STATE OF CALIFORNIA

EDMUND G. BROWN JR., Governor

COMMISSION FOR TEACHER PREPARATION AND LICENSING
OFFICE OF THE EXECUTIVE SECRETARY
1020 O STREET
SACRAMENTO, CALIFORNIA 95814

(916) 445-0184

February 27, 1979

78-7940

TO: PARTICIPANTS IN THE VALIDITY STUDY

FROM: Peter L. LoPresti, Executive Secretary

SUBJECT: Summary of Validity Study of the N.T.E. Science Area Examination

In May I sent you a summary of the results of validity study of the N.T.E. Common, Social Science, Mathematics and English Examinations. At that time, I promised to send you a copy of the results of the analysis of the science data as soon as it has been completed. The enclosed summary reports the final part of this study.

If after reviewing this paper, you have any questions, please feel free to call Mrs. Marjorie Brodt. Her phone number is (916) 322-2304 or ATSS 8-492-2304.

Again let me thank you for your participation in this important Commission activity.

Enclosure

SUMMARY OF VALIDITY STUDY FOR PHYSICAL AND LIFE SCIENCES

A. PURPOSE

The purpose of the standardized study was to investigate whether there exists a rational basis for the use of the National Teacher Examinations for teacher certification purposes in the State of California, consistent with applicable legal standards and E.T.S. guidelines. Within this general framework are two specific goals: (1) to assess the appropriateness of the examinations as measures of the knowledge required as a qualification to teach in the state; and (2) to develop information that the Commission can consider in establishing rational rules for evaluating the N.T.E. scores of applicants for certification of teachers.

The design of the study requires that the judgements of educators within the state be combined into a "statewide standard" as to content appropriateness and minimum score.

Two types of panels were used in this study. Content Review Panels evaluated the content of the tests in relation to the content of appropriate college programs and in relation to the Commission-adopted Scope and Content Statements. Knowledge Estimation Panels estimated the percentages of minimally-knowledgeable students who would know the answers to individual questions and estimated the number of questions out of each major category listed in E.T.S. specifications that a minimally-qualified applicant should be able to answer correctly.

As part of their study, E.T.S. staff collected and analyzed data for the General Education Test in the Common Examinations and for three Area Examinations: English Language and Literature, Mathematics, and Social Studies. At the request of the Commission, data were also collected, but not analyzed, for two other Area Examinations: Biology and General Science; and Chemistry, Physics, and General Science. The Commission issued a Request for Proposals for the analyses of the data for these two examinations, and a contract was subsequently awarded to Educational Evaluation and Research, Inc. (E.E.R.I.). Since E.T.S. collected the data and since both the E.T.S. and E.E.R.I. results are to be used in the same way and for the same purposes, the E.E.R.I. approach to the analyses of the Science data replicates the appropriate analyses made by E.T.S. for the other examinations. Portions of the E.T.S. study were based on a standardized model that has been used by E.T.S. in other studies of the N.T.E.

B. CONTENT REVIEW RESULTS

1. When comparing relative emphasis on major content topics, the index of comparison ranges from 0-100. Low values represent a high degree of correspondence between relative emphasis in test and curriculum. The index of difference in relative emphasis for Biology and General Science Area Examination was 32 while for Chemistry, Physics and General Science it was 22.

2. To study the overall comprehensiveness of the test, judges were asked to list topics in their college programs not covered by the test. Five topics were listed as being in a Biology and General Science college program and not being covered by the examination, however, only one topic was noted by more than one judge. To be significant, an area must be listed by more than one judge. While in the Chemistry, Physics, and General Science area, 13 topics were noted by the judges as being in their college programs and not covered by the test. However, none of the topics were significant since they were not listed by more than one judge.

3. When judging overall similarity between tests and college programs, the test was considered appropriate when more than half the judges rated the test as very closely parallel or some differences but not appreciable. Using this standard, both the Biology and General Science and Chemistry, Physics, and General Science Area Examinations were considered appropriate since over half the judges rated the list as very closely parallel or some differences but not appreciable.

4. 96% of the questions in the Biology and General Science Examination studied were judged content appropriate for students who have followed standard college programs and 99% of the questions in the Chemistry, Physics, and General Science Examination were judged content appropriate.

C. EVALUATION OF THESE RESULTS

Conclusions about the Correspondence between the Content of the Area Examinations and Program Content in Particular Teaching Fields:

	Degree of Correspondence				
Examination	Very Closely Related	Closely Related	Reason-ably Related	Prob-ably Related	Not Re-lated
Biology & General Science		X			
Chemistry, Physics & General Science	X				

The results for the individual components of the Area Tests are described below:

1. Biology and General Science curriculum is Closely Related to the examination. The judges found 96% of the questions content appropriate. Five topics in the curriculum were listed as not being included in the examination, however, only one was significant since it was listed by more than one judge. The relative emphasis in major content topics in the test and the curriculum

appears to be generally similar; the index of difference was 32. Seven of the nine judges who responded to the question about overall similarity said that the test curriculum are closely parallel or do not differ appreciably.

