

**2007**  
**Grade Eight Proficiency**  
**Assessment (GEPA)**

**TECHNICAL REPORT**  
**March Administration**



PTM#1506.61

Copyright © 2007 New Jersey  
Department of Education

All rights reserved.





<b>CHAPTER 1: INTRODUCTION</b> . . . . .	<b>1</b>
1.1 Description of the Grade Eight Proficiency Assessment (GEPA) . . . . .	1
1.2 Purpose of the GEPA . . . . .	4
1.3 GEPA Organizational Support. . . . .	5
<b>CHAPTER 2: TEST DEVELOPMENT</b> . . . . .	<b>6</b>
2.1 Test Specifications. . . . .	6
Language Arts Literacy . . . . .	7
Mathematics . . . . .	8
Science . . . . .	8
2.2 Development of Test Items . . . . .	9
2.3 Item Review Process. . . . .	12
2.4 Operational Test Development . . . . .	18
2.5 Review and Approve Operational Test Forms . . . . .	20
2.6 Test Materials for Visually Impaired Students . . . . .	20
<b>CHAPTER 3: TEST ADMINISTRATION</b> . . . . .	<b>21</b>
3.1 Participation. . . . .	21
3.2 Test Security Procedures . . . . .	22
3.3 Test Administration Procedures. . . . .	23
3.4 Test Accommodations . . . . .	24
3.5 Results for Special Education Students and Section 504 Students Tested with Accommodations or Modifications . . . . .	26
<b>CHAPTER 4: SCORING</b> . . . . .	<b>30</b>
4.1 Multiple-choice Items . . . . .	30
4.2 Open-ended Items . . . . .	30
Scorer Selection . . . . .	31
Rangefinding . . . . .	31
Development of Scoring Guides . . . . .	32
Team Leader Training and Qualifying . . . . .	32
Scorer Training and Qualifying . . . . .	33
Scoring Procedures and Paper Flow . . . . .	34
Scorer Monitoring . . . . .	35
Agreement Between Scorers for the Writing Tasks and Open-Ended Items . . . . .	35
4.3 Quality Control Procedures in Data Preparation . . . . .	37
Scanning and Scoring . . . . .	37
NJDOE Quality Control of Score Reporting . . . . .	38
<b>CHAPTER 5: STANDARD SETTING</b> . . . . .	<b>39</b>
5.1 Overview of the Process . . . . .	39
5.2 Procedures . . . . .	39
Language Arts Literacy–Proficiency Level Descriptors . . . . .	40
Mathematics–Proficiency Level Descriptors . . . . .	41
Science–Proficiency Level Descriptors . . . . .	42

<i>Judge Selection Process and Criteria</i> . . . . .	43
<i>Holistic/Paper Sorting Methodology</i> . . . . .	43
<i>Overview of the 8-Step Plan</i> . . . . .	44
<b>5.3 Results</b> . . . . .	46
<b>CHAPTER 6: SCALING AND EQUATING</b> . . . . .	48
<b>6.1 Scaling</b> . . . . .	48
<b>6.2 Equating</b> . . . . .	48
<i>Mathematics and Science Equating Design</i> . . . . .	49
<i>Language Arts Literacy Equating Design</i> . . . . .	50
<b>CHAPTER 7: TEST STATISTICS</b> . . . . .	52
<b>7.1 Reliability of the Test Scores</b> . . . . .	52
<b>CHAPTER 8: ITEM-LEVEL STATISTICS</b> . . . . .	55
<b>8.1 Classical Item Statistics</b> . . . . .	55
<b>8.2 Speededness</b> . . . . .	58
<b>8.3 Intercorrelations</b> . . . . .	60
<b>CHAPTER 9: TEST VALIDITY</b> . . . . .	62
<b>9.1 Content and Curricular Validity (Evidence Based on Test Content)</b> . . . . .	64
<b>9.2 Construct Validity (Evidence Based on Response and Evidence Based on Internal Structure)</b> . . . . .	65
<b>9.3 Criterion-Related Validity (Evidence Based on Relations to Other Variables)</b> . . . . .	66
<b>9.4 Consequential Validity Evidence (Evidence Based on Consequences of Testing)</b> . . . . .	69
<b>CHAPTER 10: REPORTING</b> . . . . .	72
<b>10.1 Information on the Reports</b> . . . . .	72
<b>10.2 Types of Reports</b> . . . . .	74
<i>Cycle I Reports</i> . . . . .	74
<i>Cycle II Reports</i> . . . . .	79
<b>Appendix A: Scoring Rubrics and 3rd Reader Score Calculation Charts</b> . . . . .	83
<b>Appendix B: Cycle II Test Results</b> . . . . .	94
<b>Appendix C: Modifications of Test Administration Procedures for Special Education Students and Students Eligible Under Section 504 of the Rehabilitation Act of 1973</b> . . . . .	114
<b>Appendix D: Raw to Scale Scores Conversions</b> . . . . .	119
<b>Appendix E: Scale Scores with Frequencies</b> . . . . .	124
<b>Appendix F: Reporting Limited English Proficiency (LEP) and Special Education (SE)</b> . . . . .	129
<b>References</b> . . . . .	132

