DOCUMENT RESUME

ED 478 072 TM 035 013

AUTHOR	DeMauro, Gerald E.
TITLE	Construct Properties of New York State English Language Arts and Mathematics Examinations, 1998-1999-2000.
INSTITUTION	New York State Education Dept., Albany.
PUB DATE	2001-01-00
NOTE	58p.
PUB TYPE	Reports - Research (143)
EDRS PRICE	EDRS Price MF01/PC03 Plus Postage.
DESCRIPTORS	Ability; *Achievement Tests; *Construct Validity; Elementary Secondary Education; *English; *Mathematics; *Test Format; Test Items
IDENTIFIERS	*New York

ABSTRACT

Several analyses of the construct validity of the fourth-grade, eighth-grade, and commencement-level English and Mathematics examinations of New York state were performed. The analyses present construct and differential construct elaboration both across tests and within tests. Results show strong relationships among different question types, open-ended and multiple choice, within the same tests and weaker relationships for similar types of questions in diferent tests. These findings indicate that the tests are much more sensitive to skills they are designed to measure than they are to the format of the questions. Simply stated, there is greater evidence that it is mathematics and English that are being measured rather than the ability to answer multiple choice or essay or rubric-scored formats.In particular, the evidence suggests that in the ranges of skills needed to pass the Regents (commencement-level) examination or to achieve competency (proficiency level 3) performance on the fourth and eighth grade tests, the skill intended to be measured is the predominant skill being measured.Two appendixes contain details of the analyses. (Contains 15 tables and 6 references.) (SLD)

Construct Properties of New York State English Language Arts and Mathematics Examinations 1998 - 1999 - 2000

Office of State Assessment
Gerald E. DeMauro
January 2001

Executive Summary

Several analyses of the construct validity of the fourth-grade, eighth-grade, and commencement-level English and Mathematics examinations of New York State follow. The analyses present construct and differential construct elaboration both across tests and within tests.

Results show strong relationships among different question types, open-ended and multiple choice, within the same tests and weaker relationships for similar types of questions in different tests. These findings indicate that the tests are much more sensitive to skills they are designed to measure then they are to the format of the questions. Simply stated, there is greater evidence that it is mathematics and English that are being measured rather than the ability to answer multiple choice or essay or rubric-scored formats. Such findings support the construct validity of the instruments. In particular, the evidence suggests that in the ranges of skills needed to pass the Regents (commencement-level) examination or to achieve competent (proficiency level 3) performance on the fourth and eighth-grade tests, on the skill intended to be measured is the predominant skill measured.

Table of Contents

Table of Contents - continued

Construct Properties of New York State English Language Arts and Mathematical Examinations, 1999-2000, 2000-2001

G. DeMauro, Office of State Assessment

Overview

The fourth and eighth-grade English Language Arts (ELA) and Mathematics examination in New York State are developed, administered, and scored under contract with CTB/McGraw Hill. The Mathematics and English Regents (High school commencement level) examinations are developed by the State Education Department and administered and scored in the schools. Each of these six examinations contains both open-ended or constructed response questions scored with reference to rubrics and multiple choice questions.

Classical test theory conceives of test scores as having components. Validity demands that the components that are irrelevant to the trait or construct being measured have a minimal contribution to the observed score, while the components related to the trait or the construct being measured have the largest contribution to the test score. The irrelevant components may be related to characteristics of the examination, such as the item type, open-ended or multiple choice, in which the question is posed, or to characteristics of the examinees, such as ethnicity. Irrelevant should not have a systematic relationship to individual examinees' capacity to respond. Construct validity, then, is often concerned with the relative contributions of these components to test scores and differential construct validity is concerned with how these relative contributions vary with respect to the demographic characteristics or skill levels of the examinees.

Convergence and Discrimination

One way to estimate the relative contribution of relevant and irrelevant factors to the children's test scores is to examine the convergent and discriminant properties of the examinations. Basically, examinations with greater construct validity yield performances or scores that are demonstrably related to scores or performances on instruments of the same trait. For example, results of one mathematics test should have a clear relationship to results of another mathematics test in the same subject matter. This type of evidence addresses the convergent properties of the examination, or convergent validity.

A second criterion follows from this: Scores on measures of different traits or constructs should not be as well-related to each other as are those from measures of the same or similar constructs. For example, scores on a test of English Language Arts should not be as well related to the scores on a mathematics examination, as are scores from another mathematics examination. This type of construct validity addresses the disciminant properties of the examination, or discriminant validity.

The current examinations can be divided according to the item types that compose them to assess multi-trait multi-method relationships. We hypothesize that performances on different item types measuring the same trait (multi-method) should be better related to each other than performances on the same item types, e.g., multiple choice, across tests of different traits (multitrait). Therefore, the mathematics multiple choice items and rubric-scored or open-ended items should yield results that are better related to each other than they are to the performances on either of these types of questions to the corresponding item types on the English Language Arts examinations. We hypothesize the same for the multiple choice and rubric-scored item types used on the English language Arts examinations.

Professional testing standards (AERA, APA, & NCME, 1999) describe these relationships and their meaning for validity using the following example:

> "For example, within some theoretical frameworks, scores on a multiple choice test of reading comprehension might be expected to relate closely (convergent evidence) to other measures of reading comprehension based on other methods, such as essay responses; conversely, test scores might be expected to relate less closely (discriminant evidence) to measures of other skills, such as logical reasoning." p. 14.

The reader should note that this paper refers to items as rubric-scored because it is the cognitive skill of recall and the scoring provisions for partial credit that most clearly demarcate these item types across tests.

Examinations

The fourth-grade and eighth-grade English Language Arts and Mathematics examinations are pattern-scored. In this paradigm, each possible scale score is associated with an array of probabilities of answering each question correctly or of achieving each score point on the scoring rubrics. Each child's observed pattern of right and wrong answers is then matched to the scoring probabilities and the scale score that maximizes agreement between the child's pattern and the predicted probabilities is assigned to the child.

The Regents examinations consist of the Regents Comprehensive Examination in English (CEE) and a variety of mathematics examinations depending upon the class of the student and the year in which the examination was taken. These examinations are scored on a one to one conversion of raw score totals to scale scores in which 65 is passing.

The Regents Mathematics A examination (M-A) is the newest form of the mathematics examinations, and encompasses about a year and a half of the former mathematical curriculum sequence. Normally, then, students would attempt the M-A sometime during the course of their sophomore year. The test consists of 20 multiple choice and about 15 open-ended questions scored with reference to two-, three-, and four-point rubrics. Mathematics A was first administered in June 1999.

Most students are in the process of the three course mathematics curriculum sequence, however, and are still eligible to meet the mathematics requirement through passing the older versions of the mathematics tests: Course I (M-1), Course II (M-2), and Course III (M-3). Therefore, students who have taken both the CEE and a mathematics Regents (M-1, M-2, M-3, or M-A) will most likely have taken one of the older mathematics examinations. Because the CEE is normally administered in the junior year, the mathematics Regents that is most commonly administered to students who are also taking the CEE is the M-3.

The CEE consists of four sections. Each section is associated with a stimulus that is common to the questions of that section. Each of the four sections contains a long open-ended question that is scored with a 0-6 point rubric. The first three sections also contain six, ten, and ten multiple-choice questions, respectively. M-1 consists of 25 multiple choice and seven open-ended questions. There were not enough students who took M-1 and also took the CEE to permit analyses of the relationship between the two examinations.

M-2 consists of 35 short answers and seven longer open-ended questions. M-3 consists of the same combination of item types a M-2. The newest examination, M-A consists of 20 multiple-choice items and 15 open-ended questions scored on 2-point, 3-point, and 4-point rubrics.

<u>Construct Validity Criteria</u>

Several analyses were employed to evaluate the construct validity of the examinations particularly with respect to the relationships among components of the examinations. In general, the question that was in common to all of the analyses was whether the trait of focus accounted for more of the observed scoring variance than the methods of measurement.

This investigation used the available data. Because the State Education Department (SED) does not collect the scores on the Regents examinations, a sample of Regents papers was solicited from a few school districts. The SED does, however, collect a sample of ten percent of June Regents papers to review, and a sample of these papers were analyzed as a follow-up of the studies to examine within-test convergent and discriminant properties. As well, a special April 2000 administration of CEE for seniors only was scored by SED and provides more within-test data, although the special nature of this large April sample restricts its generalizability. Obviously, there is no claim that the across-test data (e.g., matched mathematics and CEE groups) represent the State, so the analysis of the Regents in this study has a more limited generalizability than the analyses of the fourth- and eighth-grade instruments, for which all item and test level data for the whole state population of examinees are available within and across tests. The available Regents data consisted of item-level data for both the mathematics examinations and the CEE and whole test scores, only, for the matched samples taking both the CEE and the Mathematics tests.

