DOCUMENT RESUME

ED 146 629 CS 203 772

TITLE SLATE (Support for Learning and Teaching of English). Volume 2.
INSTITUTION National Council of Teachers of English, Urbana, Ill.
PUB DATE 77
NOTE 22p.; Newsletter, v2 n1 through 8

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Bilingual Education; Elementary Secondary Education; *English Instruction; Language Arts; *Performance Based Education; *Phonics; *Standardized Tests; *Testing Problems; Testing Programs; *Test Interpretation

ABSTRACT

These newsletters, produced during the second year of activity of a standing committee of the National Council of Teachers of English, focus on standardized testing as it affects the teaching of English. The first newsletter describes committee activities and priorities and includes a bibliography of information on tests and evaluation. The other seven deal in turn with specific issues: minimal competencies and measures of competence, how to interpret standardized test results, phonics instruction, competency testing and bilingual/bicultural students, implications of the national Assessment of Educational Progress, strategies of response to the minimum competency movement, and implications of the Scholastic Aptitude Test score-decline report. (AA)

Vol. 2, No. 1 : November 1976

SLATE

Support for Learning and Teaching of English

EXECUTIVE COMMITTEE CONTINUES SLATE FOR 1977

The NCTE Executive Committee, at its May 1976 meeting, extended SLATE for one year. The decision to continue SLATE as a special committee of the Council was taken in order to allow more dialogue among members as to what SLATE can and should do, to discover how many contributors SLATE will have and how permanent their support will be, and to give further study to the question of NCTE's tax-exempt status.

In deciding to continue SLATE, the Executive Committee adopted the following stipulations:

1. The present committee, elected in San Diego, should function for another year.
2. SLATE should keep a clear and complete record of all money spent.
3. SLATE will make no policy on its own, but it may make recommendations to the Executive Committee.

SLATE PRIORITIES

Nearly 3000 English teachers responded to SLATE's questionnaire on important issues facing the profession. Each respondent was asked to check two issues, yielding the following order of priorities:

teaching load in English and the language arts
competency-based education
standardized testing
censorship of instructional materials
sexism in learning materials
students' right to their own language
restrictive copyright legislation
racism in learning materials.

Although teaching load received the largest number of votes (about a hundred more than the next issue), the SLATE committee believes that this is an area where its energies would have little immediate effect. Up for consideration and adoption at the 1976 NCTE Convention are three new policy statements on equitable teaching loads in English—elementary, secondary, and college. If these policy statements are endorsed in Chicago, SLATE will supply them, on request, wherever teachers believe the statements might help in reducing unreasonable assignments. The NCTE Executive Committee has also accepted a SLATE request that NCTE sponsor research on the effect of teacher load on student learning and teacher morale.

The next two issues—competency-based education and standardized testing—are both related to the general problem of measuring progress in English by objective, quantifiable instruments, determined and administered by outside agencies. Combined, these two problems received nearly twice as many votes as any other single issue. On the recommendation of the SLATE committee, the NCTE Executive Committee therefore charged SLATE to focus during the next year on the impact of testing programs on the teaching and learning of English.

SLATE will oppose the definition of English as only those matters which can be measured in quantitative terms. It will publicize stands NCTE has already taken on the socio-political elements involved in large-group testing, regardless of test-makers' claims of validity and reliability. SLATE may also, within the limits of its funding, establish a monitoring system to discover where action involving testing programs is being proposed; write to those responsible for the proposed actions; distribute documents, reports, new releases, etc., that might influence the decision; send experts to testify; provide parents, concerned citizens, legislators, and government officials with information on the ways in which IQ tests, aptitude tests, achievement tests, and entrance tests can label and damage students, and work with other associations alarmed over the misuse and misinterpretation of testing instruments.

Concentration on testing, of course, does not mean that SLATE has lost interest in other issues that affect English teaching. Concentration means only that limited energy and limited money will be directed toward what seems to most members to be most urgent.

"Pharis' First Law of Educational Evaluation: If you cannot quantify anything of value then assume that what you can quantify is related to value."—William L. Pharis, in a speech at the May 1976 Invitational Conference on Measurement in Education.

STARTER SHEETS AVAILABLE FROM NCTE

During the summer and early fall of 1976, the SLATE steering committee published a series of Starter Sheets dealing with the "back-to-the-basics" cry. Each of the short pamphlets—none is more than two pages—begins with a statement of the issues, explains professional viewpoints on those issues, and ends with some suggested strategies for action.

Seven were prepared and distributed as SLATE Newsletters: "What Are the 'Basics' in English?", "Back to the Basics: Grammar and Usage," "Back to the Basics: Composition," "Censorship of Instructional Materials," "Back to the Basics. Language and Dialect," "Back to the Basics. Spelling," and "Back to the Basics: Reading."

The Starter Sheets may be reproduced in quantity by anyone who would like to use them; single copies of the set of seven Starter Sheets are available at $1.00 from SLATE at NCTE, 1111 Kenyon Road, Urbana, Illinois 61801.

Anyone making a tax-deductible donation of $10 or more to support the activities of SLATE will receive a free set of Starter Sheets, as well as future issues of the SLATE Newsletter.

TESTING IS BIG BUSINESS

According to figures in *Publishers' Weekly* (April 19, 1976), sales of standardized tests in 1975 came to more than 36 million dollars, an increase of 45 percent from sales recorded in 1971. These statistics were compiled from reports issued by the Association of American Publishers.

Sherwood Davidson Kohn (*National Elementary Principal*, July/August 1975, pp. 11-23) gives what he calls some "scattered, incomplete information" as to where this money comes from:

- Six million elementary and secondary school children take SRA achievement tests every year, including the Iowa Tests of Educational Development.
- Two million high school upperclassmen take SAT every year.
- More than 4,000 of the country's approximately 17,000 school systems use Harcourt Brace Jovanovich's Stanford Achievement Test.
- Other testing companies, including McGraw-Hill's California Test Bureau and Westinghouse Learning Corporation, refuse to reveal the number of tests they process annually.

In the light of these figures, it is hardly surprising that the testing industry has responded with indignation to calls for a moratorium on standardized testing.

"Because the results are expressed in numbers, it is easy to make the mistake of thinking that the intelligence test is a measure like a foot rule or a pair of scales. . . . But intelligence is not an abstraction like length and weight; it is an exceedingly complicated notion which nobody has as yet succeeded in defining. . . . [I]f the impression takes root that these tests really measure intelligence, that they constitute a sort of last judgment on the child's capacity, that they reveal scientifically his [sic] predetermined ability, then it would be a thousand times better if all the intelligence testers and all their questionnaires were sunk without warning in the Sargasso Sea."—Walter Lippman, in *New Republic*, 1922.

MYERS REPLACES SUHOR

Charles Suhor, who was elected to the SLATE steering committee in San Diego last November, has resigned his position on the SLATE committee as a result of his selection as Deputy Executive Secretary of NCTE, succeeding John Maxwell who will leave the post in June 1977. Suhor will be replaced on the SLATE committee by Miles Myers, Castlemont High School, Oakland, California. Myers was active in the formation of SLATE and, following the 5 persons elected, was the candidate receiving the next greatest number of votes in the 1975 election. He will join the SLATE committee in its two-day preconvention session in Chicago.

MINIMAL COMPETENCY TESTS PROLIFERATE

Five states—California, Florida, Colorado, Virginia, and Maryland—have now passed laws requiring minimal competency testing in their public schools. The requirements vary from high school equivalency tests which permit students to leave school and effectively lower the age of compulsory attendance, to "basic proficiency" tests which must be given regularly and which require districts to provide remedial or tutorial services for low-scoring students.

According to *Update II: Minimal Competency Testing* (July 1976), a bulletin issued by the Education Commission of the States (ECS), four other states—Louisiana, Michigan, New Jersey, and Pennsylvania—have similar legislation pending.

The extent of concern with minimal competency testing was dramatically illustrated last spring when ECS and National Assessment sent out invitations to what it thought of as a small preliminary conference 32 states sent 125 representatives to the meeting. There the representatives considered the feasibility of creating a centralized bank of ready-made test exercises and concerned themselves with the problem of keeping the test items "secure."

In September 1976, an NCTE Ad Hoc Committee on Competency Testing met for three days in Urbana, in response to the Executive Committee's concern about appropriate NCTE action related to the movement for minimum competency testing. Among the immediate actions of the committee were the drafting of brief guidelines designed to counter the worst abuses and misuses of competency testing and the preparation of a resolution to be presented to the November NCTE convention. The committee also requested one of its members, Alan Purves, to draft a SLATE Starter Sheet on the issues of competency testing.

"The future of the movement [minimal competency testing] could on one hand lead toward a general improvement of education for all students or—in a punitive sense—return the schools to a screening role for society that was evident 50 years ago."—Chris Pipho, Associate Director of Research and Information Services, Education Commission of the States, 1976.

APPLEBEE APPOINTED AS STAFF COORDINATOR FOR SLATE

Acting on a request from the SLATE steering committee for more help at headquarters, the Executive Committee of NCTE has appointed Arthur N. Applebee to serve as staff coordinator for SLATE. Applebee, who joined the headquarters staff in August, is responsible for answering inquiries, writing news releases, coordinating the publication of SLATE materials, tapping the resources of NCTE as they are needed to support SLATE's work, and keeping records on finances and contributors.

". . . the SAT was not designed to measure school performance and should not be used that way. To single out the schools as being responsible for the decline is, by the nature of the test, unwarranted, unfair and scientifically unfounded."—College Board President Sidney Marland speaking to Chief State School Officers in San Diego, 1976.

ASSOCIATIONS CALL FOR MORATORIUM ON STANDARDIZED TESTING

Among the organizations which have recently called for a moratorium on one form or another of standardized testing are the Association for Supervision and Curriculum Development (ASCD); the National Association for the Advancement of Colored People (NAACP); and the National Association of Elementary School Principals (NAESP). The National Education Association (NEA), which called for a moratorium on standardized tests in 1972, temporarily set it aside in 1975 "in order to concentrate its energies in this area on lending support to affiliates as they implement strategies to challenge standardized testing; for example, initiating court actions on behalf of students or teachers, attacking specific test instruments, seeking alliances with other groups which have a vested interest in countering test abuse, cross-committee planning for remediation of problems related to testing, developing negotiation procedures and language dealing with testing issues."

At its March 1976 meeting, the Board of Directors of ASCD unanimously passed the following resolution:

> A norm referenced standardized test is any test that compares performance to predetermined norms and is administered in identical form to large numbers of students. The interpretation of tests is usually based on several fallacious assumptions about education. It is recommended that ASCD through its elected officers, its staff, publications, and programs become an advocate for a general moratorium on norm referenced standardized testing, and further, that the report of the Executive Director given in 1977 reflect the status of this effort.

At an invitational conference on minority testing, sponsored by NAACP in September 1975 and partially supported by CEEB and ETS, the first recommendation was

> That there be a moratorium on all current standardized tests unless such instruments conform to the recommendations set forth in this report.

The report offers an analysis of the ways in which standardized tests discriminate against minority students and makes a number of recommendations intended to reduce that discrimination. The report also includes disclaimers from CEEB and ETS, saying that although they believe standardized tests need improvement, they cannot support a moratorium. (*NAACP Report on Minority Testing,* NAACP Special Contribution Fund, 1790 Broadway, New York, New York 10019.)

The vigorous stand on testing taken by NAESP prompted one high-ranking testing official to label that association as "ill-formed, wrong-headed and irrelevant" and to call its publication, *The National Elementary Principal,* "spiteful to a degree unprecedented in a professional journal."