## Tables

<b>TABLE 1.1</b>	<b>Number of Items and Approximate Times</b>	<b>1</b>
<b>TABLE 1.2</b>	<b>Total Student Group Testing in 2007</b>	<b>2</b>
<b>TABLE 2.1</b>	<b>Total Points Possible for the Language Arts Literacy Component of the GEPA</b>	<b>10</b>
<b>TABLE 2.2</b>	<b>Total Points Possible for the Mathematics Component of the GEPA</b>	<b>10</b>
<b>TABLE 2.3</b>	<b>Total Points Possible for the Science Component of the GEPA</b>	<b>10</b>
<b>TABLE 2.4</b>	<b>District Factor Groups (DFG) Represented on the GEPA Content and Sensitivity Committees</b>	<b>12</b>
<b>TABLE 2.5</b>	<b>Number of Items Field Tested in 2007</b>	<b>13</b>
<b>TABLE 2.6</b>	<b>GEPA 2007 Statistical Committee Meetings</b>	<b>13</b>
<b>TABLE 2.7</b>	<b>LANGUAGE ARTS LITERACY - READING – Number of Field Test Items Approved During Statistical Review</b>	<b>17</b>
<b>TABLE 2.8</b>	<b>LANGUAGE ARTS LITERACY - WRITING – Number of Field Test Items Approved During Statistical Review</b>	<b>17</b>
<b>TABLE 2.9</b>	<b>MATHEMATICS – Number of Field Test Items Approved During Statistical Review</b>	<b>17</b>
<b>TABLE 2.10</b>	<b>SCIENCE – Number of Field Test Items Approved During Statistical Review</b>	<b>18</b>
<b>TABLE 2.11</b>	<b>Operational Test Specifications</b>	<b>19</b>
<b>TABLE 3.1</b>	<b>Setting Accommodations and Modifications for Special Education Students and Section 504 Students</b>	<b>27</b>
<b>TABLE 3.2</b>	<b>Setting Accommodations for Special Education Students and Section 504 Students</b>	<b>27</b>
<b>TABLE 3.3</b>	<b>Scheduling Accommodations for Special Education Students and Section 504 Students</b>	<b>28</b>
<b>TABLE 3.4</b>	<b>Test Materials Modifications for Special Education Students and Section 504 Students</b>	<b>28</b>
<b>TABLE 3.5</b>	<b>Test Procedures Modifications for Special Education Students and Section 504 Students</b>	<b>29</b>
<b>TABLE 4.1</b>	<b>Number of Writing Prompts and Open-ended Items Scored</b>	<b>30</b>
<b>TABLE 4.2</b>	<b>Consistency Between Raters Scoring GEPA Writing Tasks and Open-Ended Items - March 2007</b>	<b>36</b>
<b>TABLE 5.1</b>	<b>Proficiency-Level Cut Scores</b>	<b>47</b>
<b>TABLE 5.2</b>	<b>Percentage of Students Achieving Each Performance Level</b>	<b>47</b>
<b>TABLE 6.1</b>	<b>Summary of Equating Statistics</b>	<b>51</b>
<b>TABLE 6.2</b>	<b>Cut Scores and Associated Thetas for Proficiency Levels</b>	<b>51</b>

<b>TABLE 7.1 Reliability Estimates and Standard Errors of Measurement (SEMs) for Content Areas and Clusters - 2007</b> . . . . .	<b>53</b>
<b>TABLE 7.2 Reliability Estimates and Standard Errors of Measurement (SEMs) for Dichotomously Scored Items Within Content Clusters - 2007</b> . . . . .	<b>54</b>
<b>TABLE 8.1 Item Difficulty and Discrimination Summary Statistics for Dichotomously Scored and Open-Ended Items by Test Section and Cluster - 2007</b> . . . . .	<b>55</b>
<b>TABLE 8.2 Frequency Distributions of Item Difficulty and Item Discrimination by Content Cluster - 2007 Language Arts Literacy</b> . . . . .	<b>56</b>
<b>TABLE 8.3 Frequency Distributions of Item Difficulty and Item Discrimination by Content Cluster - 2007 Mathematics</b> . . . . .	<b>57</b>
<b>TABLE 8.4 Frequency Distributions of Item Difficulty and Item Discrimination by Content Cluster - 2007 Science</b> . . . . .	<b>58</b>
<b>TABLE 8.5 Percentage of Students Omitting the Last Items of Each Test Part - 2007</b> . . . . .	<b>59</b>
<b>TABLE 8.6 Intercorrelations Among Major Content Clusters and Item Types - 2007</b> . .	<b>60</b>
<b>TABLE 8.7 Intercorrelations Among Content Areas and Clusters - 2007</b> . . . . .	<b>61</b>
<b>TABLE 9.1 LANGUAGE ARTS LITERACY - Percentages of Students Across GEPA/HSPA Proficiency Levels</b> . . . . .	<b>67</b>
<b>TABLE 9.2 MATHEMATICS - Percentages of Students Across GEPA/HSPA Proficiency Levels</b> . . . . .	<b>67</b>
<b>TABLE 9.3 NAEP Results for New Jersey Grade 8 Students</b> . . . . .	<b>68</b>