Sample Sizes

As explained above, the entire fourth- and eighth-grade test populations were used for the analyses of those tests. Data were available from the 1998-1999 and the 1999-2000 academic years for the fourth- and eighth-grade examinations. Data across tests on the Regents examinations were available for the June 1999 administration-within test Regents data from Department review were available from June 1999, April 2000 (CEE only) and June 2000. Sample sizes are given on the next page:

Table 1

Sample Sizes for the Construct Analyses
Of New York State English Language Arts
and Mathematics Examinations

Test	Group	1998-1999 Number	1999-2000 Number	April 2000 Number
Matched Grade 4:	African American	36,993	41,693	
	American Indian/N.Amer.	572	801	
	Asian American	9,427	10,412	
	European American	108,407	117,561	
	Hispanic American	31,700	34,891	
	All Groups	196,808	206,127	
Matched Grade 8:	African American	32,096	39,428	
	American Indian/N.Amer.	566	717	
	Asian American	8,356	10,129	
	European American	112,696	130,726	
	Hispanic American	25,435	31,342	
	All Groups	185,299	214,000	
CEE/M-3	All Groups	29		
Across Test Level Regents Analyses				
CEE/M-1	All Groups	130		
CEE/M-2	All Groups	64		
CEE/M-3	All Groups	117		
CEE/M-A	All Groups	53		
Within Regents Analyses				
CEE	All Groups	488	1,787	6,825
Mathematics A	All Groups	385	1,294	

Methods

Construct Properties of the Examinations

The primary means of evaluating the construct properties of the tests in this study was a multitrait-multi-method analysis (Campbell & Fiske, 1959) with a variety of follow-up procedures. In particular, this analysis examines the convergent and discriminant properties of the instruments, as described earlier.

For the Regents examinations, the fourth-grade, and the eighth-grade instruments, a 4x4 correlation matrix was computed for the total points achieved on the short answer, multiple choice or non-rubric-scored questions and for the total points achieved on the open-ended rubric-scored questions for the English and mathematics examinations. For the fourth- and eighth-grade examinations, data were available for the following self-identified ethnic groups: African American, American Indian/Native American, Asian American, European American and Hispanic American. Data were also available by six school district community types: New York City, Big Four Large Cities, Urban/Suburban High Needs, Rural High Needs, Average Needs, and Low (affluent) Needs.

For both the open-ended and multiple-choice point totals, reliability was estimated using Cronbach's alpha (Lord & Novick, 1968). Because individual item-level data were not available for the Regents examinations on the open-ended rubric-scored questions, reliability could not be directly estimated for the totals on these questions. Reliabilities for the CEE and for M-A, only, were estimated based on the Department Review process, which is a random sample of papers that are rescored by trained consultants. This sample included about 500 test papers from 1999 and about 1200 test papers from 2000 for each of these two subjects (exact sample sizes given earlier).

As described above, classical testing theory holds that each score or point total is composed of the true score of the student and some randomly distributed error component. The greater the proportion of the true score to the observed score, the greater the reliability of the score. Construct validity analyses, such as these, are ultimately concerned with the true score relationships of parts of the tests and parts of different tests. When the degree of relationship is estimated, the observed correlations among parts of tests are adjusted when possible to account for the unreliability or the error components of the observed student performances. The random distribution of error (see Lord & Novick, 1969) has the effect of suppressing, or attenuating the correlation among the parts of the tests because it adds to each score a component that is uncorrelated and random. Therefore, wherever possible, the correlations cited in the analyses have been disattenuated, or corrected for unreliability (see Thorndike & Hagen, 1969, for example).

Multitrait-multi-method analyses have four criteria:

1. Non-nominal correlations (greater or equal than .35) of traits (English or mathematics) across methods;
2. Higher correlations within traits across methods (open-ended or short or multiple choice) than across both traits and methods;
3. Higher correlations within traits across traits within methods;
4. Relationships among traits that follow the same pattern regardless of method.

These analyses require partitioning of the scores into total points achieved in relation to item types. Because data were only available on the whole test for many Regents examinees, this was impossible, so a second series of analyses was performed on the whole test data available on the Regents examinations. The focus of these analyses was to estimate the degree to which the skills measured on one test intruded on the performance on another test, e.g., communication skills measured by the CEE on the mathematics skills measured by the four Regents examinations. Specifically, we were interested in whether or not a proficient or passing performance in English was necessary to pass the mathematics tests. This would indicate that the relationship between the two instruments puts the students into a "double jeopardy" situation in which the second cannot be passed without passing the first.

<u>Dimensionality of the Tests</u>

Previous analyses of the Mathematics A and the CEE (AES, 1999) indicate that these instruments are unidimensional. That is, each of these instruments measures one predominant factor. By design, these factors would be English, as delineated by the New York State Learning Standards and mathematics, also as delineated by the Learning Standards. We suspect, therefore, that we should see clear evidence of good convergent properties within the tests. That is, that all of the items predominantly measure the same trait. However, recent trends in mathematics instruction emphasize the student's ability to discern the important elements in solving a problem from information that is less important. This requires reading skills.

Nevertheless, such reading to do math is often demanded in a social environment like the classroom where clarification of exceptional reading or communication demands is available that is not available in a high stakes testing environment. Even attempts to provide real world contexts for mathematics must be mindful that the testing environment does not permit the collaborative possibilities of other environments. For this reason, the discriminant properties of the test would require that the reading encountered on the mathematics examinations not be so difficult that it contaminates good measurement of the mathematics skills. These construct validity properties were evaluated using multilinear regression analyses and post hoc planned quantitative comparisons (Myers, 1972).

Specifically, the CEE scores were divided into four categories: below 55 (the score that could be used to meet requirements for a local diploma), 55-64 (meeting local, but falling short of a Regents diploma), 65-84 (meeting the Regents diploma requirement), and 85-100 (meeting the requirement of graduation with distinction). General Linear Model regression analyses were employed to identify the relationships between achievement of these categories and scores on the Mathematics A examination. Quantitative post hoc analyses were also employed (Myers, 1972) to further elucidate the relationship. In particular, the discriminant validity demands could be met by showing that, while certain increases in the communication skills measured in CEE would benefit the student on the Mathematics A examination, that beyond a certain modicum of skills more skill does not confer an additional advantage. This would be demonstrated by significant nonlinear relationships between scores on the two examinations.

Follow Up Analyses

For the fourth- and eighth-grade examinations further analyses examined the precision with which the reading component of the English Language Arts (ELA) examinations predicted rubric-scored ELA performance and multiple-choice mathematics performance. This was examined for various community types and ethnic groups again to elaborate the construct validity properties of the examination and examine the differential properties.

Within test analyses were also made of the Regents CEE and Mathematics A examinations. These analyses evaluated how the cognitive demands of the components of each examination varied and the nature of the interrelationship of these demands.

Results

Grade 4 and 8

Tables 2-5 provide the multitrait-multi-method analyses. Given the four criteria, the tables provide the disattenuated (corrected) correlation coefficients and the internal consistency reliabilities for each group on the grades four and eight English and mathematics examinations, including validity correlations, which are the underlined trait correlations (mathematics with mathematics, English with English), within item types correlations (multiple choice English to multiple choice mathematics, rubric-scored mathematics), and correlations of totals across both traits and methods. The four evaluation criteria are summarized below:

1. The validity correlations for each grade level, for both mathematics and English for both years all exceed .35;
2. The validity correlations are higher in each case in both English and mathematics than across both traits and methods (e.g., English rubric-scored and mathematics multiple choice or English multiple choice and mathematics rubric-scored);
3. The validity correlations, both for English and mathematics, were higher in both years for the total eighth-grade group than were the correlations across tests either for multiple choice items or for open-ended items; they were not higher in fourth grade CEE, where the multiple-choice component correlations across tests exceeded the rubric-scored to multiple-choice correlations;
4. In fourth grade, for European American students, the English validity correlation was lower than the correlation of multiple choice totals across tests for 1998-1999 year and the 1999-2000 year. For all other groups, the disattenuated validity correlations were the two highest, and the correlations of English multiple choice totals and mathematics open-ended totals were the lowest. For eighth grade, both validity correlations were the highest in all cases. In all other cases, for both years except for English for the small sample (n=428) of American Indians/Native Americans for the 1999-2000 year, the next highest correlations were for open-ended questions across tests, and the least high correlations were across both tests and methods. Clearly there are discernable patterns in both grade levels, and for both years.