Although NCTE has not called for a moratorium on standardized testing, it has joined the North Dakota Study Group on Evaluation, representing 25 other groups, in calling for "new processes of assessment that are more fair and effective than those currently in use and that more adequately consider the diverse talents, abilities, and cultural backgrounds of children"; greater involvement of parents and educators in the planning and process of assessment; an explanation of the limitations of the assessment instrument used in all reports to the public; terms broader than single-score national norms (which can be misleading) in reporting educational achievement; the sharing of information about assessment among professionals, policy makers, and the public so that appropriate improvements and reforms can be discussed; the return of every standardized test taken by a child to the school for analysis by teachers, parents, and the child; and making any standardized test used in any community publicly available so that citizens can understand and review the tests.

Further, NCTE has included in its statement on Teacher Preparation and Certification the requirement that English teachers should have knowledge of "the uses and abuses of testing procedures and other evaluative techniques for describing students' progress in the handling and understanding of language" (*A Statement on the Preparation of Teachers of English,* NCTE, 1976).

"There is no evidence or reason to believe that taking standardized tests is a positive learning experience for children. Quite the contrary, the test situation can be a dehumanizing experience for children in which they are subjected to a standardized stimulus in an authoritarian setting which allows for virtually no creativity or individuality and subjects the child to an anxiety-ridden atmosphere."—Michael Q. Patton, representative of the North Dakota Study Group on Evaluation, 1976.

USEFUL INFORMATION ON TESTS AND EVALUATION

The profession has produced considerable information on what standardized tests can and cannot do, and has suggested some alternative ways of measuring progress. Notable among these publications are the following, all of them available through NCTE. (Stock numbers given in parentheses will expedite your order.)

Common Sense and Testing in English, by the Task Force on Measurement and Evaluation in the Study of English, Alan Purves, chair, 1975, $1.00 (No. 07737R). Describes the types of tests commonly given, discusses their limitations, and suggests ways to make these limitations known; lists criteria for selecting standardized tests in English and for interpreting and using test results; includes a three-page "Citizen's Edition" which can be reproduced and distributed to administrators and parents.

National Assessment and the Teaching of English, by the Committee to Study the NAEP, John C. Mellon, chair, 1975, $3.60 to NCTE members (No. 32235R). Gives the results of the first National Assessment of Educational Progress in Writing, Reading and Literature, interprets the factual data from a number of professional perspectives.

Reviews of Selected Published Tests in English, Alfred H. Grommon, editor, 1976, $3.90 to NCTE members (No. 41218R). In part one, considers complaints that many tests are culturally biased or based on outdated curricula and suggests questions that should be asked during test selection. In part two, evaluates more than 50 widely used English tests.

Equivalency Testing: A Major Issue for College English, Forest D. Burt and Sylvia King, 1974, $1.75 to NCTE members (No. 15640R). Discusses the role of ETS and CEEB, the policies and procedures of CLEP, the role of college professors, testing centers, and administrators in implementing equivalency testing programs. Includes the California report in its entirety and resolutions from the ADE-Bradley conference which focused on the politics of CLEP.

Testing in Reading: Assessment and Instructional Decision Making, Richard L. Venezky, 1974, $1.00 to NCTE members (No. 53321R). Provides ten "canons" for developing assessment procedures in reading and for using the results in ways which will most benefit students; considers problems related to program assessment and to the distribution and protection of assessment results; emphasizes that assessment must be secondary to instruction.

Measures for Research and Evaluation in the English Language Arts, William T. Fagan, Charles R. Cooper, and Julie M. Jensen, 1975, $5.50 to NCTE members (No. 30992R). Describes more than 100 unpublished instruments for research and evaluation in language development, listening, literature, reading, "standard English" as a second language or dialect, teacher competency, and writing.

Measuring Growth in English, Paul B. Diederich, 1974, $2.50 to NCTE members (No. 31093R). Shares a lifetime of experience and research in testing to show that teachers don't need wide knowledge of statistics to measure students' learning; shows how to increase the reliability of essay grades, how to measure growth in writing ability, how to reduce the time and anxiety involved in measurement, and how to use results to improve relations between students and teachers.

SLATE ADDED TO COSPONSORED SPEAKER PROGRAM

The NCTE Executive Committee has added members of the SLATE steering committee to the list of cosponsored speakers available to all NCTE affiliates during the 1976-77 academic year.

Under the terms of the program, affiliates may request a speaker for an official affiliate meeting. The local affiliate contributes what it can to the speaker's travel expenses and NCTE assumes responsibility for remaining expenses. Speakers agree to serve without honorarium.

Affiliates wishing to take advantage of this service send a Speaker Request Form to Nancy Prichard, at NCTE, who makes the rest of the arrangements.

"Unfortunately, the rank and file of test users do not appear to be particularly alarmed that so many tests are either severely criticized or described as having no validity. Although most test users would probably agree that many tests are either worthless or misused, they continue to have the utmost faith in their own particular choice and use of tests regardless of the absence of supporting research or even of the presence of negating research."—Oscar Buros, in the preface to the *Sixth Mental Measurements Yearbook*

CUNY COMMITTEE CRITICIZES COLLEGE BOARD'S NEW USAGE TEST

This fall the College Board decided to add a short objective test of "Standard Written English" to the SAT. Scores are not incorporated in the SAT scores but are listed separately.

CAWS, an acronym for the CUNY Association of Writing Supervisors, examined the test when it was presented in trial form by ETS and administered it to students at one of the CUNY branches. The committee found that the level of the test was inappropriate for CUNY freshmen, that the exam did not focus on the kind of skill problems CUNY wished to detect, and that the short-answer editing questions did not indicate students' ability to generate sentences in their own writing. (*College Composition and Communication*, October 1976, pp. 287-9.)

"No one should interpret the current controversy over standardized tests as an effort to abandon assessment. Rather, it is an effort to develop assessment procedures that are more in keeping with a new set of educational and social assumptions that we as a society are working on: that the purpose of education is not to sort people but to educate them; that in a knowledge society we need to expose as many people to education as possible, not to exclude them from it; that human beings are marvelously variegated in their talents and abilities, and it is the function of education to nurture them wisely and carefully; and, not least, that education has an overriding responsibility to respect and draw on cultural and racial diversity. Assessment of students must begin to reflect that philosophy, and that is the true reason for the current call for test reform and an end to IQ testing!"—Paul L. Houts, in *Phi Delta Kappan*, June 1976.

HOW SLATE SPENDS ITS MONEY

SLATE's income for fiscal year 1976 was $18,146.00—a $10,000 appropriation from NCTE and more than $8,000 in contributions from 651 individuals and affiliates. During the same period, SLATE expenses came to $11,614.65, the majority of that sum going to headquarters services—production and distribution of Newsletters and Starter Sheets, staff services, etc. The 5-member Steering Committee met twice, once in January and once in May, and expenses for these meetings were also covered by SLATE funds.

At its May meeting, SLATE authorized two special expenditures. The committee agreed to underwrite Robert Hogan's expenses in testifying before the National Assessment of Educational Progress in Washington, D.C., and appropriated $250 to distribute copies of *Common Sense and Testing in English* to school boards, superintendents, legislators, or other concerned groups.

SLATE Newsletter will be published periodically by the National Council of Teachers of English and is directed to NCTE members who contribute $10 or more to support SLATE. Contributions should be sent to SLATE, NCTE Headquarters, 1111 Kenyon Rd., Urbana, IL 61801.

SLATE needs your support—have you contributed yet this year?

Name _____

Address _____
Number Street

_____
City State Zip

Yes, I want to contribute. Enclosed is my donation of $10.00 ___ ; $15.00 ___ ; $20.00 ___ ; $25.00 ___ ; Other ___.

First contribution ☐ Contributed previously ☐ NCTE member ☐ nonmember ☐

Make checks payable to NCTE/SLATE and mail to SLATE, NCTE Headquarters, 1111 Kenyon Road, Urbana, Illinois 61801. All contributions are tax-deductible.

Suggestions for SLATE activities? Send them to SLATE at NCTE.

Vol. 2, No. 2 December 1976

SLATE

Support for Learning and Teaching of English

MINIMAL COMPETENCIES AND MEASURES OF COMPETENCE

The Issues

Over the past few years there has been an increasing demand at local, state, and national levels for the establishment of a list of essential skills or 'minimal competencies' which every child in school should be expected to attain. The reasons for this demand are many: among them the desire of educational systems to have some sort of accountability system, the dissatisfaction of many with the still large number of citizens who appear unable to read or write; the apparent decline in students' performance on standardized tests, the uneasiness of many with an elective curriculum, and the increasing lack of local support for schools as reflected in tax and bond referenda. As a result certain administrative units have set forth requirements for verified competencies of secondary school students—states like Arizona, California, Oregon and communities like Gary (Indiana), Duval County (Florida), Westside Schools (Omaha, Nebraska) are but a sample of the many states and districts using minimal competency tests. At the national level, the National Association of Secondary School Principals has endorsed the recommendation of its Task Force that there be uniform graduation requirements verified by competency measures as well as by units or credits.

The implications of this movement are not entirely clear, but certainly it raises a host of potential problems. These include:

1. the possibility that scores on competency tests will determine promotion or non-promotion and thus will lead to a return to grade-repeating, a practice which disappeared from American schools after research pointed to its ill-effects on students and on schools;
2. the possibility that there will emerge diplomas of different 'classes' dependent on the student's scores on a test, with attendant danger of social stigmatism;
3. the possibility that statements of competence will lead to a circumscribing of the curriculum to a point where it will be merely preparing students for the test rather than educating them broadly;
4. the possibility that statements of competence will lead to didactic teaching as a shortcut to improving student performance, when research has indicated that such teaching is not as beneficial as more varied approaches to teaching;
5. the possibility that educational resources will be concentrated upon those students who have trouble attaining the minimal level of competence, and will be diverted from the gifted and the broad range of students who can attain the minimal level fairly easily but need to go beyond it;
6. the possibility that educators and the public will be satisfied with bringing students to a minimal level of competency rather than to a level of mastery that is appropriate to their age, their view of themselves, and their aspirations;
7. the possibility that statements of competence and measures of competency will make schools less responsive to the cultural and linguistic diversity of this country.

Obviously these implications bear social and political dimensions which can only be explored through observation of the long-term effects of competency programs and tests.

Professional Viewpoints: NCTE/Research

There has been a great deal of study of various aspects of testing as well as of aspects of goal-setting, competency definition, and behavioral objectives.

The research on testing is too extensive to be dealt with justly in a brief space; its major areas have concerned reliability (how effective a test is in measuring a phenomenon accurately) and validity (how well a test measures a phenomenon or acts as a predictor of subsequent behavior). Reliability studies usually deal with whether a test is internally consistent, whether it measures a student's performance the same way when it is given to the student again, or whether two versions of the same test produce the same scores when a particular student takes them. Reliability research indicates that standardized tests are generally consistent in one of these ways, but such research does not indicate anything about the usefulness of the tests concerned.

Validity studies seek to determine whether an objective test in writing, for example, is as good a measure of writing performance as is the grading of a set of papers from a particular student. Some validity studies use experts to examine the test and to say whether the questions are dealing with appropriate material, or following some established theory; such studies determine the "content validity" of the test. Examples of such studies can be found in O. Buros' *Mental Measurements Yearbook* (N.J.: Gryphon Press) and A. Grommon, *Review of Selected Published Tests in English* (Urbana: NCTE, 1976). The other form of validity research is called "empirical" or "criterion" validity. It seeks to determine whether a test measures some attribute of learning that can be observed in practice. A test of reading interests might be validated against the library withdrawals of a particular

student, or a test of spelling against a count of spelling errors in a student's writing. Validity studies are much more difficult to construct than are reliability studies, very few get beyond content validity to empirical validity. A major validity study of the College Entrance Examination Board's Achievement Test in English showed that the test was somewhat valid but that its validity could be increased by the addition of a piece of writing that was graded by trained readers (Goldschalk, Swineford, and Coffman, *The Measurement of Writing Ability*. New York. CEEB, 1966). Other empirical validity studies can be found in various articles in *Research in the Teaching of English*, and in Fagan, Cooper, and Jensen, *Measures for Research and Evaluation in the English Language Arts* (Urbana: NCTE and ERIC/RCS, 1975).