2. Chemistry, Physics, and General Science curriculum is Very Closely Related to the examination. The judges found 99% of the questions in the list as content appropriate. The index of difference in relative emphasis on major content topics was 22. Low values on a scale from 0-100 represented high degree of correspondence. Thirteen topics were listed by only one judge as not being covered by the test. To be significant, a topic must be listed by more than one judge. Five of the seven judges reported close parallelism or no appreciable difference between test and the college course sequence.

D. KNOWLEDGE ESTIMATION RESULTS

Judges were asked to estimate the percentage of a group of minimally-qualified candidates who would know the answer to each test question. Individual questions were analyzed and results of analyses were summarized. The average scaled scores as determined by these panel judgements appear in column one below:

Test	Column 1	Undergrad. Program Norms	1976-77 Calif. % Equiv.	Present Commission Score Stand.	Stand. Error of Measure.
Area Examinations		% Equiv.			
Biology & General Science	'57	94	95	676	19
Chemistry, Physics & General Science	680	81	80	628	21

The standard error of measurement can be used to judge the range of score theoretically achieved by any person taking the test. There are 68 chances out of 100 that the examinee's observed score will be within one standard error of measurement of the true score. Ninety-five chances that the score will be within two standard errors of measurement. Translating this to the findings of the judges for each test studied:

Test	Average Scaled Scores Determined By Judges	S.E. Meas.	*Score Estimates Adjusted By Applying + or - One Standard Error of Measurement
Area Examinations			
Biology & General Science	757	19	738-776
Chemistry, Physics & General Science	680	21	659-701

*Theoretical Range of Scores for 68 out of 100 students when Observed Score is Average Scaled Score.

E. STANDARD SETTING STUDY

The purpose of this study was to obtain reasoned judgements from qualified professionals regarding the number of questions on selected N.T.E. tests that minimally-qualified applicants should be able to answer correctly. The questions in the tests studied were grouped into the major categories in the E.T.S. test specifications, and the task of these panelists was to estimate the number of questions from each group that a minimally-qualified applicant should be able to answer correctly. Panelists were asked to think of a hypothetical person with minimum knowledge to complete the academic program and teach effectively.

Panel Estimates

The average number of questions in each category of a test and the average number of the total questions judged that should be correctly answered is reported in Table 12. In all tests the data suggests that a minimally-qualified applicant should answer correctly about 75% of the questions. In addition, the variance between number of questions that should be answered in subtest is very limited, never exceeding 19% difference.

Converted Scaled Score Estimates Based on Judgements of Total Panels:

Test	Col. 1	% Equiv. Grad. Prog. Norms	76-77 Calif. % Equiv.	Present Comm. Score Stand.	Stand. Error Meas.
Area Examinations					
Biology & General Science	737	89	90	676	19
Chemistry, Physics & General Science	720	92	91	628	21

F. SCOPE AND CONTENT REVIEW STUDY

The purpose of this study was to ascertain the extent to which selected portions of the N.T.E. Examinations satisfy the test specifications implied by the C.T.P.L. adopted Scope and Content Statements which are outlines of minimum knowledge and competencies to be achieved by a candidate in a given subject matter area. Judges were to place each question within one or more categories of the structure provided by the Scope and Content Statement or to indicate that the question did not fit within that structure. They were also to determine the extent to which specific topics in the Scope and Content Statement were represented by questions in the test. When the necessary information was available in the Scope and Content Statement, they were to determine the degree of correspondence between the relative coverage of various topics specified in the Scope and Content Statement and the judged content distribution in the test.

Fit of Question Content

A question was defined to fit a Scope and Content Statement if more than half the judges who classified the question placed it in one or more of the topical categories. In the Biology and General Science Examination, an average on both forms of 91% of the questions were judged to fit the Biology and General Science Scope and Content Statement and 100% of the Chemistry, Physics, and General Science questions were judged to fit that Scope and Content Statement.

Test Representation of Specific Topics

If at least half the judges placed no questions in a category for a topic, the topic was considered to be unrepresented in the test. All topics in both tests were represented in both test editions.

Estimated Test Score Means for the Scope and Content Statement

Minimal test scores were estimated by using the data from the Knowledge Estimation Panels, and using only those items that were judged to fit the Scope and Content Statement. These scaled scores are shown for each test edition.

Score	Edition 1	Edition 2	Average
	Biology and General Science		
Scaled	730	744	737
	Chemistry, Physics, and General Science		
Scaled	671	692	682

A comparison of the results from each of the studies is shown on the following page. The data indicates that the National Median is below our present score standard and that both of these scores are far below the recommended standards set by the Knowledge Estimate Panels.

SUMMARY OF DATA FROM KNOWLEDGE ESTIMATE AND STANDARD SETTING PANELS

	Biology & General Science			Chemistry, Physics & General Science		
	Score	ETS Norm %ile	1976-77 Calif. %ile Equiv.	Score	ETS Norm %ile	1976-77 Calif. %ile Equiv.
Median 50%	640	50	49	598	50	40
CTPL Standard	676	66	67	628	64	54
KEP Cur College	757	94	95	680	81	80
KEP Scope & Content	737	89	90	682	82	81
SSP Adjusted	737	89	90	720	92	91
SE Meas	19			21		