Table 2

Multitrait Multimethod Correlation Matrix
New York State Examinations, 1998-1999
Grade 4 English Language Arts and Mathematics

Attenuated Above Diagonal, Disattenuated Below Diagonal
(Reliabilities in Parentheses)

All Students (n = 196,808)

			ELA		Math	
	Mean	S.D.	M.C.	O.E.	M.C.	O.E.
ELA-M.C.	20.59	4.68	(.817)	.656	.672	.660
ELA-O.E.	8.41	2.80	.814	(.795)	.644	.658
Math-M.C.	22.82	4.97	.813	.788	(.837)	.821
Math-O.E.	26.75	7.76	.777	.784	.954	(.884)

African Americans (n = 36,993)

			ELA		Math	
	Mean	S.D.	M.C.	O.E.	M.C.	O.E.
ELA-M.C.	18.18	4.76	(.794)	.661	.625	.621
ELA-O.E.	7.26	2.84	.826	(.806)	.629	.651
Math-M.C.	19.88	5.23	.774	.774	(.819)	.791
Math-O.E.	22.06	8.04	.742	.773	.931	(.880)

Asian Americans (n = 9,427)

			ELA		Math	
	Mean	S.D.	M.C.	O.E.	M.C.	O.E.
ELA-M.C.	21.21	4.22	(.787)	.658	.646	.648
ELA-O.E.	9.36	2.70	.845	(.771)	.616	.637
Math-M.C.	24.54	4.17	.811	.781	(.808)	.807
Math-O.E.	29.49	6.86	.783	.778	.963	(.870)

Table 2 **1998-1999** (continued)

European Americans (n=108,407

			ELA		Math	
	Mean	S.D.	M.C.	O.E.	M.C.	O.E.
ELA-M.C.	22.12	3.89	(.766)	.548	.581	.560
ELA-O.E.	9.09	2.49	.728	(.740)	.555	.571
Math-M.C.	24.39	4.07	.748	.726	(.788)	.571
Math-O.E.	29.19	6.43	.697	.722	.946	(.844)

Hispanic Americans (n-31,700)

			ELA		Math	
	Mean	S.D.	M.C.	O.E.	M.C.	O.E.
ELA-M.C.	17.94	4.90	(.802)	.685	.645	.633
ELA-O.E.	7.20	2.90	.847	(.817)	.647	.656
Math-M.C.	20.39	5.19	.794	.789	(.822)	.801
Math-O.E.	23.00	8.00	.753	.773	.942	(.881)

Native Americans/American Indians (n-572)

			ELA		Math	
	Mean	S.D.	M.C.	O.E.	M.C.	O.E.
ELA-M.C.	19.81	4.73	(.811)	.612	.611	.608
ELA-O.E.	7.45	2.74	.763	(.793)	.609	.612
Math-M.C.	21.58	4.87	.752	.758	(.813)	.792
Math-O.E.	24.90	7.54	.723	.749	.941	(.870)

Table 3

Multitrait Multimethod Correlation Matrix
New York State Examinations, 1998-1999
Grade 8 English Language Arts and Mathematics

Attenuated Above Diagonal, Disattenuated Below Diagonal
(Reliabilities in Parentheses)

All Students (n = 185,299)

			ELA		Math	
	Mean	S.D.	M.C.	O.E.	M.C.	O.E.
ELA-M.C.	19.48	4.26	(.852)	.679	.573	.586
ELA-O.E.	10.87	3.44	.843	(.796)	.549	.597
Math-M.C.	17.96	5.28	.666	.660	(.868)	.832
Math-O.E.	20.12	9.98	.669	.706	.941	(.901)

African Americans (n = 32,096)

			ELA		Math	
	Mean	S.D.	M.C.	O.E.	M.C.	O.E.
ELA-M.C.	17.58	4.67	(.794)	.670	.561	.583
ELA-O.E.	9.21	3.44	.837	(.806)	.540	.606
Math-M.C.	14.71	4.99	.695	.665	(.820)	.767
Math-O.E.	13.18	8.50	.697	.719	.903	(.881)

Asian Americans (n = 8,356)

			ELA		Math	
	Mean	S.D.	M.C.	O.E.	M.C.	O.E.
ELA-M.C.	20.23	4.02	(.787)	.694	.625	.646
ELA-O.E.	12.21	3.41	.890	(.772)	.588	.648
Math-M.C.	20.39	5.01	.783	.744	(.809)	.841
Math-O.E.	24.55	10.00	.780	.790	1.001	(.872)

Table 3- **1998-1999** (continued)

European Americans (n=112,696)

			ELA		Math	
	Mean	S.D.	M.C.	O.E.	M.C.	O.E.
ELA-M.C.	20.40	3.68	(.769)	.626	.479	.490
ELA-O.E.	11.55	3.17	.829	(.742)	.450	.500
Math-M.C.	19.27	4.77	.593	.568	(.847)	.812
Math-O.E.	22.94	9.05	.629	.652	.991	(.792)

Hispanic Americans (n=25,435)

			ELA		Math	
	Mean	S.D.	M.C.	O.E.	M.C.	O.E.
ELA-M.C.	17.43	4.71	(.803)	.679	.575	.591
ELA-O.E.	9.45	3.44	.837	(.818)	.542	.602
Math-M.C.	15.26	5.07	.708	.661	(.823)	.773
Math-O.E.	14.33	8.65	.702	.709	.908	(.881)

Native Americans/American Indians (n=566)

			ELA		Math	
	Mean	S.D.	M.C.	O.E.	M.C.	O.E.
ELA-M.C.	18.65	4.27	(.814)	.655	.494	.525
ELA-O.E.	10.03	3.25	.814	(.796)	.454	.519
Math-M.C.	16.73	4.91	.607	.565	(.813)	.786
Math-O.E.	17.72	8.74	.623	.623	.934	(.872)

Table 4
Multitrait Multimethod Correlation Matrix
New York State Examinations, 1999-2000
Grade 4 English Language Arts and Mathematics

Attenuated Above Diagonal, Disattenuated Below Diagonal
(Reliabilities in Parentheses)

All Students (n = 206,127)

			ELA		Math	
	Mean	S.D.	M.C.	O.E.	M.C.	O.E.
ELA-M.C.	21.74	4.59	(.824)	.666	.703	.691
ELA-O.E.	8.79	2.71	.825	(.790)	.647	.667
Math-M.C.	23.41	5.09	.837	.786	(.856)	.845
Math-O.E.	26.66	8.04	.810	.798	.960	(.882)

African Americans (n = 41,693)

			ELA		Math	
	Mean	S.D.	M.C.	O.E.	M.C.	O.E.
ELA-M.C.	19.52	5.08	(.827)	.661	.678	.654
ELA-O.E.	7.66	2.76	.814	(.797)	.633	.780
Math-M.C.	20.51	5.70	.807	.768	(.853)	.796
Math-O.E.	21.62	8.15	.771	.650	.924	(.871)

Asian Americans (n = 8,396)

			ELA		Math	
	Mean	S.D.	M.C.	O.E.	M.C.	O.E.
ELA-M.C.	22.45	4.10	(.790)	.663	.678	.693
ELA-O.E.	9.65	2.57	.852	(.768)	.611	.654
Math-M.C.	25.46	4.14	.839	.767	(.828)	.802
Math-O.E.	29.52	7.14	.840	.804	.950	(.861)

Table 4

Multitrait Multimethod Correlation Matrix
New York State Examinations, 1999-2000
Grade 4 English Language Arts and Mathematics

Attenuated Above Diagonal, Disattenuated Below Diagonal
(Reliabilities in Parentheses)

European Americans (n = 117,561)

			ELA		Math	
	Mean	S.D.	M.C.	O.E.	M.C.	O.E.
ELA-M.C.	23.15	3.58	(.756)	.561	.703	.691
ELA-O.E.	9.48	2.39	.752	(.736)	.611	.654
Math-M.C.	24.98	3.88	.837	.767	(.785)	.802
Math-O.E.	29.38	6.56	.810	.804	.950	(.835)

Hispanic Americans (n = 34,891)

			ELA		Math	
	Mean	S.D.	M.C.	O.E.	M.C.	O.E.
ELA-M.C.	19.42	5.15	(.829)	.691	.697	.678
ELA-O.E.	7.57	2.82	.846	(.806)	.655	.668
Math-M.C.	20.96	5.61	.827	.788	(.857)	.814
Math-O.E.	22.67	8.12	.796	.796	.939	(.875)

Native American/American Indian (n = 801)

			ELA		Math	
	Mean	S.D.	M.C.	O.E.	M.C.	O.E.
ELA-M.C.	20.54	4.89	(.831)	.652	.677	.654
ELA-O.E.	7.90	2.75	.802	(.796)	.620	.634
Math-M.C.	22.16	5.43	.804	.751	(.855)	.789
Math-O.E.	24.39	7.89	.768	.761	.914	(.872)

Table 5

Multitrait Multimethod Correlation Matrix
New York State Examinations, 1999-2000
Grade 8 English Language Arts and Mathematics

Attenuated Above Diagonal, Disattenuated Below Diagonal
(Reliabilities in Parentheses)

All Students (n=214,000)

			ELA		Math	
	Mean	S.D.	M.C.	O.E.	M.C.	O.E.
ELA-M.C.	19.67	4.11	(.815)	.674	.680	.682
ELA-O.E.	11.04	3.56	.837	(.796)	.632	.674
Math-M.C.	18.01	5.43	.816	.767	(.852)	.845
Math-O.E.	21.36	10.60	.792	.793	.960	(.909)

African Americans (n=39,428)

			ELA		Math	
	Mean	S.D.	M.C.	O.E.	M.C.	O.E.
ELA-M.C.	17.47	4.48	(.804)	.662	.622	.626
ELA-O.E.	9 .49	3.35	.828	(.794)	.588	.635
Math-M.C.	14.41	5.12	.773	.734	(.807)	.788
Math-O.E.	14.19	9.13	.743	.758	.935	(.882)

Asian Americans (n=10,129)

			ELA		Math	
	Mean	S.D.	M.C.	O.E.	M.C.	O.E.
ELA-M.C.	20.52	3.94	(.821)	.699	.686	.695
ELA-O.E.	12.30	3.31	.865	(.796)	.641	.680
Math-M.C.	20.10	5.13	.817	.776	(.858)	.854
Math-O.E.	26.00	10.57	.804	.798	.966	(.912)