As far as can be determined, there have been no empirical or criterion validity studies of competency tests.

Research in competency statements is generally meagre, though they could be validated against the opinion of experts or against the kinds of tasks actually performed by competent citizens. In mathematics, such a validation might begin with a look at the kinds of mathematics people employ in different work and family situations, and continue with an analysis of what might be general levels of competence and what might be more specialized levels (for a mechanic, say, or an engineer). In English, a similar validation procedure could take place. What kinds of writing do people actually do? Although competence statements have been attempted for the National Assessment of Educational Progress and in the course of studies of mastery learning (J.H. Block, *Mastery Learning* [N.Y.. Holt, Rinehart, and Winston, 1971]), they have come from experts' opinions rather than from a careful scrutiny of what goes on in the world.

Strategies for Action

Teachers in those states or districts that are considering legislation related to competency programs have a different problem from those in locations where legislation already exists. In the first case, teachers and teachers' groups have an opportunity to lobby concerning the direction of legislation. A useful source for potential lobbyists is the publication of the American Library Association, *Who Me, A Lobbyist?*, available from the ALA Washington Office, 110 Maryland Avenue N.E., Washington, D.C. 20002. In the second case, teachers have to work to make sure that the criteria of competence and the measures of competence meet their own professional standards. The following guidelines, drafted by an NCTE *ad hoc* Committee on Minimal Competency Testing, may provide a useful starting point.

GUIDELINES FOR THE DEVELOPMENT OF COMPETENCY STATEMENTS AND MEASURES OF COMPETENCE IN ENGLISH

Competency Statements Must Be Sufficiently Comprehensive So as to Cover the Many Facets of English

1. Competency statements in English must include the broad areas of reading, writing, oral language, and media literacy.

2. Competency statements in English must include a variety of contexts (e.g., statements should not be limited to business communication).

3. Competency statements should emphasize those aspects of English which are critical in effective communication, as opposed to those which might be trivial or less significant (e.g., clarity of expression in a letter of application is more critical than using the appropriate state abbreviation in an address).

4. Competency statements in English must reflect the developmental level of the students (e.g., criteria for making inferences about reading must consider the potential ability of students at a given age to make inferences of a particular level of abstraction).

Measures of Competence in English Must Be Many and Varied

1. Measures of student competence in English must include not only tasks dealing with analysis but tasks calling for performance in oral and written language.

2. Measures of student competence in English must include tasks that are to be evaluated for their overall effectiveness by teachers or other professionals in English.

3. Measures of student competence in English should include language presented orally, in print, and through electronic media.

4. Measures of student competence in English must include tasks which are to be observed and evaluated by English teachers on a continuing basis (e.g., there should be some report on the student's use of language in daily situations as well as in specific test situations).

5. Measures of student competence in English must include provisions for superior performance (e.g., a measure of student performance in writing must include the possibility that a student could be judged an excellent writer so that incentives to excellence can remain in the curriculum and the testing process).

Alan C. Purves
(For the SLATE Steering Committee)

Vol. 2, No. 3 March 1977

SLATE

Support for Learning and Teaching of English

STANDARDIZED TESTING: HOW TO READ THE RESULTS

The Issues

Standardized tests continue to loom very large in decision-making processes affecting the world of English teaching. Serious consequences arise from "misreadings" of standardized test results: unwarranted inferences and conclusions about student learning, teacher performance, and curriculum change accrue from misconceptions about the nature and the meaning of results from standardized, multiple-choice tests. In order to counter misuses of such tests, English teachers need to become "close readers" of standardized tests and the manuals accompanying them.

Professional Viewpoints and Research

Several research studies have indicated that many professionals (teachers, counselors, administrators, and education agency officials) are relatively naive about the limitations and appropriate uses of standardized test results. David Goslin's survey of teachers and counselors in 75 secondary schools found most of the teachers to be relatively uninformed; similar findings emerged in a companion study of teachers in 800 elementary schools (*Teachers and Testing* N Y: Russell Sage Foundation, 1967). Examining uses of standardized test data by Michigan administrators and education agencies, House, Rivers, and Stufflebeam reported a number of questionable practices in the development and use of standardized tests in the Michigan Accountability System. These investigators found "serious errors" in the use of test results, reflected in such practices as: (1) tying district funding to the gain scores of low achievers, (2) using test scores as a major criterion in evaluating teacher performance, and (3) interpreting test scores as if they were literal indicators of what is taught and learned in school. House, Rivers, and Stufflebeam's expert opinion about the effectiveness of standardized tests as measures of school learning is especially illuminating.

> Test results are not good measures of what is taught in school, strange as it may seem. They are good indicators of socioeconomic class and other variables. But, unless one teaches the tests themselves, they are not very sensitive to school learning. (*Phi Delta Kappan* 55.10 [1974], 663-69.)

Ronald P. Carver demonstrates in another study why achievement tests are inappropriate measures of school learning. In his appraisal of the standardized tests used in the famous Coleman Report (the STEP Tests), Carver explains that these tests, designed according to traditional test-making principles, eliminate those items that all students got correct during development of the tests. Even though these items might be very good measures of achievement, they would be very poor items for producing variations in student performance. Hence, the "good" achievement test that is built upon psychometric principles is not set up to reveal what "everybody knows" about a subject; rather it is arranged to maximize variation in performance in order to produce student rankings. Carver concludes that achievement tests so constructed actually emphasize *aptitude* instead of *achievement*. (*American Educational Research Journal* 5.4 [1975], 77-86.)

A growing national concern over widespread misuses of standardized tests has led representatives of 40 major educational organizations to organize a National Symposium on Testing under the auspices of the North Dakota Study Group on Evaluation and the National Association of Elementary School Principals. This group, with which NCTE is affiliated, hopes to undertake a massive national study of the construction, the content, the use, and the effects of various standardized tests, if it can find funding sources.

Strategies for Action

Countering the excesses and abuses in the use of standardized measures requires two steps. (1) self-instruction in the "basics" of scientific measurement, (2) public exposure of the fallibility of standardized tests and of the injuries their misuse may cause.

The person made uneasy by the language of mathematics and confused by statistical reporting schemes can begin to make sense of the matter by consulting NCTE's brief booklet on *Common Sense and Testing in English*. Here the complexities of standardized testing procedures are explained clearly in non-technical ways, limits and dangers of testing are identified, alternatives are proposed, patterns for reporting test results to the public are suggested.

The second strategy for action—uncovering the fallibility of standardized tests—will take up the rest of this brief report.

Measurement in education is not an exact science. We must never forget that statements about "achievements" in learning made in the language of mathematics, like statements made in everyday speech, are only *approximately* true within *defined limits*. There is no such thing as error-free measurement in physical science or in the human "sciences." Lancelot Hogben points out in *Mathematics for the Millions* that correct measurements cannot be represented by a single number: true measurement involves the statement of two limits, between which the "real" value can be expected to fall. This principle obviously has implications for reporting test results.

Errors may enter into any phase of the standardized testing enterprise. (1) in selecting the sample of language "behaviors" to be measured, (2) in creating and keying test items, (3) in administering the test, (4) in scoring it by hand or machine, (5) in interpreting test results. It is therefore essential to consider all potential sources of error when using any particular standardized

test *before* deciding whether its results are dependable. Once the "human" errors in making and using standardized tests are detected, we still are left with "chance" error, which is invariably present in statistically treated data. A statement of the chance factor, called the "standard error of measurement," should accompany any report of an individual test score. Otherwise, serious misinterpretation of scores will occur. The more responsible companies make clear in their manuals how "obtained" scores will vary through pure chance, but sometimes this information is inaccessible to teachers, being buried in technical manuals in the test administrator's office.

Educational Testing Service does present, in readily available manuals, non-technical explanations of the error factor. On their Scholastic Aptitude Test, for example, if a student's "obtained" score is 600, the odds are 2 to 3 that his or her "true" score will fall between 570 and 630. In other words, if this student were to take this same test again, the chances are 2 out of 3 that the score obtained would fall somewhere within this 60 point range. There is still another error factor to "read" into a comparison of the scores of two persons. The "standard error of the difference" on the SAT indicates that a 72 point difference between two persons' scores on the math section or a 66 point difference on the verbal section is so statistically insignificant that "it cannot be taken seriously." In regard to error of measurement, House, Rivers, and Stufflebeam make the point that even on "highly reliable" tests, individual gain scores "can and do fluctuate wildly for no apparent reason by as much as a full grade-equivalent unit." The implication of this characteristic of standardized test results is clear: *decisions based on test results must take into account measurement error. Scores unaccompanied by measurement error data should be rejected.*

Let us now move one step back to see what lies behind the test score: what goes into a standardized test score? In a standardized reading test, for example, various kinds of test items are bunched together under broad headings such as "vocabulary" or "comprehension." The "raw" score (before it is transformed into grade level equivalents or some other scale) is often the sum of the correct answers, regardless of the level of thinking each item may require. In accumulating the points that make up a reading comprehension score, all responses are counted as equal, whether the item required a high level generalization or merely an act of literal recall. The same score can be achieved by countless combinations of right answers. Thus, a single, uninterpreted score reveals nothing of the strengths or weaknesses that produced the comprehension score.

Let us suppose that we have a mini-comprehension test with only 10 items. Two people each make scores of 7 on the test. The score "7" could be achieved by any of 120 different combinations of answers (according to a mathematician). There is never a guarantee that identical scores arise from the same sequences of performance or that they represent the same pools of knowledge. Yet, identical scores on standardized tests are equated, and educational decisions are made upon the basis of this falsely assumed equivalency. People rarely see any need to ask the question, "*Which* seven questions did the person get correct?"

It is important to remember that test scores are summations. Once the summation is made, important differential information (often the really meaningful, individualizing information) is lost. Even computerized readouts of all right/wrong responses do not provide the truly significant differential information. Even with the item-by-item record of right/wrong responses, we still do not know the "logic" of the error. A recent *Peanuts* cartoon epitomizes the "logic" of the wrong response, of the potential mismatch between the intent of the question-maker and the perception of the answer-giver. In one panel we see Peppermint Patty musing over the teacher's question, "What was the author's purpose in writing the story?" In the next panel comes Patty's answer: "Maybe he needed the money." Deborah Meier's *Reading Failure and the Tests* (New York, N.Y.: Workshop Center for Open Education, 1973) provides a sobering record of what happened when third grade children in PS 144, Manhattan, were asked what particular questions meant to them in the 1970 and 1971 Metropolitan Achievement Tests. The reasons for their "wrong" answers actually make very good sense, exposing levels of comprehension hidden by right/wrong tabulations.

The final topic of this report deals with the major issue in reading the results of any standardized test: How closely does the content of the test fit the program of instruction? Any standardized test of 85 to 150 items is only a sampling of a universe of instruction. So the question of the adequacy of the sample inevitably arises, that is, is this test a representative sample of the behaviors it purports to measure, in terms of the emphases in instruction in a particular classroom or school district? Consider, for example, the typical "Test of Written Expression," where the only things measured are the things that happen *inside* sentences. Often taken by the public and some members of the profession as a true test of composition, it may dwell on the improprieties of "lie/lay," "who/whom," "couldn't hardly," and other mechanical problems. The class that has concentrated on the real tasks of writing (selecting a subject, pursuing a specific intention, addressing a particular audience) may suffer a disadvantage in such a narrow sampling of "written expression." Richard Braddock (in Grommon, 1976) raises an interesting question about these "objective" tests of writing when he asks, "What is the difference between a test of reading ability and a multiple-choice test of 'writing ability'?" A "good reading" of the results of standardized tests in English requires one to determine whether the labels on tests and the content of the items have any significant bearing on the domain they purport to measure.

