore and Indianapolis. While ETS moves toward further inquiry, hundreds of thousands of students are being measured on archaic scales and found wanting. □

John Ryor
President, NEA
Today's Education
November-December 1977

Documents such as this are a part of our effort to help teachers teach. The size and the complexity of the teaching profession are reflected in the range and the diversity of these publications.

The realities of the classroom and the problems which constantly confront the professional teacher are often deeply rooted in larger social, political, and economic issues. For this reason a number of our publications, as well some other program activities, are directed at very specific groups within the profession.

The National Education Association works through its Instruction and Professional Development (IPD) unit to provide members with information, service, and resources related to the Association's goal area of Professional Excellence. For more information, write to the IPD Information Center, National Education Association, 1201 16th Street, N.W., Washington, DC 20036.

John D. Sullivan, Director
NEA Instruction and
Professional Development