Table 5

Multitrait Multimethod Correlation Matrix
New York State Examinations, 1999-2000
Grade 8 English Language Arts and Mathematics

Attenuated Above Diagonal, Disattenuated below Diagonal
(Reliabilities in Parentheses)

European Americans (n=130,726)

			ELA		Math	
	Mean	S.D.	M.C.	O.E.	M.C.	O.E.
ELA-M.C.	20.74	3.38	(.762)	.606	.618	.626
ELA-O.E.	11.74	3.07	.794	(.762)	.566	.622
Math-M.C.	19.67	4.71	.784	.718	(.816)	.813
Math-O.E.	24.61	9.53	.762	.756	.957	(.886)

Hispanic Americans (n=31,342)

			ELA		Math	
	Mean	S.D.	M.C.	O.E.	M.C.	O.E.
ELA-M.C.	17.69	4.58	(.816)	.682	.635	.641
ELA-O.E.	9 .65	3.37	.846	(.797)	.594	.640
Math-M.C.	14.93	5.13	.781	.740	(.811)	.800
Math-O.E.	15.40	9.38	.754	.762	.945	(.884)

American Indians/Native Americans (n=717)

			ELA		Math	
	Mean	S.D.	M.C.	O.E.	M.C.	O.E.
ELA-M.C.	18.87	4.05	(.796)	.611	.632	.642
ELA-O.E.	10.05	3.31	.767	(.797)	.570	.618
Math-M.C.	16.65	5.27	.775	.699	(.836)	.812
Math-O.E.	17.98	9.94	.759	.731	.937	(.899)

Construct Elaboration: Grades Four and Eight

Because the multiple choice questions for the Fourth Grade ELA are all reading questions, it was hypothesized that the strong relationships between these questions and the totals on each of the two components of the fourth-grade Mathematics examination totals reflect the heavy reading demands of the mathematics items. In particular, for younger children, mathematics items that are scored dichotomously, e.g., multiple choice questions, may depend even more on reading skills because there is no partial credit that can be assigned for proper procedure after an initial misinterpretation.

A review of the correlation coefficients in Tables 2 through 5 show that, for European American students, the magnitude of the correlations between the multiple choice (reading) English Language Arts questions and each type of mathematics questions supports this hypothesis. However, for all other students, these coefficients are lower than either of the within test correlations indicating the appropriate use of discernible mathematics and English language skills. Also note that for European American students, the correlation coefficients between ELA multiple choice and mathematics questions was actually lower than it was for other students, as were all of the correlations, in general.

Because ethnicity is distributed disproportionately according to community type, or needs resource category, as community type is expressed in New York State, a secondary analysis was undertaken in which the questions scored by rubrics on the fourth-grade and on the eighth-grade ELA examinations were identified as listening, reading, independent writing, and writing mechanics, according to the test blueprints. A seven by seven correlation matrix examined the interrelationships among each of these categories separately, the ELA multiple choice questions (reading), the Mathematics multiple choice questions, and the Mathematics questions scored by rubrics. These analyses were conditioned on community type, or needs resource categories. The matrices were examined to determine which components of the fourth-grade ELA increase with district affluence.

Note first from Tables 6 and 7 that the correlations between the ELA multiple choice sections and the two components of the Mathematics examinations in both grades are higher than those between the ELA multiple choice sections and each of the rubric-scored components of ELA. This is somewhat to be expected from the restricted range of the four rubric-scored ELA components.

Table 6

Correlation Coefficients between Multiple Choice and Rubric-Scored Questions on the Grade 4 ELA and Mathematics Questions by Needs Resource Categories, 1999-2000 Administrations (Reading totals underlined)

Needs Resource		ELA-MC	LIST.	IND. WRIT	WRIT. MECH.	READ.	MATH MC	MATH OE
NYC	ELA-MC	1.000						
	LIST.	.567	1.000					
	IND. WRIT.	.511	.469	1.000				
	WRIT. MECH.	.594	.530	.668	1.000			
	READ	.635	.538	.506	.589	1.000		
	MATH MC	.727	.541	.492	.581	.609	1.000	
	MATH OE	.714	.556	.505	.589	.637	.840	1.000
BIG FOUR	ELA-MC	1.000						
	LIST.	.462	1.000					
	IND. WRIT.	.411	.416	1.000				
	WRIT. MECH.	.470	.465	.571	1.000			
	READ.	.528	.467	.422	.498	1.000		
	MATH MC	.615	.449	.406	.477	.494	1.000	
	MATH OE	.578	.436	.396	.452	.520	.753	1.000
URB./SUB. HIGH	ELA-MC	1.000						
	LIST.	.439	1.000					
	IND. WRIT.	.374	.375	1.000				
	WRIT. MECH.	.452	.421	.540	1.000			
	READ.	.505	.439	.394	.460	1.000		
	MATH MC	.617	.426	.374	.443	.467	1.000	
	MATH OE	.584	.436	.386	.448	.507	.754	1.000

Table 6

Correlation Coefficients between Multiple Choice and Rubric-Scored Questions on the Grade 4 ELA and Mathematics Questions by Needs Resource Categories, 1999-2000 Administrations (Reading totals underlined)

Needs Resource		ELA-MC	LIST.	IND. WRIT.	WRIT. MECH	READ	MATH MC	MATH OE
RURAL HIGH	ELA-MC	1.000						
	LIST.	.392	1.000					
	IND. WRIT.	.325	.341	1.000				
	WRIT. MECH.	.396	.379	.466	1.000			
	READ	.489	.419	.372	.440	1.000		
	MATH MC	<u>.562</u>	.371	.326	.378	<u>.448</u>	1.000	
	MATH OE	<u>.542</u>	.377	.335	.387	<u>.477</u>	.696	1.000
AVERAGE	ELA-MC	1.000						
	LIST.	.391	1.000					
	IND. WRIT.	.340	.336	1.000				
	WRIT. MECH.	.407	.378	.512	1.000			
	READ.	.472	.411	.376	.436	1.000		
	MATH MC	<u>.580</u>	.373	<u>.339</u>	.400	<u>.440</u>	1.000	
	MATH OE	<u>.562</u>	.383	<u>.353</u>	.408	<u>.478</u>	.729	1.000
LOW	ELA-MC	1.000						
	LIST.	.324	1.000					
	IND. WRIT.	.290	.298	1.000				
	WRIT. MECH.	.341	.333	.490	1.000			
	READ.	.399	.348	.327	.387	1.000		
	MATH MC	<u>.526</u>	.310	<u>.293</u>	.352	<u>.376</u>	1.000	
	MATH OE	<u>.533</u>	.317	<u>.307</u>	.354	<u>.412</u>	.702	1.000

Table 7

Correlation Coefficients between Multiple Choice and Rubric-Scored Questions on the Grade 8 ELA and Mathematics Questions by Needs Resource Categories, 1999-2000 Administrations (Reading totals underlined)

Needs Resource		ELA-MC	LIST.	READ	IND. WRIT.	WRIT. MECH	MATH MC	MATH OE
NYC	ELA-MC	1.000						
	LIST.	.633	1.000					
	READ	.612	.637	1.000				
	IND. WRIT.	.522	.535	.508	1.000			
	WRIT. MECH.	.557	.559	.544	.669	1.000		
	MATH MC	.680	.575	.567	.471	.517	1.000	
	MATH OE	.681	.610	.601	.505	.546	.841	1.000
BIG FOUR	ELA-MC	1.000						
	LIST.	.543	1.000					
	READ	.549	.574	1.000				
	IND. WRIT.	.465	.514	.460	1.000			
	WRIT. MECH.	.505	.551	.500	.651	1.000		
	MATH MC	.635	.492	.496	.426	.476	1.000	
	MATH OE	.630	.512	.523	.445	.490	.800	1.000
URB./SUB. HIGH	ELA-MC	1.000						
	LIST.	.570	1.000					
	READ	.539	.597	1.000				
	IND. WRIT.	.465	.513	.468	1.000			
	WRIT. MECH.	.494	.518	.496	.609	1.000		
	MATH MC	.647	.535	.499	.447	.484	1.000	
	MATH OE	.653	.583	.557	.489	.524	.823	1.000

Table 7 - continued

Correlation Coefficients between Multiple Choice and Rubric-Scored questions on the Grade 8 ELA and Mathematics questions by Needs Resource Categories, 1999-2000 Administrations (Reading totals underlined)

Needs Resource		ELA-MC	LIST.	READ	IND. WRIT.	WRIT. MECH	MATH MC	MATH OE
RURAL HIGH	ELA-MC	1.000						
	LIST.	.532	1.000					
	READ	.516	.582	1.000				
	IND. WRIT.	.410	.475	.452	1.000			
	WRIT. MECH.	.434	.491	.472	.563	1.000		
	MATH MC	.606	.487	.482	.391	.426	1.000	
	MATH OE	.614	.535	.537	.432	.459	.803	1.000
AVERAGE	ELA-MC	1.000						
	LIST.	.514	1.000					
	READ	.491	.553	1.000				
	IND. WRIT.	.420	.462	.437	1.000			
	WRIT. MECH.	.443	.476	.466	.574	1.000		
	MATH MC	.613	.476	.461	.401	.433	1.000	
	MATH OE	.620	.523	.508	.440	.466	.809	1.000
LOW	ELA-MC	1.000						
	LIST.	.476	1.000					
	READ	.456	.527	1.000				
	IND. WRIT.	.397	.430	.406	1.000			
	WRIT. MECH.	.415	.431	.423	.555	1.000		
	MATH MC	.591	.437	.424	.372	.402	1.000	
	MATH OE	.599	.475	.463	.404	.428	.802	1.000

The correlations between the reading multiple-choice questions and the two components of mathematics decrease monotonically across community types until the average needs districts, where they rise. In fact, the heaviest concentrations of European American students are in these districts, accounting for 47.8 percent of the fourth grade examinees and 48.6 percent of the eighth grade examinees. In contrast, among all other students, the largest concentrations are in New York City, accounting for 70.0 percent of the fourth grade examinees and 70.7 percent of the eighth grade examinees.