A person can, of course, perform poorly on a standardized test. But, sometimes, what appears to be poor performance actually is an artifact of the test structure and content, or a misreading of the test results. We need to be able to distinguish one condition from the other. We need then to become better readers of test results and thus wiser consumers of commercially prepared "one-right-answer" standardized tests.

Leo Ruth
(For the SLATE Steering Committee)

Resources

American Psychological Association. *Standards for Educational & Psychological Tests.* Washington, D.C.: American Psychological Association, 1974.

Block, N.J. and Gerald Dworkin, Editors. *The IQ Controversy: Critical Readings.* New York: Pantheon Books (Random House), 1976.

Grommon, Alfred H., Editor. *Reviews of Selected Published Tests in English.* Urbana, Illinois: NCTE, 1976.

Maloney, Henry, Editor. *Accountability and the Teaching of English.* Urbana, Illinois: NCTE, 1972.

Purves, Alan, Editor. *Common Sense and Testing in English.* Urbana, Illinois: NCTE, 1975.

Purves, Alan. "Evaluating Growth in English," in *The Teaching of English, The Seventy-sixth Yearbook of the National Society for the Study of Education, Part I,* James Squire, Editor. Chicago: University of Chicago Press, 1977.

Samuda, Ronald. *Psychological Testing of American Minorities: Issues and Consequences.* New York: Harper and Row, Publishers, 1975.

SLATE Starter Sheets are offered as part of the Newsletter series as resources for dealing with current issues affecting the teaching of English language arts. Reproduce these sheets and use them to help promote better understanding of the goals of English teaching.

SLATE Newsletter is sent to individuals who contribute $5 or more to support the activities of the SLATE Steering Committee. Send your contribution to NCTE/SLATE, 1111 Kenyon Rd., Urbana, Illinois 61801.

Vol. 2, No. 4 May 1977

SLATE

Support for Learning and Teaching of English

PHONICS: MEANS OR MADNESS?

Issues

Phonics can stir strong emotions.

Advocates of a *phonics method* to teach reading often favor traditional education and a return to discipline in the face of so-called progressivism and permissiveness. Most recently, a phonics method has been recommended by educators supporting the "back to the basics" movement. In practice, there is indeed a phonics method, replete with left-to-right "sounding out" of letters and numerous "rules" to cover alleged irregularities between sounds and letters. Old and new materials advocating such a phonics method crowd the market, the majority of beginning reading materials, for example, rely on phonics (R. Aukerman, *Approaches to Beginning Reading*, N.Y.: Wiley, 1971).

The other side of the issue is not quite opposite. Few reading specialists contend that *no* phonics should be taught. They caution, however, against sole reliance on phonics at any stage of learning to read. The issue is between phonics as the only method and phonics as one means among others. To resolve the issue, we must ask. (1) What is known about the writing system? (2) What is known about the learner's interaction with print? (3) What is known about the learner's involvement in achieving literacy?

Professional Viewpoints

What is known about the writing system? Until recently, many reading authorities viewed written English as an attempt to represent sounds directly with letters. Since there is evidence that such representation is far from perfect, some authorities have argued that phonics—the teaching of letter-sound associations—lacks utility.

Bloomfield and Barnhart's *Let's Read* (Detroit: Wayne State University Press, 1961) contended that the match between letters and sounds is more consistent than had been believed. The book brought forth a flood of "linguistic" materials in the 1960s, emphasizing the systematic alphabetic nature of the writing system.

This position seemed to be supported by a computer study done at Stanford. Given a sequence of sounds and sets of "rules," the computer represented those sounds with letters, achieving a fairly high percentage of correct spellings (P. Hanna et al., *Phoneme-Grapheme Correspondences as Cues to Spelling Improvement*, Washington, D.C.: USOE, 1966). On the other hand, a series of studies showed that commonly taught generalizations about sound-letter correspondences lack utility. The English writing system is not adequately described by rules such as "when two vowels go walking, the first one does the talking" (G. and E. Spache, *Reading in the Elementary School*, Boston: Allyn and Bacon, 1977).

While researchers sought to improve such phonics rules, linguist Noam Chomsky and his followers suggested that our spelling system sometimes trades letter-sound consistency for meaning-unit consistency. For example, the past tense marker *ed* is pronounced in a number of ways (*called*, *stated*, *paced*), but its meaning does not vary. The unvarying spelling pattern (*ed*) records the meaning rather than the sound variations.

You need not sit back until the arguments about our writing system are resolved. Examine well-developed materials in practice today and you should find. (1) careful introduction of letter patterns and their respective sound patterns, (2) critical use of phonics rules with emphasis on those whose utility has been verified, (3) the attempt to include meaning units useful to the learner in his or her decoding strategy.

What is known about the learner's interaction with print? Studies comparing children learning to read English with those learning to read a native language possessing higher sound-letter consistency than English have shown that beginning readers of English are slower in learning to decode. Their comprehension, however, is equal to that of their counterparts who speak other languages. At the intermediate stage of learning to read, English-speaking children excel their counterparts in rate and comprehension. At maturity, the groups do not differ in their reading (W. Gillooly, *Reading Research Quarterly* 8 [1973], 167-99).

Hence the beginning reader of English appears to be slowed by a writing system that reflects meanings as well as sounds. Along the way, though, the dual system seems to contribute to the search for meaning through print. Such a finding emphasizes the need to consider long-range results of a reading program, not to base judgment only on results at the end of grades one, two, or three!

Beginning readers can also be helped to attend to basic spelling patterns. That is, they can be helped to discover that certain letter combinations are common in English writing. The term *discover* is important. The best controlled experiments to teach children to decode through spelling patterns, sound-letter associations, or recognition of meaning-units indicate that teaching must be inductive if the new information is to "take." The teacher can aid the discovery, but the discovery itself must be the learner's own (E. J. Gibson, in H. Singer and R. Ruddell, eds, *Theoretical Models and Processes of Reading*, IRA, 1976, pp. 186-216).

SLATE Starter Sheets are offered as part of the Newsletter series as resources for dealing with current issues affecting the teaching of English language arts. Reproduce these sheets and use them to help promote better understanding of the goals of English teaching.

What is known about the learner's involvement in achieving literacy? "Culture conflict" between learner and teacher may lead to rejection of reading. If teachers or materials insist on phonemic distinctions that do not exist in the reader's dialect, the implication is that the dialect is unacceptable. (For example, *i* and *e* in *pin* and *pen* do not represent contrasting vowel sounds in all dialects.) Rigid "correction" of dialect-based miscues seems unwarranted in view of research findings that these sound-based miscues often do not change meaning or reduce comprehension (C. Burke and K. Goodman, *Elementary English* 47 [1970], 121-29). If meaning-units and context clues are balanced with phonics to reflect an acceptance of dialects, the odds are that the divergent-speaker will respond positively to reading instruction.

The child who is taught phonics to the exclusion of other methods is likely to employ a phonics strategy at the expense of other cue systems (R. Barr, *Reading Research Quarterly* 10 [1974-75], 555-82). Teachers' phonics methods beget learners' phonics strategies. By the same token, the teacher whose method balances phonics with other decoding strategies equips the learner with a repertoire of strategies to help in achieving literacy.

Examining the learner's involvement in achieving literacy, J. Johns and D. Ellis found that many American children perceive reading as a "sounding-out" process related only to school activities (*Reading World* 16 [1976], 115-27). Few children view reading as a meaning-deriving process related to daily life. The child who views reading as meaningless "sounding-out" has little reason to want to read. In contrast, the child who perceives reading as silent communication, a means for discovery and delight, will seek opportunities to read. To develop this perception of reading, the teacher needs to draw extensively upon meaning-based techniques which help place phonics and other skills in context as means to reading for understanding, not as ends in themselves.

Strategies for Action

☐ Evaluate materials for inclusion of a variety of strategies to be used in decoding. Present reasons to colleagues and parents for the need to incorporate these strategies in materials and methods.

Viewpoints cited above indicate, as one author suggested, "phonics in proper perspective." (See Heilman below.) Patterns of sound-letter relationships need to be taught by helping the reader discover such relationships, using those generalizations and spelling patterns whose usefulness has been demonstrated. In addition, the consistency in spelling patterns at the meaning-unit level of the writing system justifies concurrent teaching of such units. Rather than asking *how much* phonics is taught in a reading program, one might better ask whether recent information about the writing system and the learner has been utilized in designing the program. (See Burmeister below.)

Longitudinal research has been neglected in reading. In its absence, schools must take responsibility for evaluating their programs for long-range effects. One may ask whether a primary reading program produces results at the end of the primary years. A more important question, however, is whether the program contributes to a balance of decoding skills, comprehension, and positive attitude toward reading at *all* levels, including those beyond the primary grades.

☐ For help with divergent dialects, adapt pronunciations and emphasize a variety of language experiences in conjunction with phonics.

Phonics instruction must be flexible particularly when dialect differences are involved. If it is not, the learner may reject reading through "culture conflict," the impression that his or her dialect is wrong. It follows that reading teachers must be informed about dialect variations, in order to modify the phonics program appropriately. (See Zintz below.)

☐ Read and discuss research that indicates that reading is much more than sequential "letter-sounding." (See Kolers below.) Discuss definitions of reading with students, colleagues, and parents. Elicit students' aid in formulating purposes for reading and encourage students to devise reading activities in keeping with the goals of the reading program.

There is little logic in the argument that a "code emphasis" program precludes meaning. There *is* logic in the premise that beginning readers, like all readers, need to know why reading is important and what is to be gained from it. Without this understanding, decoding programs replete with phonics will still fall short of the goals of literacy.

Sam Leaton Sebesta
Mary Ellen Pfrimmer
(For the SLATE Steering Committee)

Resources

Artley, A.S. "Phonics Revisited," *Language Arts* 54 (1977): 121-126.

Burmeister, L.E. *Words—From Print to Meaning.* Reading, Massachusetts: Addision, Wesley, 1975.

Heilman, A. *Phonics in Proper Perspective.* Columbus, Ohio. Charles E. Merrill, 1976.

Kolers, P. "Three Stages of Reading," in F. Smith, ed., *Psycholinguistics and Reading.* New York. Holt, Rinehart and Winston, 1973.

Zintz, M.W *The Reading Process: The Teacher and the Learner,* 2nd ed., Dubuque, Iowa. William C. Brown, 1975. Chapters 10 and 16.

SLATE Newsletter is sent to individuals who contribute $5 or more to support the activities of the SLATE Steering Committee. Send your contribution to NCTE/SLATE, 1111 Kenyon Rd., Urbana, Illinois 61801.

Volume I of the SLATE Starter Sheet/Newsletter series (emphasis: back-to-the-basics) is now available from the NCTE order department. Stock number 98619-76, $1.00 per set. Prepaid orders only, please.

Vol. 2, No. 5 August 1977

SLATE

Support for Learning and Teaching of English

COMPETENCY TESTING AND BILINGUAL/BICULTURAL STUDENTS

The Issues

Bilingual/bicultural populations are usually both minority populations and economically disadvantaged populations. In the United States, there has been a long history of assimilation of minorities, probably as an outgrowth of the "melting pot" metaphor. Only in relatively recent times has the educational system (and indeed the population in general) begun to recognize the essential values of cultural diversity.

The history of assimilation practices is well documented with respect to such populations as Native Americans, various Spanish-speaking groups, Blacks, and so-called "French-Creole" populations. Unfortunately, it is less well-documented for significant numbers of other populations from South America, Europe, Asia, Africa, and the Middle East. Any or all of these populations may be present in a large metropolis like New York or Los Angeles, but they are also present in many smaller cities in all parts of the United States.