In grade four, as the needs resources designations change from New York City through High Needs Rural, the correlations between the ELA multiple-choice section and the two components of the Mathematics examinations drop steadily, as do the correlations between the ELA multiple-choice section and each of the rubric-scored components. At the Average Needs Resource category, however, the correlation between the multiple-choice reading section of the ELA examination and Independent Writing and the Writing Mechanics totals rises, as does the correlation between the multiple-choice reading and the two components of the Mathematics examination. On the other hand, the correlations between each of the two mathematics components and each of these two writing components, which drops steadily as affluence increases, become higher in the Average Needs districts than in the Rural High Needs districts. This pattern is replicated with the grade eight data (Table 7). Clearly further analyses are needed to explain this complex pattern.

Changes in Relation to Scoring Range

An evaluation was made of where in the scoring ranges were multiple-choice reading items most sensitive to the differential improvements in the rubric-scored ELA, multiple-choice mathematics, and rubric-scored mathematics questions. That is, for what levels of proficiency are reading skills more important. To evaluate this, the multiple-choice mathematics totals, the rubric-scored mathematics totals, and the open-ended English Language Arts totals for grades four and eight, for the 1999-2000 administration were regressed onto the ELA multiple choice (reading) totals. An analysis of the residuals that is, of the differences between the predicted total based on the regressions and the observed totals, could then determine the precision of the prediction of these totals from the reading measure at different points in the mathematics scales. In this way, it could be determined whether reading for the European American population predicts multiple-choice mathematics performance or rubric-scored mathematics performance better than it predicts rubric-scored English performance throughout the whole range of mathematics scoring, or is reading more important at certain score ranges of mathematics. In effect, the analysis addresses the differential utility of reading for different populations for mathematics performance.

To evaluate this, the squared residuals for the multiple-choice mathematics totals were compared to those for the rubric-scored ELA totals as a repeated measure in a General Linear Regression Model, and evaluated, as well, for European Americans and all other students according to proficiency levels in mathematics. Because this procedure was designed to evaluate the observed component correlations shown in Tables 2-5, and the correlations between components of the tests were computed for these groups separately, the initial regressions of multiple-choice reading onto multiple-choice and rubric-scored mathematics and rubric-scored ELA were also computed separately for two ethnic groups, European Americans and all others. The regressions were also computed separately within ethnic groups for each level of mathematics proficiency. While this has the certain effect of restricting range and reducing prediction accuracy, it controls against the residuals merely reflecting ethnic group and proficiency level differences related to distance from the overall scoring means.

The analyses were simplified to compare only European Americans to all other students because European American students manifested the different pattern of interrelationships among test components. Finally, squared residuals rather than positive and negative residual values were chosen as measures of the precision of the regression for each student. That is, it is the difference of each student from the prediction based on reading that is of primary interest. The reader will recognize the square root of the mean squared residuals as the standard error of estimation.

The results of the regressions are shown in Table 8

Table 8

Regression of Multiple Choice English Language Arts (Reading) Totals onto Multiple Choice Mathematics and Rubric-Scored English Language Arts, by Grade and Ethnicity (European American or Non-European American), 1999-2000 Administrations

Grade Level	Math Level	Ethnicity	Independent Variable	Number	Slope	Intercept	R-Square
4	1	Eur.-Am.	ELA-OE	2,839	0.208	2.047	0.211
		Non-EA	ELA-OE	14,045	0.272	0.939	0.326
	2	Eur.-Am.	ELA-OE	20,217	0.202	3.508	0.135
		Non-EA	ELA-OE	32,473	0.224	2.959	0.178
	3	Eur.-Am.	ELA-OE	63,890	0.261	3.378	0.131
		Non-EA	ELA-OE	33,850	0.275	3.035	0.170
	4	Eur.-Am.	ELA-OE	31,767	0.283	3.879	0.097
		Non-EA	ELA-OE	8,290	0.273	4.178	0.109
	1	Eur.-Am.	Math-MC	2,839	0.211	9.081	0.096
		Non-EA	Math-MC	14,045	0.283	8.052	0.165
	2	Eur.-Am.	Math-MC	20,218	0.183	16.563	0.058
		Non-EA	Math-MC	32,473	0.203	15.794	0.073
	3	Eur.-Am.	Math-MC	63,890	0.178	21.268	0.061
		Non-EA	Math-MC	33,850	0.165	24.493	0.063
	4	Eur.-Am.	Math-MC	31,767	0.105	25.621	0.029
		Non-EA	Math-MC	8,290	0.096	25.958	0.031

Table 8

Regression of Multiple Choice English Language Arts (Reading) Totals onto Multiple Choice Mathematics and Rubric-Scored English Language Arts, by Grade and Ethnicity (European American or Non-European American), 1999-2000 Administrations

Grade Level	Math Level	Ethnicity	Independent Variable	Number	Slope	Intercept	R-Square
4	1	Eur.-Am.	Math-OE	2,839	0.268	6.794	0.098
		Non-EA	Math-OE	14,045	0.355	5.050	0.159
	2	Eur.-Am.	Math-OE	20,218	0.119	18.581	0.016
		Non-EA	Math-OE	32,473	0.139	17.350	0.023
	3	Eur.-Am.	Math-OE	63,890	0.258	23.388	0.048
		Non-EA	Math-OE	33,850	0.223	23.383	0.045
	4	Eur.-Am.	Math-OE	31,767	0.192	31.481	0.035
		Non-EA	Math-OE	8,290	0.161	32.014	0.032
8	1	Eur.-Am.	ELA-OE	14,366	0.329	2.713	0.233
		Non-EA	ELA-OE	34,344	0.388	1.877	0.325
	2	Eur.-Am.	ELA-OE	45,335	0.338	4.057	0.159
		Non-EA	ELA-OE	30,680	0.352	3.871	0.189
	3	Eur.-Am.	ELA-OE	58,594	0.427	3.398	0.141
		Non-EA	ELA-OE	15,639	0.418	3.751	0.156
	4	Eur.-Am.	ELA-OE	12,427	0.431	4.501	0.081
		Non-EA	ELA-OE	2,607	0.436	4.640	0.105

Table 8

Regression of Multiple Choice English Language Arts (Reading) Totals onto Multiple Choice Mathematics and Rubric-Scored English Language Arts, by Grade and Ethnicity (European American or Non-European American), 1999-2000 Administrations

Grade Level	Math Level	Ethnicity	Independent Variable	Number	Slope	Intercept	R-Square
8	1	Eur.-Am.	Math-MC	14,366	0.253	7.227	0.111
		Non-EA	Math-MC	34,344	0.277	6.403	0.143
	2	Eur.-Am.	Math-MC	45,335	0.188	13.532	0.044
		Non-EA	Math-MC	30,680	0.186	13.079	0.048
	3	Eur.-Am.	Math-MC	58,593	0.232	17.219	0.053
		Non-EA	Math-MC	15,639	0.216	17.269	0.052
	4	Eur.-Am.	Math-MC	12,427	0.086	23.642	0.012
		Non-EA	Math-MC	2,607	0.117	22.982	0.029
8	1	Eur.-Am.	Math-OE	14,366	0.290	3.262	0.095
		Non-EA	Math-OE	34,344	0.382	1.315	0.172
	2	Eur.-Am.	Math-OE	45,335	0.316	12.574	0.050
		Non-EA	Math-OE	30,680	0.298	12.162	0.049
	3	Eur.-Am.	Math-OE	58,593	0.504	19.120	0.068
		Non-EA	Math-OE	15,639	0.395	21.194	0.050
	4	Eur.-Am.	Math-OE	12,427	0.161	34.679	0.013
		Non-EA	Math-OE	2,607	0.167	34,759	0.018

The reader will note that a certain degree of imprecision is attributable to the greater restriction in scoring range of the rubric-scored ELA totals as compared to mathematics multiple-choice totals. The smaller sample of these questions also restricts their reliability (compare the reliabilities given in Tables 2-5, for example). This is reflected in the somewhat lower correlations.

The general linear models for the residuals of the two grade levels are given in Appendix A. There were significant effects in both grades four and eight for:

1. group membership (European American compared to others),
2. proficiency level of mathematics proficiency,
3. the interaction of these two variables,
4. the type of residual (rubric-scored ELA compared to multiple choice mathematics),
5. the interaction of type of residual and group membership,
6. the interaction of type of residual and level of mathematics proficiency.
7. the interaction of type of residual, level of mathematics proficiency, and group membership.