As bilingual/bicultural legislation (and resultant educational policy) has evolved, it has applied generally to the largest and most visible populations. In part this is a function of the fact that, traditionally, these populations have tended to cluster geographically, however, in more recent times, improved transportation facilities and changing patterns of migration have begun to alter the population distribution.

Only since the Supreme Court handed down its decision in the Lau case has there been a significant national effort to provide bilingual/bicultural education for all populations entitled to such education. In his majority opinion, Justice William O. Douglas wrote:

> Where inability to speak and understand the English language excludes national origin-minority group children from effective participation in the educational program offered by a school district, the district must take steps to rectify the language deficiency in order to open its instructional program to these students. [*The Linguistic Reporter*, 16:3 (March 1974), pp. 6-7.]

Now, under the provisions of the so-called "Lau Guidelines" issued by the Department of Health, Education, and Welfare, bilingual/bicultural programs are mandated under the provisions of Title VI of the 1964 Civil Rights Act, as amended. (The original guidelines are reprinted in *The Linguistic Reporter* 18 [October 1975], pp. 1, 5-7.)

Professional Viewpoints

Despite popular conceptions of bilingual/bicultural programs based on the largest and most visible populations, there are in fact at least quite different groups involved. One group consists of individuals who are legal residents of the United States, while another group consists of individuals who reside in the United States on a temporary basis (diplomats, foreign students, United Nations employees, and their children). Although members of these groups share common difficulties in the use of the English language and would benefit from similar programs, only legal residents are entitled to participate in federally funded programs. As a result, some school and community college districts have had to establish parallel programs which compete for limited instructional resources.

There is, potentially, a third bilingual/bicultural population consisting of speakers of Black English at that point where it differs most from Standard English. While this population is in many ways quite different from the other two groups, teaching strategies developed for other bilingual/bicultural groups would still be useful. This group, however, is also excluded from participation in programs funded for the other populations.

NCTE has spoken to the issues in its 1974 resolution on the students' right to speak and learn their own languages and in the subsequent publication *Students' Right to Their Own Language*, as well as in its endorsement of a position paper developed by the Interassociational Group on the Social and Political Concerns of Minority Groups. (The position paper is available from SLATE, 1111 Kenyon Road, Urbana, IL 61801.) The "rights of all peoples" recognized in this position paper include:

- the right to maintain their own patterns and varieties of culture, language, and dialect; and
- the right to expect public educational institutions to recognize the validity of cultural and linguistic diversity.

While the Department of Health, Education, and Welfare has developed certain criteria for bilingual/bicultural programs, there is still confusion about (1) testing for English (or other) language competence and (2) testing for subject competence through the English language.

A variety of tests have been developed to test language competence. The best known of these is the ETS Test of English as a Foreign Language (TOEFL), but even it suffers from some problems:

- it is not easily available except overseas
- it costs money
- its security has been breached
- it does not require active language production
- it cannot predict rate of learning

SLATE Starter Sheets are offered as part of the Newsletter series as resources for dealing with current issues affecting the teaching of English language arts. Reproduce these sheets and use them to help promote better understanding of the goals of English teaching.

Despite its problems, it is the most valid and reliable (in the technical sense) test available. Other available tests include the *Michigan Test*, the *American Language Institute Georgetown University Test* (ALIGU), the *CELT Test* (McGraw-Hill), the *Ilyin Oral Interview Test* (Newbury House), and a number of older tests. (See O. Buros. *Mental Measurements Yearbook*. New Jersey: Gryphon Press, various years, and SLATE 2:2 [December 1976].) Basically, these are all designed for easy scoring and, except for the Ilyin test, require no active production of connected language, either oral or written. Each of the tests was originally designed for a particular educational level (for example, the Ilyin test was intended for use in adult education, the CELT for use in high school), but because of the lack of good alternatives all of them are often used indiscriminately for a variety of levels. This is an unfortunate practice—almost as unfortunate as the practice of testing bilingual/bicultural students with tests designed for native-speakers of standard English.

The discussion so far has focused on measures of language proficiency. Similar problems relate to competency measurement in specific subject areas. Since subject-field tests do not generally exist in a variety of languages, standardized tests written in English are used instead. This not only ignores important linguistic differences but also ignores the problems resulting from contrasting instructional emphases (e.g., the mechanics and jargon of arithmetic are not universal, and the concepts underlying history and political science ["Civics"] vary dramatically). Attempts to use standardized tests discriminate against linguistically different populations in several ways:

- —students may not understand the question itself, because of difficulty with the grammar or the meanings of words;
- —students may not understand the presuppositions of the question, because of cultural differences in outlook or experience;
- —students may not understand the processes implicit in the question, because of differences in the way in which the subject field has been approached or conceptualized.

Strategies for Action

- ☐ Find out how the Department of Health, Education and Welfare "Lau Guidelines" apply to your state and your school district.
- ☐ Examine the linguistic and cultural diversity in your school and community. (The Lau Guidelines require a periodic census of non-English-speaking and limited-English-speaking students.)
- ☐ Lobby in your state for assurances of adequate criteria for the training and certification of teachers in bilingual/bicultural and English-as-a-Second-Language programs.
- ☐ Work toward the adoption of reasonable measurement instruments. (The issue of competency testing is mentioned but not resolved in the Lau Guidelines.)

Robert B. Kaplan
(For the SLATE Steering Committee,
with the cooperation of the Interassociational Group
on the Social and Political Concerns
of Minority Groups)

Resources

Brickman, W.W. and S. Lehrer. *Education and the Many Faces of the Disadvantaged*. New York. John Wiley, 1972.

Butler, Melvin A., chair. *Students' Right to Their Own Language*. Fall 1974 special issue of *College Composition and Communication*. Urbana, Illinois. National Council of Teachers of English.

Gibson, M.A., "Approaches to Multicultural Education in the United States," *Anthropology and Education Quarterly* 7.4 (November 1976), pp. 7-18. The entire issue is devoted to Multicultural Education.

National Association of Bilingual Educators, 500 South Dwyer Street, Arlington Heights, Illinois.

National Bilingual Resource Center, Center for Applied Linguistics, 1611 North Kent Street, Arlington, Virginia 22209.

National Center for Education Statistics, 400 Maryland Ave., S.W., Washington, D.C. 20202.

Navajo Bilingual Education Center, c/o The Navajo Nation, Window Rock, Arizona 86515.

Office of Bilingual Education, Department of Health, Education, and Welfare, 7th and D Streets, S.W., Washington, D.C. 20202.

Teachers of English to Speakers of Other Languages (TESOL), 455 Nevils Building, Georgetown University, Washington, D.C. 20057.

vonMaltitz, F.W. *Living and Learning in Two Languages*. New York: McGraw-Hill, 1975.

Williams, F. (ed.) *Language and Poverty*. Chicago: Markham, 1970.

SLATE Newsletter is sent to individuals who contribute $5 or more to support the activities of the NCTE/SLATE Steering Committee on Social and Political Concerns. Send your contribution to NCTE/SLATE, 1111 Kenyon Road, Urbana, Illinois 61801.

Volume I of the SLATE Starter Sheet/Newsletter series (emphasis: back-to-the-basics) is now available from the NCTE order department. Stock number 98619-76; $1.00 per set. Prepaid orders only, please.

Vol. 2, No. 6 October 1977

SLATE

Support for Learning and Teaching of English

THE NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS

The Issues

The National Assessment of Educational Progress (NAEP) is federally funded through the National Center for Education Statistics and is administered through the Education Commission of the States (ECS), which in turn is a compact of the several states to deal with common educational concerns. The assessment is essentially a representative census of the educational achievement of Americans aged 9, 13, and 17. In the first cycle of the assessment young adults (26-36 years old) were included, but practical problems have led to the elimination of this group. Assessments are made periodically in ten areas, three of which comprise the Language Arts Curriculum—Reading, Literature, and Writing.

The aims of NAEP are to provide a reliable and comprehensive profile of students' knowledge and abilities and to record changes in knowledge and ability as reported in the different test cycles. These are research activities, and NAEP is explicitly prohibited from collecting data in a way that would allow individuals or even school districts to be identified. Results are reported by region, sex, color, educational background of the parents, and type of community. Individual states have used NAEP materials (which are in the public domain) for testing programs which can be used to judge districts or individuals, but the NAEP procedures are designed to provide information about groups of students, not individuals. Usually the term "assessment" is used instead of "testing" to emphasize the general scope of the information to be gathered.

In developing assessments NAEP tries to be sensitive to the concerns of the lay public and of teachers in the subject matter areas as well as to the technical requirements of educational statisticians and exercise designers. Test objectives are developed in consultation with lay groups, reviewed by teachers, and ultimately approved by lay directors of ECS. These are not merely pro forma efforts. The reports on writing mechanics, for example, are in part a response to requests from the lay public. Objectives change as additional ideas are forthcoming. The second cycle of testing in writing, for example, placed additional emphasis on the ability of individuals to formulate feelings as well as business, social, and academic transactions, and objectives tested in the assessments of reading and literature have also been amended.

Once the objectives are approved, the construction of exercises is contracted to consultants and testing agencies. These exercises are field tested and analyzed, revised as necessary, and retested. When they are ready, exercise packets are administered under contract by another external agency, which also constructs the test sample. The resulting papers are scored by yet another group, the results are compiled for interpretation by specialists in the subject matter areas.

Although NAEP and its consultants determine the research questions to be asked and the format of the answers obtained, strictly speaking NAEP does not interpret the results. That is a major issue. On one hand the Government does not wish to seem to be establishing a national watchdog, on the other hand, defining the problems to be studied and the means of studying them also determines much of what will be revealed. Choosing the system of stylistic description, for example, may determine the apparent sophistication of the language user by calling attention to one kind of language feature and ignoring another. Still, NAEP has tried to bring in people from the discipline to examine the results and report judgments under their own names.

Tabulated results and, sometimes, professional opinions are released in NAEP technical reports. The data are partial. Approximately half of the material is withheld in order to provide papers for exact comparison with materials produced in the next cycle. The great expense involved in scoring and reporting complex data is another reason that much raw material is held for later review. NAEP hopes that outside researchers with outside funds will be willing to make additional studies.

The limited and tentative nature of the reports is often belied by newspaper reports. NAEP is legally bound to release materials without its own commentary and depends for funds upon the public sense that the activity is worthwhile. Popular accounts tend to drop out the qualifications and hesitations of researchers; what the public hears is more definite than the results usually deserve. Careful research may be significant without being electrifying. Still, uninterpreted data encourages some people to try to be electrifying, and the lay public is generally not able to read through the viewing-with-alarm in order to find the underlying fact.

Professional Viewpoints

Reading

The reading assessment was administered in 1970-71 to reveal how well students (1) comprehended what they read, (2) analyzed what they read, (3) utilized what they read, (4) reasoned logically from what they read, and (5) made judgments from what they read. Strictly only the first aim deals with reading as such; the others deal with the use of materials gathered in reading and thus provide inferential evidence about comprehension.

SLATE Starter Sheets are offered as part of the Newsletter series as resources for dealing with current issues affecting the teaching of English language arts. Reproduce these sheets and use them to help promote better understanding of the goals of English teaching.

The results of the first cycle of the reading assessment have been published in great detail but should be read with caution. At best they provide baseline data, in no category do we know what someone *ought* to be able to do. The second assessment in reading was administered in 1974-75, and results are just beginning to become available. Comparisons between the first and second cycles have been made in the general areas of literal comprehension, inferential comprehension, and reference skills. The most encouraging finding is that nine-year-olds during the second assessment read significantly better than did nine-year-olds four years earlier. The improvement was noted in all reading skills, but was greatest for reference skills. Black nine-year-olds showed the most dramatic improvement, a finding which a panel of reading specialists convened by NAEP attributed to intervention programs that have been implemented at the primary level since the first assessment.