Tables 9 and 10 show the standard errors of estimates. The larger the standard error, the more independent that total is of the multiple-choice reading total. The largest standard errors of estimate for both groups in both grades is for mathematics rubric-scored questions, indicating that this measure is least precisely predicted by multiple-choice reading. Most interesting, however, is the great disparity in the standard errors of estimate for the two mathematics components between the students in the lowest levels of mathematics proficiency and those in levels 3 and 4, especially for the European American students. These analyses suggest that reading skills employed by the European American students to score higher in mathematics are insufficient to achieve level 3, but are used by the students at or above level 3 to achieve higher scores. This is true of all students, but the differences are most dramatic for the European American students.

Table 9

Standard Errors of Estimate for
Projected Rubric-Scored ELA, Mathematics Multiple Choice,
and Rubric-Scored MathematicsTotals
Grade Four, 1999-2000
by Ethnicity

		European Americans	Non-European Americans	Both
Level 1	ELA - OE	1.99	1.93	1.94
	Math - MC	3.20	3.15	3.16
	Math - OE	4.01	4.03	4.03
Level 2	ELA - OE	1.92	1.93	1.93
	Math - MC	2.78	2.90	2.85
	Math - OE	3.47	3.62	3.57
Level 3	ELA - OE	1.91	1.94	1.92
	Math - MC	1.97	2.03	1.99
	Math - OE	3.26	3.28	3.27
Level 4	ELA - OE	1.71	1.77	1.72
	Math - MC	1.20	1.22	1.20
	Math - OE	1.99	2.02	2.00
All	ELA - OE	1.86	1.92	1.89
	Math - MC	2.01	2.52	2.24
	Math - OE	3.04	3.44	3.22

Table 10

Standard Errors of Estimate for Projected Rubric-Scored ELA, Mathematics Multiple Choice, and Rubric-Scored Mathematics Totals Grade Eight, 1999-2000 by Ethnicity

		European Americans	Non-European Americans	Both
Level 1	ELA - OE	2.42	2.39	2.40
	Math - MC	2.89	2.90	2.90
	Math - OE	3.62	3.59	3.60
Level 2	ELA - OE	2.32	2.38	2.34
	Math - MC	2.63	2.69	2.65
	Math - OE	4.12	4.26	4.18
Level 3	ELA - OE	2.30	2.31	2.30
	Math - MC	2.14	2.20	2.15
	Math - OE	4.07	4.12	4.08
Level 4	ELA - OE	2.01	1.99	2.01
	Math - MC	1.08	1.05	1.07
	Math - OE	1.93	1.94	1.93
All	ELA - OE	2.29	2.36	2.32
	Math - MC	2.34	2.66	2.47
	Math - OE	3.88	3.91	3.90

Whole Test Analyses of the Regents Examinations, 1999

The CEE results were divided into four performance levels based on scale scores: 0-51 (not passing), 55-64 (local passing, eligible for a local diploma in some school district), 65-84 (passing), and 85-100 (passing with distinction). General Linear Model regressions examined the scores on the four mathematics Regents for students within each of those four CEC categories. The results are summarized in Table 11.

Note that the mean CEE scale score of students achieving 65 to 84 scale score in M-A was 59.42. Evidently a high level of English skills is associated with passing the mathematics examination, but the average student, even in this high-skilled sample, passes M-A with a lower passing performance in CEE.

It should be noted as well that the population that volunteered data is more highly skilled in mathematics than the general population. For example, the Department Review of a random sample of the June test-taking population (n=386 for M-A and n=488 for CEE), shows that the average scale score is 58.38 (std. = 17.29) for M-A and 67.94 (std. = 10.86) for CEE. These compare with the sample statistics for the study group of 68.64 (std. = 13.86) for M-A and 78.92 (std. = 16.55) for CEE ($\underline{t}$ (df=437) = 4.14, p<.001, for M-A, and (df=539)=0.42, ns for CEE). Post hoc analyses, using quantitative contrasts for M-A further elucidated the relationship between CEE and mathematics performance. The mean CEE scale scores in each of the four categories, 0-54, 55-64, 65-84, and 85-100 were: 47.33, 61.75, 70.58, and 94.35, respectively.

Table 11

General Linear Regression Results, by Type of Mathematics Regents Examination, of Scoring Levels on the Comprehensive Examination in English, 1999

	Mean Mathematics Scores					
Mathematics Examination	**CEE: 0-54**	**CEE: 55-64**	**CEE: 65-84**	**CEE: 85+**	**F-Ratio**	**df**
M-1	39.14	59.23	71.08	84.87	4796.17	3, 126
M-2	54.00	57.71	61.14	63.31	0.76	3, 60
M-3	-	42.50	69.31	78.28	14.10	3, 114
M-A	54.00	56.08	59.42	80.39	40.93	3, 49

Using these mean CEE values as spacing functions, quantitative contrasts revealed a significant linear ($F(df=1, 49)=117.12$, $p<.001$) component. The quadratic component ($F(df=1, 49)=1.20$, ns) was not significant. Using 54, 64, 84, and 100 as spacing functions, the linear, quadratic, and cubic components were $F(df=1, 49)=104.21$, $p<.001$; $F(df=1, 40)=0.29$, ns; and $F(df=1, 49)=18.30$, $p<.001$; respectively.

Multitrait-Multimethod Analysis of the Regents Examinations

Table 12 shows the multitrait-multimethod analysis of the M-2 and M-3 with CEE. Item level data were not available for populations that took both CEE and M-A. The correlations involving M-2 and M-3 results are attenuated, so the reader is advised to evaluate them with caution.

The analysis shows that the validity correlations (correlations of measures of the same traits) within the mathematics tests are higher than all other correlations. For M-3, the second highest correlation is the CEE validity correlations this CEE validity correlation is not as high for students who took M-2 as the correlation of the mathematics and CEE rubric-scored questions. Again, results from the small, restricted samples must be viewed with caution. When the validity coefficients were computed based on the 488 students who were scored on the CEE as part of the Department Review, the CEE correlation was .481, considerably higher than the .358 CEE validity correlation (multiple choice to rubric-scored) for the Course 2 sample that took M-2. Nevertheless, taken as a whole, even with these restricted samples, these results provide considerable evidence of the convergent and discriminant properties of the mathematics and CEE examinations.

Table 12

Regents Attenuated Correlation Coefficients of Course 2 or Course 3 Mathematics with the Comprehensive Examination in English, Open-Ended (O.E.) and Multiple Choice (M.C.) Totals, June 1999

	Validity		Within Types		Eng. O.E.	Eng. M.C.
Math Course	**Eng.**	**Math**	**O.E.**	**M.C.**	**Math M.C.**	**Math O.E.**
Course 2 (M-2)	.358	.626	.544	.228	.310	.037
Course 3 (M-3)	.441	.647	.101	.233	.283	.293

Within Examination Analyses of CEE

Analyses were performed on the internal structure of the CEE and M-A. For CEE, the structure of the examination was studied in two phases. First, the examination was divided into seven components dependent on four stimuli, as follows:

1. A listening passage with
   (a) six associated multiple choice questions and
   (b) one open-ended (rubric-scored) question;
2. Two reading passages each associated with
   (a) ten multiple choice questions and
   (b) one open-ended (rubric-scored)question; and
3. One reading passage associated with an open-ended (rubric-scored) question.

Each multiple choice question is worth one point maximum and each rubric-scored question is worth six points maximum.

Because the rubric-scored questions involve either listening and writing or reading and writing, they are not pure measures of writing. The multiple choice listening and reading questions involve less contamination of those traits.

As a consequence, three types of relationships were hypothesized as follows:

1. Weak: Listening multiple choice (questions 1-6) and reading multiple choice (questions 7-16 and 17-26) listing multiple choice and rubric-scored reading-then-writing (rubric-scored questions 2-4), reading multiple choice (questions 7-16 and 17-26) and rubric-scored listening.
2. Partial: Listening multiple choice and rubric-scored listening-then-writing (question 1), reading multiple choice and rubric-scored reading-then-writing, rubric-scored listening then writing and rubric-scored reading-then-writing.
3. Strong: Reading multiple choice across sections (questions 7-16 and 17-26), rubric-scored reading-then-writing (questions 2 through 4).

The reader will note that item type is confounded with the listening, reading , and writing traits. More importantly, the previous factor analytic work suggests that this is a unidimensional examination. Nevertheless, as the test samples the New York State Learning Standards, it is expected that the pattern of correlations overall exhibit convergent and discriminant properties in relation to the separability and the dimensionality of these standards. Most important to these characteristics are the hierarchy of cognitive linguistic demands. For example, multiple choice listening questions require retrieval based on matching the salient features of the stimulus and the questions. In contrast, rubric-scored reading-the-writing questions require recall without match of the features and then integration of skills for writing production.

To demonstrate this, cognitive hierarchy the correlations were converted to z-scores to provide a proper scale for analysis and two General Linear Regression Models were performed:

1. Within each administration (June 1999, April 2000, June 2000) estimating degree of relationship (weak, partial, or strong) as a main effect;
2. Across administrations estimating degree of relationship and administration as main effects and the interaction of these two.

The correlation matrices for the three administrations are given in Table 13 . Note that the correlation coefficients, sometimes depending on one open-ended question, are not corrected for attenuation. Table 14 shows the results of the regression analyses, and the mean correlation coefficients converted back from z-score means to correlation coefficients. The analyses showed significant differences related to degree of relationship for June 1999 ($F(df=2,39)=12.31$, $p<.001$), April 2000 ($F(df=2,39)=6.58$, $p<.01$), and June 2000 ($F(df=2.39)=8.92$, $p<.001$), respectively.

In the one analysis for all three administrations there were significant main effects for degree of relationship ($F(df=2, 117)=27.02$, $p<.0001$) and for administration $F(df=2,117)=9.18$, $p<.001$), but not for the interaction of the two ($F(df=4,117)=0.14$,ns).

Post hoc contrasts revealed that the April 2000 and June 2000 administrations yielded higher overall correlations than the June 1999 correlation. This is reasonable in view of the smaller sample size and the greater heterogeneity of the June 1999 administration, which was the first for the CEE.

Table 13

Interrelationships of Regents Comprehensive Examination in English Totals, June 1999, April 2000 and June 2000

June 1999	Means	S.D.	Listening M.C. List	Listening Write	Read M.C. 1	Read M.C. 2	Read/Write O.E. 2	Read/Write O.E. 3	Read/Write O.E. 4
Listening									
M.C.	4.55	1.26	-						
then Writing	3.53	1.08	.307	-					
Reading									
M.C. 1	7.36	1.60	.260	.315	-				
M.C. 2	7.14	1.70	.306	.366	.403	-			
Reading then Writing									
O.E. 2	3.31	0.98	.208	.561	.308	.366	-		
O.E. 3	3.36	1.01	.190	.546	.306	.334	.528	-	
O.E. 4	3.18	1.09	.184	.506	.196	.317	.447	.543	-
April 2000									
Listening									
M.C.	3.71	1.42	-						
then Writing	2.63	1.00	.454	-					
Reading									
M.C. 1	6.80	2.24	.486	.484	-				
M.C. 2	6.41	2.33	.401	.425	.441	-			

Reading then Writing									
O.E. 2	2.13	0.99	.334	.664	.431	.367	-		
O.E. 3	2.63	0.99	.320	.502	.371	.536	.495	-	
O.E. 4	2.66	1.07	.291	.482	.346	.346	.487	.718	-
			Listening		Read		Read/Write		
June 2000	Means	S.D.	M.C. List	Write	M.C. 1	M.C. 2	O.E. 2	O.E. 3	O.E. 4
Listening									
M.C.	5.09	1.05	-						
then Writing	3.89	0.98	.354	-					
Reading									
M.C. 1	8.62	1.62	-	-					
M.C. 2	7.90	1.54	.357	.386	-				
Reading then Writing									
O.E. 2	3.33	1.05	.283	.560	.401	.309	-		
O.E. 3	3.60	1.00	.307	.640	.398	.357	.539	-	
O.E. 4	3.59	1.15	.308	.643	.429	.368	.555	.660	-

Table 14

Mean Correlation Coefficients for Strongly Related, Partially Related, and Weakly Related Sections of the Regents Comprehensive Examination in English June 1999, April 2000, and June 2000 (same grouping Roman numeral indicates not different at p<.05)

Administration Date	Degree of Relationship	Correlation	Grouping
June 1999	Weak	.264	I
	Partial	.375	II
	Strong	.480	III
	All	.360	
April 2000	Weak	.393	I
	Partial	.561	I, II
	Strong	.722	II
	All	.461	
June 2000	Weak	.353	I
	Partial	.457	I, II
	Strong	.452	II
	All	.441	
All	Weak	.338	I
	Partial	.369	I
	Strong	.544	II
	All	.422	

Within Examination Analyses of Mathematics A

Like the Regents CEE, the Regents Mathematics A examination has several different item types that are confounded with content. That is, there are slight intended differences in the cognitive and content focus of the measures that are related to item types. The multiple-choice questions are dichotomously scored, either zero for an incorrect answer or two for a correct answer. In all, there are 20 of these items. There are also four types of polytomous, or rubric-scored items: five ranging from zero to two, five ranging from zero to three, and five ranging from zero to four.

Again, the Department Review of the Regents examinations from June 1999 (n=386) and June 2000 (n=1,284) were the sources of the data analyzed to describe the within test structure. Because we expected the Mathematics A test to be unidimensional (c.f. AES, 1999) the content of the test was considered, similar to the conceptual design of the CEE, as having components of greater cognitive relationship to each other and of weaker relationships.

A structure for characterizing the inter-component relationships was determined by reference to content analyses performed by the State Education Department's Office of Curriculum and Instruction to identify test units for the development of the component retesting program. In all, the New York State Learning Standards specify seven key ideas for mathematics: Mathematical Reasoning, Numbers and Numeration, Operations, Modeling/Multiple Representation, Measurement, Uncertainty, and Patterns and Functions. Of these seven, the content analysis identified the first three as prerequisites for each of the final four. The final four, then, are more distinct in terms of specialized skills and knowledge, while the first three are more diffused as they are shared to some extent in each of the final four. Cognitively, a modicum of achievement of the first three is prerequisite to achievement of the final four. While the more basic skills, or the better-developed problem solving skills can be applied to the first three key ideas successfully, only the better-developed skills can be successfully applied to the last four.

Correlations were performed on units defined within each of the seven key ideas and delineated as multiple choice, 2-point rubric-scored, 3-point rubric-scored, and 4-point rubric-scored. In all, there were 18 distinct configurations of items yielding 153 bivariate correlation coefficients in 1999 and 21 distinct configurations of items yielding 171 correlation coefficients in 2000.

The correlations were classified as basic/basic (between configurations that each measure one of the first three key ideas), basic/distinct (between configurations that measure one of the first three and one of the last four key ideas), and distinct/distinct (between configurations that each measure one of the last four key ideas). This formulation would lead to the prediction that the basic/basic correlation coefficients should be lowest and the distinct/distinct correlation coefficients should be highest because the latter content areas demand application of a narrower range of higher order cognitive skills.

The correlation coefficients were transformed to z-scores, and a General Linear Model was computed in which the transformed correlations were the dependent variable and the independent variables were year of testing (1999 or 2000), agreement (termed "validity") vs. non-agreement in the key idea of the component, type of relationship (basic, mixed or distinct), and item type (both components are multiple choice, one is multiple choice and the other is rubric-scored, or both components are rubric-scored).

A summary table of the regression is given in Appendix B. There were significant main effects for item type and type of relationship. There were also significant interaction effects related to year by type of relationship and year by type of relationship by validity. Table 14 provides the mean values of the correlations converted back from z-scores to correlation coefficients. Post hoc Tukey comparisons show that the mean correlation coefficients among open-ended components (.353) were larger than those among different item types (.292) or among multiple choice components (.269).

Post hoc Tukey comparisons also revealed that, as hypothesized, the relationships involving the four distinct content areas (mean=.348) were higher than either those involving the distinct and basic components (mean=.283) or those involving the basic components with other basic components (mean=.239). The interaction effect showed, that while the ordering of these relationships (distinct highest, mixed next, basic lowest) was consistent both in 1999 and in 2000, the differences among the three types of relationships were much greater in 1999 than they were in 2000. This might suggest the more specific application of skills in Mathematics A related to greater development of the Mathematics A curriculum over the past year.

Table 15

Mean Correlation Coefficients for Interrelationships among Sections and Item Types on the Mathematics A Administrations, June 1999 and June 2000 (Computed on z-scores and converted back)

Type of Relationship	Item Types	Same Content 1999	Same Content 2000	Same Content Both	Different Content 1999	Different Content 2000	Different Content Both	Both 1999	Both 2000	Both Both
basic/basic	oe/oe	-----	.390	.390	.230	.312	.298	.230	.325	.312
	oe/mc	.183	.263	.237	.201	.216	.212	.195	.233	.220
	mc/mc	-----	-----	-----	.257	.153	.206	.257	.153	.206
	all	.183	.289	.260	.226	.236	.232	.217	.249	.238
basic/distinct	oe/oe	-----	-----	-----	.341	.296	.310	.341	.296	.310
	oe/mc	-----	-----	-----	.315	.235	.270	.315	.235	.270
	mc/mc	-----	-----	-----	.279	.246	.245	.262	.245	.262
	all	-----	-----	-----	.315	.261	.283	.315	.270	.283
distinct/dist.	oe/oe	.488	.295	.389	.516	.272	.386	.511	.276	.386
	oe/mc	.385	.243	.308	.377	.258	.316	.379	.254	.314
	mc/mc	-----	-----	-----	.302	.278	.290	.302	.278	.290
	all	.431	.265	.344	.438	.267	.349	.437	.294	.362
all	oe/oe	.488	.319	.396	.449	.315	.367	.454	.316	.370
	oe/mc	.350	.265	.302	.334	.248	.287	.336	.251	.289
	mc/mc	-----	-----	-----	.283	.242	.290	.283	.242	.262
	all	.406	.285	.336	.368	.278	.317	.373	.278	.319

Conclusions

This paper examines the convergent and discriminant properties of the New York State fourth-grade, eighth-grade, and commencement-level tests. In general, the evidence is very supportive of the construct validity of the tests examined. Of all the correlations presented, there were only two cases in which the relationships among similar item types exceeded the validity correlations. Both cases involved the validity correlations of the fourth-grade ELA examinations. The validity correlations within the mathematics examinations were always the highest.