For the other two age groups tested, results were less consistent. Both groups showed a slight improvement in literal comprehension and a slight decrease in inferential comprehension. Students at all ages had little difficulty comprehending straightforward, literal material but their comprehension dropped off quickly as the reading tasks became more difficult.

Preparations for the third cycle of the reading assessment are already underway, although some delays have resulted from a plan to combine the assessments in reading and literature. The results should provide further information about trends in reading skills, although the reading assessment has been complicated by conflicting theories about the nature of the reading process, which in turn leads to conflicting views about the most productive techniques of assessment.

Literature

The literature assessment, also administered in 1970-71, was designed by the Educational Testing Service to reveal whether students (1) read literature of excellence, (2) became engaged in, discovered meanings in, and evaluated works of literature, and (3) developed a continuing interest and participation in literature and literary experience.

The results of the assessment were reported in four convenient but not exclusive categories, each represented in a separate technical report. These are (1) understanding imaginative literature (for example, sensitivity to mood and tone or metaphor), (2) responding to literature, (3) recognizing literary works and characters, and (4) reading habits.

The objectives of the second cycle were drastically recast to put more emphasis upon literary language—the ability to use it and respond to it—as opposed to knowledge about literature. These new standards are far more detailed and probe into areas which seem to be difficult to assess even though they may be closer to what are generally cited as reasons for studying literature. They could be a good framework for a course in literary criticism.

The new literature objectives like the new objectives for the assessment in reading suggest a greater concern for mental process and thus the two assessments seem to be naturally complementary in ways not much stressed in the first cycle. Some conventional concepts of instruction in reading fall short of any concern with literary language despite its pervasiveness even in ordinary social relationships, and some views of literature are so concerned with literary works that they ignore the problems of processing language. The combined tests have not yet been developed, but the danger in the combination is that the difficulty of evaluating skills in processing literary language may tempt the assessors to avoid it and thus seem to say that literature is unimportant.

The results of the first cycle of the literature assessment are necessary to establish a baseline for comparisons with later cycles and thus must not be over-read, but often the specific information gathered in the assessment can be readily examined in terms of the classroom. A list of titles and characters recognized at different age levels implicitly describes what the schools are teaching and suggests other works which might be included. A survey of attitudes can be directly interpreted. Observations about skills in reading literature are less satisfactory, partly because the tests are less adequate and partly because of the lack of an external benchmark to use in interpreting the data.

Writing

The most provocative assessment is that of writing. Two cycles have been completed—1969 and 1974—and the third is well along in preparation. Not only were the objectives refined after the first cycle, but methods of assessment more in accord with the function of NAEP were developed in response to objections that existing methods of scoring papers merely showed that some papers were better than others. And, of course, the papers themselves are very important data which can be re-examined as new methods are developed.

The aims of the first assessment were to communicate adequately in social, business, and academic situations and to appreciate the value of writing. Some of the exercises required little more than simple reportage or opinion, but the most useful tasks required substantial writing. The objectives of the second and third assessment include an emphasis upon writing to express personal feelings and ideas, so the new assessment is probably more closely representative of skills of composition as they affect most people in their daily lives.

The results of the writing assessments have been reported in several ways. Perhaps the most useful is Report 10 in the first cycle, for it contains a computer printout of all of the responses to one exercise at each age level. This is raw data, but it allows the reader to experience what the respondents have created and thus make more effective judgments about papers written for the classroom. Since much of the public comment about NAEP scores comes from people who have never actually seen a large number of responses, resultant battles often seem not to refer to any substantial reality. This report can be a useful corrective. Its existence can make one wish that more exercises could be called off tape, for the computer provides the possibility of many kinds of detailed analyses.

The reports on the general quality of the writing are tantalizing, but frustrating. The general report on the first cycle is based on holistic scoring of essays and fairly mechanical reports on the value assigned to writing and on skill in supplying information suitable to occasions. Results on supplying information are tainted by the degree to which respondents may simply not have felt the imperatives of real need. For example, a person who would have supplied a return address for a real letter may not have bothered, in an exercise, to give fictitious names and addresses. The holistic scores merely rank papers, and the examples of papers of different quality suggest more than they actually show.

The second cycle report on expressive writing is much more informative because it contains detailed accounts of what the scores represent, and it deals with a limited set of writing situations. Those who want to disagree with the report or reinterpret it have more specific observations to work with. The frustration lies in knowing that the focus is achieved by putting aside large amounts of material for later study. As it stands this report makes an excellent complement to Report 10 in Cycle 1—the anthology of responses—for (like a piece of literary

criticism) it helps the reader examine the material, but the results are not readily captured by a statistician.

The other kind of report on writing deals with writing mechanics. This subject is especially important to the lay public, and generally the material can be quantified. Still, many readers will differ about the actual definitions of "errors", and modern linguists probably will want to use different systems of description to report problems. Nonetheless, one is probably justified in saying that the comparisons show little change in the ability to handle manuscript problems. Error levels seem to be constant among the age groups and over the five year period even though the particular errors change; this suggests that most people develop a sense of error tolerance which prods them to seek alternative forms of expression when the number of errors becomes large enough to seem improper.

In conjunction with the study of mechanics some loss in the conventional markers of coherence was observed—transitions, topic sentences, and the like. This was not a problem for the best writers in the 17-year-old group—the ones apparently bound for college. Thirteen year olds showed the greatest loss in this respect between the two assessments. The nine-year-olds seemed better in all aspects of writing skill, as judged on an exercise which is basically expressive.

Speculations about the causes of the changes are inconclusive, although probably one should observe that skills of language usage are developed outside the school as much as in. For some subjects, almost all of the students' knowledge is obtained in school and results can be related to changes in the schools. For writing, one probably should look more to the society in general. Possibly recent changes in social conditions are also reflected in the social aspirations of young people—especially those in the middle range of competency—and thus the desire to produce writing for an anonymous test may be attentuated. For nine-year-olds the task may have seemed like fun and television may have enlarged the experiential base from which writing is generated. It is also possible that the larger place of the electronic media in the lives of the young may be subtly changing their sense of what is coherent, leading them from the abstract categorical and hierarchical patterns of our print oriented tradition to the more visual and associative pattern of stop-action films.

Strategies for Action

NAEP is essentially a research organization in a political atmosphere. Responsiveness to the society is a virtue so long as it is protected from demagoguery. Some are so fearful of the demagogues that they want to resist any discovery which might support a change in present practice. Others are eager to take tentative studies as definitive. Our challenge is to find a balance, and that may require that we know enough about the work going on that we can add complications to the demagogue's simple assertions.

NAEP is developing a bank of raw material which has been hardly examined. Some materials need to be put into a form which is economically available to researchers at a distance. Specifically, many would be helped by having actual essays; others would be helped by having more convenient access to technical studies which do not appear as reports but can be obtained by those who happen to have heard of the work. Although the total NAEP budget is finite and precludes some scholarly efforts, still the Assessment is sensitive to active interest by researchers. It has a stake in getting its data examined and reported on. Those with research grants from other sources might discuss with NAEP officials ways of using NAEP data.

Resources

Anyone just beginning to study NAEP should use John C. Mellon's report on the first assessment, *National Assessment and the Teaching of English*, published by NCTE in 1975. NAEP reports are obtained from the Superintendent of Documents in Washington, D.C., 20402, and current lists of publications may be obtained from NAEP, 1660 Lincoln Avenue, Denver, Colorado 80203.

Richard Lloyd-Jones,
with bibliographic help from
Fred C. White
(For the NCTE/SLATE Steering Committee
on Social and Political Concerns)

Vol. 2, No. 7 October 1977

SLATE

Support for Learning and Teaching of English

THE MINIMUM COMPETENCY MOVEMENT: STRATEGIES FOR RESPONSE

What Skills Are Tested?

Minimum competency testing is nationwide. The April 1977 *Update* from the Education Commission of the States listed eight states with enacted legislation, nine state boards with adopted rulings, and thirteen other states with legislation pending. Because testing is often the tail that wags the dog of course content, English teachers need a strategy to insure that the worst effects of minimum competency testing are avoided.

But first what are some of the issues? A critical concern is what should or should not be mandated as the subject areas to be tested. Reading, writing, and mathematics are mandated by Arizona, California, Nebraska, Ohio, Oregon, Virginia, and Washington (Virginia and Washington substitute communication and language arts for writing skills). Minimum writing skills are not mandated by New Jersey, New York (though it uses the label), Maryland, or Texas. Florida and New Mexico refer only to basic skills, not mandating any specific area. Massachusetts does not mention reading or writing specifically, but it does call for skills in communications and five other areas—computation, career knowledge, social responsibility, environment, and culture.

One issue is whether a subject area will receive adequate institutional support if it is not a mandated area of competency. For example, will writing programs suffer in New Jersey, New York, Maryland, and Texas because they do not list writing as an area requiring minimum competency? Given the fact that one of the purposes of the minimum competency movement is to redefine the goals of the schools and cut the frills, the answer could by "yes." Many critics of the schools believe that over the last two decades the schools have assumed too many responsibilities - everything from solving drug abuse among teenagers to teaching everyone to drive and that schools must now invest all of their resources in only necessary, educational programs.

Too Much Testing?

On the other hand, if states mandate too many areas of minimum competency, testing time may seriously interfere with instruction. For example, in Massachusetts, with six areas, the students and teachers could find themselves spending twelve days per year on competency testing and record keeping. And even more testing days are possible in Georgia, where the State Board is considering mandating minimum proficiency standards for five "life roles"—student as learner, individual, citizen, consumer, and producer. As the list grows, the amount of time consumed by testing increases.

What Type of Tests?

The forms of tests of minimum competency are a central issue. How an area is measured and assessed may seriously influence how an area is taught. For example, if tests of writing competence contain only multiple choice items on grammar, then writing classes may not put great emphasis on actual writing assignments. Agencies responsible for developing tests of minimum competency in writing need to be familiar with the work of Paul Diederich (*Measuring Growth in English*, NCTE, 1974) and others. The Steering Committee of SLATE has responded to this problem by preparing a proposal to design a writing assessment package which uses holistic scoring and meets the needs of policy makers.

The debate over the form of minimum competency tests has centered largely on norm versus criterion references. Criterion-referenced tests appear to be on the ascendency in competency testing in Florida, Georgia, Alaska, Michigan, Ohio, Maryland, and California. States adopting a test for students who wish to test out of school early almost always use criterion references, usually matters identified as functional or survival skills. As far as the content goes, items on the criterion-referenced tests seem to be neither better nor worse than those on normed tests. The issue is whether the items on the test constitute an adequate illustration of what all students at a given grade level should be able to do.

What might be explored in the testing sections of minimum competency legislation are alternative forms of assessment. For example, suppose students graduated from (or entered) high school after successfully passing not a test but an interview with a screening committee. This committee, basing its decision on some test data, letters of recommendation, transcripts of course work, a portfolio of samples of the student's work during the previous three or four years, and responses to inquiries at the interview—could add to the assessment some personal knowledge of the student beyond the numbers and letters on competency records. Then some dimension of the student's individual growth and development would become part of the criteria for meeting competency requirements.

The question is whether the state or the profession believes that such a procedure is worth the time and money required to implement it. The total cost for graduation and early exit screening would not be small. The National Center for Education Statistics reports that twenty-two states have adopted tests for high school graduation, and twenty have adopted tests for early exit from high school (Jeanette Goor, *Statewide Development in Performance Based Education*, Report No. 1, National Center for Education Statistics, 1976).

How to Influence Competency Legislation

English teachers' ideas about what and how to test can become a part of minimum competency legislation through a strategy of influencing (1) the advisory committees at the state and local level and (2) elected officials.