The first exception is for European American students on ELA-4. It must be noted here, that the holistic rubric scoring of both the fourth- and eighth-grade ELA examinations involve evaluation of questions that are both short answer and more traditional open-ended. They are less distinct from the multiple-choice questions in response format than they are in their reference to particular stimuli, task demands, and holistic mode of scoring. These item types may present a cognitive demand that is similar to the problem solving demand of the mathematics examinations, thus raising the correlation of performance of the ELA cluster scores to performance on some mathematics items for some examinees. Follow up analyses suggest that among European American students, and among higher scoring students in general, similar skills are employed across tests to meet these increasing cognitive demands. By eighth grade, more specific skills are employed in their approach to problem solving.

The second exception is for the rubric-scored questions on the CEE and M-2. This analysis drew on small populations, so that, again, the reader is asked to review these data with caution. It is particularly important to note that there was no systematic relationship between level of CEE performance and performance on M-2. Within test analyses were also performed on the structure of both the CEE and M-A. Hypothesis were drawn about which components should be more strongly related and less well-related to each other, based on the cognitive demands related to content and item type. These hypothesis were supported.

To a large extent the results of these analyses clearly support the construct properties of the examinations. More data need to be gathered to continue these analyses, and several projects are underway to address this need.

References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. *Standards for Educational and Psychological Testing.* (Washington, D.C.: Author, 1999).

Assessment and Evaluation Services, *Item Analysis and Calibration of the New York English Language Arts and Mathematics A Regents Examinations.* (Albany, New York: Author, 1999).

Campbell, D.T., & Fiske, D.W., Covergent and Discriminant Validation be the Multitrait-Multimethod Matrix. Psychological Bulletin, 56, pp. 81-105.

Lord, F.M. & Novick, M.R., *Statistical Theories of Mental Test Scores*. (Reading, Mass.: Addison-Wesley Publishing Co., 1968).

Myers, J.L., *Fundamentals of Experimental Design (Second Edition)*. (Boston, Mass.: Allyn and Bacon, Inc., 1972.

Thorndike, R.L. & Hagen, E., *Measurement and Evaluation in Psychology and Education (Third Edition).* (New York: John Wiley & Sons, Inc., 1969).

Appendix A

General Linear Regression Analyses

Observed Minus Predicted Totals Squared

on Rubric-Scored (OE) ELA Questions

and Mathematics Multiple Choice and Rubric-Scored Questions

Using Multiple Choice ELA Questions

as the Predictor, by Grade and Ethnicity,

1999-2000 Administrations

Grade Level	Variable	Sum of Squares	Degrees of Freedom	Mean Square	F-Ratio
4	a) Ethnicity	2,372.82	1	2,372.82	22.13***
	b) Math Level	1,836,758.93	3	612,252.98	17,133.60***
	a by b	6,634.94	3	2,211.65	20.63***
	error between	22,230,827.11	207,373	3.93	
	total between	24,076,593.79	207,380		
	c) type of est.	2,535,865.25	2	1,267,932.63	12,405.13***
	a by c	1,121.22	2	560.61	5.48***
	b by c	930,707.14	6	155,117.86	1,517.63***
	a by b by c	5,215.03	6	869.17	8.50***
	error within	42,391,348.44	414,746	102.21	
	total within	45,864,257.09	414,762		
	Total	69,940,850.88	622,142		

Appendix A

General Linear Regression Analyses

Observed Minus Predicted Totals Squared

on Rubric-Scored (OE) ELA Questions

and Mathematics Multiple Choice and Rubric-Scored Questions,

Using Multiple Choice ELA Questions

as the Predictor, by Grade and Ethnicity,

1999-2000 Administrations

Grade Level	Variable	Sum of Squares	Degrees of Freedom	Mean Square	F-Ratio
8	a) Ethnicity	2,465.85	1	2,465.85	13.92***
	b) Math Level	1,176,752.96	3	392,250.99	6,644.33***
	a by b	10,124.51	3	3,374.84	19.06***
	error between	37,899,321.54	213,992	177.11	
	total between	39,088,664.87	213,999		
	c) type of est.	3,458,911.42	2	1,729,455.71	9,834.44***
	a by c	1,309.20	2	654.60	3.72*
	b by c	1,256,212.33	6	209,368.72	1,190.56***
	a by b by c	5,512.55	6	918.76	5.22***
	error within	79,985,931.54	427,984	175.86	
	total within	119,074,596.41	428,000		
	Total	2,445,550.88			

***Exceeds the p<.001 level of significance.

Appendix B

General Linear Model Summary Table
Mathematics A Correlation Coefficients

Source Of Variance	Degrees of Freedom	Sum of Squares	Mean Square	F-Ratio	Pr> F
Year	1	0.07	0.07	3.85	0.06
Validity (content)	1	0.00	0.00	0.01	0.81
Item Type	2	0.15	0.08	4.30	0.02
Bases (Basic, etc.)	2	0.13	0.07	3.66	0.03
Year * Valid	1	0.00	0.00	0.26	0.61
Year * Item Type	2	0.00	0.00	0.01	0.99
Year * Bases	2	0.11	0.06	3.19	0.04
Valid * Item Type	1	0.00	0.00	0.00	0.95
Valid * Bases	1	0.00	0.00	0.24	0.63
Item Type * Bases	4	0.04	0.01	0.59	0.67
Year * Valid * Item Type	1	0.00	0.00	0.13	0.72
Year * Valid * Bases	1	0.00	0.00	0.25	0.62
Year * Item Type * Bases	4	0.17	0.04	2.41	0.05
Valid * Item Type * Bases	1	0.00	0.00	0.08	0.77
Yr * Valid * It. T. * Bases	0	0.00			

Appendix B

General Linear Model Summary Table
Mathematics A Correlation Coefficients

Source Of Variance	Degrees of Freedom	Sum of Squares	Mean Square	F-Ratio	Pr> F
Year	1	0.09	0.09	6.54	0.01
Validity (content)	1	0.00	0.00	0.12	0.73
Item Type	2	0.13	0.07	4.85	0.01
Bases (Basic, etc.)	2	0.11	0.05	4.01	0.02
Year * Valid	1	0.01	0.01	0.67	0.41
Year * Item Type	2	0.00	0.00	0.08	0.92
Year * Bases	2	0.16	0.08	5.78	0.00
Valid * Item	1	0.00	0.00	0.01	0.03
Valid * Bases	1	0.00	0.00	0.14	0.70
Item * Bases	4	0.03	0.01	0.47	0.76
Year * Valid * Item	1	0.01	0.01	0.99	0.32
Year * Valid * Bases	1	0.01	0.01	0.46	0.50
Year * Item * Bases	4	0.24	0.06	4.48	0.00
Valid * Item * Bases	1	0.00	0.00	0.00	1.00
Yr * Valid * Item * Bases	0	0.00	-----		

In the one analysis for all three administrations there were significant main effects for degree of relationship ($F(df=2, 117)=27.02$, $p<.001$) and for administration $F(df=2,117)=9.18$, $p<.001$), but not for the interaction of the two ($F(df=4,117)=0.14$, ns).

Post hoc contrasts revealed that the April 2000 and June 2000 administrations yielded higher overall correlations than the June 1999 correlation. This is reasonable in view of the smaller sample size and the greater heterogeneity of the June 1999 administration, which was the first for the CEE.

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

ERIC
Educational Resources Information Center

REPRODUCTION RELEASE

(Specific Document)

TM035013

I. DOCUMENT IDENTIFICATION:

Title: Construct Properties of New York State English Language Arts and Mathematics Examinations 1998-1999-2000	
Author(s): Gerald E. DeMauro, Ph.D.	
Corporate Source: NYS Education Dept.	Publication Date: January 2001

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2A documents	The sample sticker shown below will be affixed to all Level 2B documents
PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 1	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2A	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2B
Level 1	Level 2A	Level 2B
[X]	[ ]	[ ]
Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.	Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only	Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here, please

Signature: Gerald E. DeMauro	Printed Name/Position/Title: Gerald E. DeMauro	
Organization/Address: NYSED 89 Washington Ave., 775 EBA, Albany, NY 12234	Telephone: 518-474-5902	FAX: 518-474-1989
	E-Mail Address: gdemauro@mail.nysed.gov	Date: 6/4/03

(Over)

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

ERIC CLEARINGHOUSE ON ASSESSMENT AND EVALUATION
UNIVERSITY OF MARYLAND
1129 SHRIVER LAB
COLLEGE PARK, MD 20742-5701
ATTN: ACQUISITIONS

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
4483-A Forbes Boulevard
Lanham, Maryland 20706

Telephone: 301-552-4200
Toll Free: 800-799-3742
FAX: 301-552-4700
e-mail: ericfac@inet.ed.gov
WWW: http://ericfacility.org