Advisory committees. Most state legislation mandating a major

change in education is preceded by a task force or advisory committee to study the issues and make recommendations. Alabama, New Mexico, New Jersey, Connecticut, Maine, Massachusetts, Michigan, and Nevada all have either state task forces or advisory committees to develop a plan for minimum competency testing or requirements for graduation. Because the recommendations from these committees will shape the major legislation on minimum competency, the state English council in each area should seek to have one of its members appointed to the advisory committee.

To secure an appointment, the state council can write to the agency or official making the appointments and request the appointment of a member of the English council. If the English council has some influence with an elected state official, ask that official to request the appointment. In either case, the SLATE liaison can initiate the action by asking the president of the state English council to write a letter of request.

The state advisory committees often need representation from English teachers in order to legitimize what the state is doing. Sometimes the state English council can use this need as a trade-off for something the profession wants. Such a trade-off happened in California when a group of English teachers threatened to resign from a state advisory committee on testing in the language arts unless the state guaranteed that it would collect samples of student writing as part of its state assessment. As a result, since 1975 California has annually collected 5,000 samples of student writing and funded an holistic scoring of these papers. In exchange the English teachers agreed to have their names listed as advisors on the machine-scored tests in reading and language.

English teachers in California and Oregon should also be involved in the local advisory committees which have the responsibility for recommending minimum competency standards and graduation requirements in these states. English councils should urge representation not only from English teachers but also from other groups in the community. If English teachers alone determine tests or standards which deny graduation or diplomas to some students, those students and their parents may very well believe that a single professional class has denied them jobs and entry to the mainstream of economic life. The responsibility for the tests and standards must be generally shared by the community, not just by English teachers.

Elected officials. State English councils may wish to change existing or pending legislation. Pending legislation is probably the most productive focus because legislators usually want to give new laws a chance to work for one or two years before they change them. Pending legislation, on the other hand, is often tomorrow's law. Most major pieces of legislation in education *began* as bills which died in committee, never going to a floor vote. Such bills should be watched carefully, because a large number are later revived.

An interesting example of a bill which died in committee but which has much potential for mischief is North Dakota's House Bill 1429. This bill called for state testing to measure minimum competency standards in communication skills, which were to include reading, phonics, syllabication, grammar, and spelling. Such legislation obviously encourages separate units, perhaps even separate courses, in reading, phonics, and syllabication, and many teachers, knowing that the state will test for phonics and syllabication as ends in themselves, would feel obliged to give both matters major attention even in classes where all the students were excellent readers. (Do *you* remember the rule for the silent e?)

The way to change this bill or others—say, legislation which omits writing—is to contact the author of the bill and ask for consideration of a specific change. The request should be made in a letter from the state English council with follow-up letters from sh teachers and parents in the legislator's district. If the author of the bill is not willing to make the change, then the state council should write to the chair of the education committee in the state legislature, requesting information on the time and place of the committee hearing on the bill and enclosing a copy of the council's position on the bill.

Whether or not the author of the bill accepts the council's changes, a representative of the state English council should appear before the committee hearing on the bill and present the position of the state's English teachers. The author of the bill, as a matter of courtesy, should always be informed of the English council's position *before* a representative appears at a hearing to testify. Up until the time the bill is finally killed or signed, the state English council should continue to inform legislators of its position through letters and phone calls. This effort is especially important if the state council is supporting the bill because the author made extensive changes. Legislators do not feel very kindly toward those who get what they want and then disappear.

Rules and Regulations

Whatever is won in the law can be lost if the English council is not attentive to the rules and regulations that are adopted to implement the law. In most states these rules and regulations are adopted by state boards. Sometimes local boards are also asked to adopt implementing rules. The experience of the California Association of Teachers of English in PPBS (Planning-Programming-Budgeting-System) illustrates the importance of regulatory agencies. After the legislature defeated an elaborate PPBS bill and passed a rather innocent two-year pilot program, the State Board of Education (after one year of lobbying from the advisory committee for the pilot program) almost decided to step into the statutory vacuum and, using its authority over the school accounting manual, adopt PPBS statewide. The opposition of the California Association of Teachers of English in testimony before the State Board and before the advisory committee played an important role in the defeat of PPBS.

One of the reasons for the success of this effort was the cooperative alliances between the California Association of Teachers of English and other teacher organizations such as AFT and NEA. To expand these cooperative alliances, the California association at one point joined in an effort to form a coordinating council of all curriculum organizations in the state. Because most curriculum organizations have a mutual interest in issues such as PPBS and testing, and because the key to effective lobbying is the comprehensive alliance, such a coordinating council, even if it meets only once each year, should be a major goal of state English councils.

Students Who Fail

The last step which state councils might consider is adoption of a policy statement or position paper on the proper standards for remediation. Many states—California, Virginia, Colorado, and Florida, for example—require that remedial or intervention courses be offered so that students who do not attain the minimum competencies or who are in danger of not graduating from high school will receive the assistance necessary to attain minimum levels. If the profession does not begin now to define the needs for remedial instruction, these remedial or intervention courses may only condemn many students to repeated failure. The legislators must turn to the profession for a definition of these standards, and it is the obligation of the profession to provide it.

Miles Myers (For the NCTE/SLATE Steering Committee on Social and Political Concerns)

Vol. 2, No. 8 November 1977

SLATE

Support for Learning and Teaching of English

THE SAT SCORE DECLINE REPORT

Background

In October 1975, the president of the College Board commissioned a blue-ribbon panel to examine the fourteen year drop in average scores on the Scholastic Aptitude Test (SAT). Their report, *On Further Examination*, was issued in August 1977, bolstered by an array of special studies commissioned by the panel. It is a readable, carefully qualified document that contributes substantially to our understanding of the nature and causes of the score decline.

But the issues involved in the report are of national interest and interpretations by the media have not always reflected the balance in the report itself. English teachers may find themselves confronted with "facts" and conclusions not justified by the report; they may also find that other parts of the report will provide good support for their efforts. This issue of *SLATE* outlines the findings and main areas of controversy for teachers who have not yet had a chance to study the report.

The SAT: What Is It?

The Scholastic Aptitude Test is a two-and-a-half hour, multiple-choice examination with separately scored Math and Verbal sections. Each section is scored on a scale that ranges from 200 to 800, constructed so that the standard will remain unchanged from year to year. Theoretically, a score of 507 on a test taken in 1977 is equivalent to a score of 507 on a test taken in 1941-1942 (when norms for the SAT were established). Special studies undertaken for the score decline panel suggest that this theoretical expectancy has been reasonably well-fulfilled; if anything, the test may be slightly easier than it once was.

The Verbal section of the test has four types of items: antonyms, analogies, sentence completion, and reading comprehension. The content of the items is *not* limited to English; materials are drawn from social, political, scientific, artistic, philosophical, and literary writing.

Some 1.4 million students took the SAT in 1976-77.

The purpose of the SAT is straight-forward and limited: it is a predictive measure of academic performance in college, particularly during the first year. The "audience" for test scores are college admissions officers. Scores on the SAT do not predict college success as well as do high school grades but, when combined with grades, provide a slightly better overall predictive index than does either alone. (In 1974, the median validity coefficient for the SAT-Verbal was .42; for the SAT-Math, .39.)

After reviewing studies on the predictive ability of the test and rejecting possible sources of bias, the panel concluded that the test accomplishes what it sets out to do, but recommended to the College Board that another panel be set up to investigate the whole matter of the function of tests in the passage from high school to college. This new panel would be concerned with examining the relevance of standards adopted in the past to the educational opportunities of the present—an issue which the score decline panel regarded as outside its scope of inquiry.

The Score Decline

The panel decided, on good evidence, that the reported score decline was real; it could not be explained away as the result of changes in test items or scoring procedures. The decline has affected both Math and Verbal scores, and has continued for approximately fourteen years. Overall, the SAT-Verbal dropped 49 points between 1963 and 1977; the SAT-Math dropped 32 points.

During the last two years the decline has slowed down, and for some populations has begun to reverse itself.

Causes of the Decline

The panels' analyses eventually led them to conclude that there have been two successive score declines with different underlying causes. The initial decline, which began about 1963 and tapered off between 1970 and 1972, was due largely to changes in the composition of the student populations taking the test. Each year this population included larger proportions of groups which traditionally have had lower scores, and which pulled the overall average down: largely as a result of a national attempt to expand and extend educational opportunities to groups previously discriminated against, increasing numbers of students of lower socioeconomic status, of women, of members of minority groups, and of students intending to go to two-year or open-admissions institutions began to take the test. The panel attributes two-thirds to three-quarters of the decline between 1963 and 1970 to such shifts.

The panel points out, however, that while this shift in the composition of the test-taking population is responsible for much of the score decline, this explanation itself has serious implications:

> What the decline reflects is the incompleteness so far of the national undertaking to afford meaningful equality of educational opportunity. This leaves the question of whether a 75 percent cross section of all young people *can* be brought up to the 11th- or 12th-grade academic attainment level previously achieved by 50 percent of them. Part of democracy's sustaining notion is that they can be. (p. 45)

After about 1970, changes in the composition of the SAT population account for no more than a quarter of the recorded decline. Instead the panel finds evidence of a "pervasive" decline in the average scores of virtually all groups, including the highest scoring. This decline is unrelated to socioeconomic level, sex, minority group membership, course of study, high school standing, or post-high school plans.

Score Decline and Overall Educational Effectiveness

The SAT is not a measure of educational effectiveness. The

tests are not administered to samples representative of the school population; they do not reflect overall goals of education; they do not predict success in life. They are simply an indication of students' probable academic accomplishments in the first years of college. (Even in this respect their effectiveness varies with the course of study followed in college.) The College Board has taken steps to make these limitations clear, but their cautions are often ignored.

Because the panel felt that the post-1970 decline might reflect a more general decline in educational attainment, they reviewed studies of trends on other standardized academic tests. Here they found a number of anomalies, including some evidence from the National Assessment and from the Armed Forces Qualification Test that scores in some areas (such as functional literacy) had been rising during the period of declining SAT scores. Their conclusion is reasonable though its implications are only partially developed in the remainder of the report:

> There are no reliable comprehensive measures yet of the comparative competence of today's youth with yesterday's. It could occur at the same time (1) that a larger percentage of young people going on to college would be less well equipped for what college has traditionally required, while (2) the general ability level of youth as a whole has increased. (pp. 23-24)

The Panel's Conjectures

To explain the "pervasive" SAT decline over the past six or seven years, the panel included a chapter entitled "Circumstantial Evidence," which begins with a sentence acknowledging that any such explanations are "essentially an exercise in conjecture." It is these conjectures which have drawn most of the attention in the media, and which contain most of the statements which can be used either to support or attack the efforts of teachers of English. In their attempts to trace the causes of the decline, the panel's speculations were wide-ranging; many have been taken out of context by later commentators. What follows is a summary of what the panel did say about a number of specific factors which some have implicated in the score decline.

Elective Programs

The panel found a complex relationship between elective programs and the SAT score decline, and warned against any oversimplistic interpretation. In studies they reviewed, they found no significant relationship between the number of electives added by a particular high school and the experience of students from that school on the SAT. They did conclude, however, that particular elective courses (they cite film making and science fiction) are less demanding than others, and may not offer as good a preparation for the SAT. Even here they caution that

> . . . the new electives are being taken less by students who are going on to college (and will therefore take the SAT) than by those who are not. It will have to be determined whether the needs and interests and developed competencies of those taking these electives are better met by a course, for example, in Radio/Television/Film or an English TV course in the refinements of the language. (p. 26)

Their reservations and qualifications emerge less clearly in the conclusions in the summary chapter. There the discussion of factors in the pervasive score decline begins:

> *One.* 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 of the number of courses that all students alike are required to take. This has been true particularly in the English and verbal skills areas. (p. 46)

It is this bald statement which has become the basis of attacks on elective programs. The full discussion is considerably more

Experimental Teaching Methods

Although the panel heard frequent suggestions that the score decline could be attributed to experimental teaching methods, open classrooms, nongraded courses, and other innovations, they found no evidence in their own or other studies of any relationship between these approaches and the SAT score decline.

Vocational and Experiential Training

On the related question of the introduction of experiential (out-of-school) training in the high schools, they found that students reporting 1 to 15 hours of outside work a week averaged *higher* SAT scores than those doing no such work at all. They also found no evidence that vocational and technically-oriented training had contributed to the score decline.

Back to the Basics

The panel's report has already been used by some editorialists to support a narrowly conceived back-to-the-basics movement. It does not. The panel warns specifically against a return "to uniform drills and exercises commended only by a traditional pedagogy" (p. 26). They argue instead, as have many English teachers, for 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 context for such instruction should be one of variety rather than of rigidity:

> The American education system is unique in its variety and its capacity to be useful to an extremely broad constituency, in which those who are going to take the SAT are a minority. 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)

Careful Writing, Critical Reading

The fundamentals which the panel continually stresses are those of careful writing and critical reading; they return to these points in many different contexts: the unfortunate reliance on tests which require little or no writing; textbooks with nothing but multiple-choice or fill-in-the-blank exercises; declining homework requirements; and electives.

English teachers can agree wholeheartedly with many of the points the panel makes, and can welcome the call for more systematic instruction in writing and critical reading. (In discussing teacher workload, and ruling it out as a factor in the score decline, the panel asks almost parenthetically for reconsideration of teachers' responsibilities "in order to achieve better writing," p. 32.) There is a tendency in the report, however, to single out English instruction for special criticism. Partly because some of the specialized studies focused upon aspects of English teaching, partly because of an implicit identification of "Verbal" scores as a measure of the success of English classes, teachers of English emerge as particularly responsible for the score decline. The report pays little attention to the fact that the critical reading and careful writing—indeed the *thinking*—that the panel consistently recommends, should be part of the whole curriculum, not just of the English class. The panel sometimes ignores what it points out in the beginning—that items on the SAT-Verbal are based on six general areas, including the physical and social sciences, not just on literature. It also ignores the fact that scores on the English Composition Test have actually been rising during the period of decline on other measures, raising a real question (as the panel does note) of the relevance of the SAT-Verbal as a measure of writing skills.

Television

While noting that the research relating television viewing to academic test scores is "entirely inconclusive" (p. 35), the panel nonetheless comes to the "essentially subjective" conclusion that television is one of the causes of the SAT score decline. They base their argument primarily on the assumption that 10,000 to 15,000 hours of precollege television viewing must be reducing the amount of time "which otherwise would have gone into the development of the skills and aptitudes measured on college entrance examinations" (p. 36). They also point out the probably beneficial effects of some children's educational television series, and speculate that television viewing may be developing other abilities or ways of learning than are measured by the SAT.

Learning Standards

The panel found a general lowering of educational standards, reflected in phenomena as diverse as high absentee rates, grade inflation, automatic promotion, reduced homework assignments, and easier textbook content. (Current 11th grade texts were generally found to be written at a 9th or 10th grade reading level.) They recognize, however, that many of these problems arose as a response to changing demands upon the schools, and they caution against simplistic answers such as rigid retention of students who do not meet some prespecified level of "competency":

> The only right answer is to vary the instructional process still more to take account of increased individual differences, but without lowering standards—which we recognize is a form of magic, but one that has been performed in this country for a long time. (p. 47)

Other Factors

The panel found a number of other factors which could be implicated in the score decline. Among the more important of these are changes in the role of the family, the disruption caused by the national traumas of the 1960s and 1970s (assassinations, Vietnam, violent protests, and Watergate), and the lack of motivation which students may have felt as a result of these general societal factors. (Again, these are discussed in the section entitled "Circumstantial Evidence.")

Conclusion

The SAT score decline panel, commissioned to examine a complex topic, found no simple answers. Their report poses many issues which need careful consideration, ranging from the role of standardized tests in the college entrance process to the ways in which all students can be held to the highest standards of which they are capable. The panel's discussions touch on many issues of direct relevance to the teaching of English: elective programs, writing instruction, critical reading, the role of television, the back-to-the-basics movement. The report is worth careful study by teachers of English language arts at all levels, if only because we are likely to suffer the consequences if we allow others to draw the inferences for us. (*On Further Examination* is available from College Board Publication Orders, Box 2815, Princeton, New Jersey 08540, for $4.)

Arthur N. Applebee
(For the NCTE/SLATE Steering Committee on Social and Political Concerns)

SLATE Starter Sheets and Newsletters are offered as resources for dealing with current issues affecting the teaching of English language arts. Reproduce these materials and use them to help promote better understanding of the goals of English teaching.

SLATE OFFERS HELP WITH ACCREDITATION TEAMS

Is your college due for accreditation or renewal of accreditation? If so, SLATE may be able to help. Send SLATE the names and addresses of the members of your accreditation committee, and SLATE will write to each of them, sending copies of the 1977 Workload Statement for College English Teachers, and urging the committee members to measure the acceptability of English Departments in terms of whether the college is meeting, or attempting to meet, those guidelines.

If the college administration is aware of the standards NCTE has set, and realizes that the adoption of those standards may have some bearing on accreditation, overworked English teachers can expect some relief—and the teaching of English at the institutions should improve.

STARTER SHEETS DISTRIBUTED BY STATE OFFICE

The Montana Office of Public Instruction has assisted the Montana Association of Teachers of English in printing and distributing Volume I of the SLATE Starter Sheets. The Montana English Consultant writes: "SLATE's activities are exciting, teachers have been grateful for the clear, efficient presentation of the Starter Sheets exhibit. You may be pleased to know that they have been widely quoted in meetings, newsletters, and at cocktail parties."

CCCC RESOLUTION ON TESTS OF WRITING ABILITY

The Conference on College Composition and Communication, at its annual meeting held in Kansas City in March 1977, passed the following resolution:

> RESOLVED: that CCCC opposes and condemns the imposition of competency, placement, or achievement tests purporting to measure the writing ability of students unless or until those specific tests, and the purposes for which they are to be used, have been examined and approved by the writing teachers at that institution.

The members present at the business meeting voted to send copies of the resolution to the "chief educational decision makers across the states and territories."

SLATE AFFILIATE REPRESENTATIVES

The SLATE Steering Committee has asked each affiliate to nominate a SLATE representative, who will be responsible for coordinating SLATE activities within the affiliate and for encouraging contributions to SLATE. By October 1, nearly 60 affiliates had sent the name of their representative to NCTE headquarters.

In some states, the SLATE representatives have already arranged special distributions of SLATE materials to affiliate members, to school administrators, and to other interested educational organizations. Special sessions to discuss SLATE Starter Sheets and SLATE activities have also been held at affiliate conferences.

COSPONSORED SPEAKERS PROGRAM

SLATE Steering Committee members have also been busy in affiliate activities. Under the auspices of NCTE's Cosponsored Speakers Program, they have spoken to affiliate groups from Florida to Hawaii. Affiliates interested in arranging a visit from a member of the SLATE Steering Committee should contact John C. Maxwell at NCTE headquarters, 1111 Kenyon Road, Urbana, Illinois 61801.

SLATE HOLDS INVITATIONAL WORKSHOP

SLATE is sponsoring a four-hour invitational workshop for affiliate representatives and other interested persons on Thursday afternoon during the NCTE convention in New York. The theme of the workshop is "Responding to Social and Political Concerns," and it will be led by Jesse Perry and Miles Myers, both members of the 1977 SLATE Steering Committee. Participants will discuss SLATE goals and organization; three local issues (censorship, student load, and the selection of classroom materials), a state issue (the minimal competency movement including a review of some sample legislation); and a national issue (ways in which SLATE might help to influence federal policy on testing).

NIE GRANTS REQUESTED

SLATE has asked the National Institute of Education to budget for small research grants ($500 to $1400) for classroom teachers, K-14, who want to explore within their own classrooms small, manageable questions on the teaching of writing, reading, and literature. The SLATE Committee believes that research, on a modest scale, should be one of the responsibilities of classroom English teachers, and that some reward for their efforts would contribute to a community of interest among researchers and practitioners.

John W. Christensen, Acting Director of the Institute, forwarded the request to the Associate Director for Basic Skills with the recommendation that it be considered in the formulation of future plans.

POLITICS AND EDUCATION

SLATE members may be interested in investigating a new publication, *Politics and Education*, to be issued bimonthly at the rate of $6 a year for institutions, $5 for faculty, and $4 for students. The publication will "provide in-depth reports on developments in higher education, critical analyses of the current process of retrenchments, and features on student and faculty political involvement." The publication is "part of the developing nationwide struggle to protect and improve higher education."

People interested in the new publication may write to Wesleyan Station, Fisk Hall, Middletown, Connecticut 06457.

COMPETENCY TESTS PROLIFERATE

SLATE's November 1976 Newsletter reported that five states had passed laws requiring minimal competency testing in their public schools, and that four other states had such legislation pending. *Update IV. Minimal Competency Testing*, prepared by Chris Pipho, Associate Director of the Research and Information Department of the Education Commission of the States, reports that by March 1, 1977, seven states had passed such legislation and another nine states had taken either state board or state department of education action to mandate some form of minimal competency activity. Even more action was expected early in 1977.

Pipho also reports that the "early out" testing now being used in Florida and California remains unique to those two states, legislative and state boards seem to be moving more toward setting standards for high school graduation or, as in the states of Washington and Virginia, toward establishing standards at the junior high or elementary school level.

If you know of actions on minimal competency testing being contemplated in your state or area, Chris Pipho would be glad to have the information. He would also find any analysis, study, or progress report on the implementation of such programs very useful. His address is 1860 Lincoln Street, Denver, Colorado 80295.

SLATE MATERIALS AVAILABLE FROM NCTE

The first two volume of *SLATE* are now available from the NCTE order department, at $1 each. Each volume contains Starter Sheets providing brief summaries of research and professional opinion related to current issues in the teaching of English language arts, and Newsletters describing related activities of SLATE and NCTE. Volume I, emphasizing appropriate responses to the back-to-the-basics movement, includes Starter Sheets on what the "basics" in English are, grammar and usage, composition, language and dialect, spelling, reading, and censorship of instructional materials.

Volume II, emphasizing current issues in standardized testing, contains discussions of minimal competency testing, standardized tests scores, minimal competency testing and bilingual/bicultural students, the National Assessment of Educational Progress, strategies for responding to minimal competency legislation, instruction in phonics, and in this issue, the SAT score decline.

To order your copies, complete the form at the bottom of this page and return it to NCTE/SLATE, 1111 Kenyon Road, Urbana, Illinois 61801.

SLATE CONTRIBUTIONS UP

By the end of July 1976, SLATE had received contributions from just over 600 individuals. A year later, the total had grown to over 1800—and it is still rising.

To make your contribution to SLATE, complete the form below and return it to NCTE/SLATE, 1111 Kenyon Road, Urbana, Illinois 61801. Contributions may be made in any amount, SLATE Newsletter is sent for one year to anyone who contributes $5 or more in support of SLATE activities.

HAVE YOU CONTRIBUTED YET THIS YEAR?

Name _____

Address _____
Number Street

_____
State Zip

Amount of donation ___ $5; ___ $10; ___ $20; ___ other.

Have you contributed to SLATE before?

___ Yes; ___ No.

SLATE Starter Sheets/Newsletters

	Quantity ordered	Total cost
Volume I (Emphasis: Back to the Basics) NCTE Stock Number 98619-76 ($1 per set.)	___	___
Volume II (Emphasis: Standardized Testing) NCTE Stock Number 98619-77 ($1 per set.)	___	___

Total payment enclosed: ___

Make checks payable to NCTE/SLATE. Prepaid orders only, please.