## Figures

<b>FIGURE 2.1: GEPA Test Development Process</b> . . . . .	<b>7</b>
<b>FIGURE 2.2: Item Approval Before Field Test</b> . . . . .	<b>12</b>
<b>FIGURE 2.3: Item Approval Before Operational Test</b> . . . . .	<b>16</b>
<b>FIGURE 9.1 LANGUAGE ARTS LITERACY - Longitudinal Graph by Economic Status</b> . .	<b>70</b>
<b>FIGURE 9.2 MATHEMATICS - Longitudinal Graph by Economic Status</b> . . . . .	<b>71</b>
<b>FIGURE 9.3 SCIENCE - Longitudinal Graph by Economic Status</b> . . . . .	<b>71</b>
<b>FIGURE 10.1: Individual Student Report (ISR Front)</b> . . . . .	<b>75</b>
<b>FIGURE 10.2: Individual Student Report (ISR Back)</b> . . . . .	<b>76</b>
<b>FIGURE 10.3: Student Roster (Mathematics)</b> . . . . .	<b>78</b>
<b>FIGURE 10.4: Cluster Means Report</b> . . . . .	<b>80</b>
<b>FIGURE 10.5: Performance by Demographic Groups</b> . . . . .	<b>82</b>



## CHAPTER 1: INTRODUCTION

### 1.1 Description of the Grade Eight Proficiency Assessment (GEPA)

The New Jersey Grade Eight Proficiency Assessment (GEPA) for the 2007 administration consisted of three content area tests – Language Arts Literacy, Mathematics, and Science. The GEPA is designed to provide an indication of the progress students are making in mastering the knowledge and skills described in New Jersey’s Core Curriculum Content Standards for these content areas.

The GEPA was administered between Monday, March 12 and Thursday, March 15, 2007, with make-up testing between Monday, March 19 and Thursday, March 22, 2007. March 2007 marked the ninth administration of the GEPA, which provides valuable information about student progress toward mastery of the skills required for high school graduation. Table 1.1 lists the number of test items and approximate testing times for the three content areas.

**TABLE 1.1**  
**Number of Items and Approximate Times**

<b>Content Areas</b>	<b>Items</b>	<b>Approximate Times</b>
Science	60 multiple-choice 4 open-ended (including field-test items)	1 hour, 57 minutes
Mathematics	40 multiple-choice 8 open-ended (including field-test items)	2 hours, 27 minutes
Language Arts Literacy (Days 1 and 2)	20 multiple-choice 4 open-ended 2 writing tasks (speculate/ picture persuade) Field-test component	2 hours, 12 minutes (per day)

The GEPA Language Arts Literacy measures both reading and writing. The Reading component requires students to read passages and to respond to related items. The passages are selected from published books, newspapers, and magazines, as well as everyday text. The Reading component includes both multiple-choice and open-ended items. The open-ended items require students to write a few sentences or a few paragraphs to answer a question about the text. The Writing component asks students to write two essays. All the tasks in the Writing component require students to write a response.

The GEPA Mathematics measures students’ abilities to solve problems using mathematical concepts. The components in this content area measure: Number and Numerical Operations; Geometry and Measurement; Patterns and Algebra; and Data Analysis, Probability, and Discrete Mathematics. Mathematics, like the Reading component of Language Arts Literacy, contains both multiple-choice and open-ended items. The open-ended items require students to solve a problem as well as explain their solution.

The GEPA Science measures students’ knowledge in Life Science, Physical Science, and Earth Science; and skill in Knowledge and Application. The Science content area contains both multiple-choice and open-ended items. The open-ended items require students to respond to a question as well as explain the answer.

Rubrics for scoring the GEPA open-ended items and writing prompts are included in Appendix A of this *Technical Report*.

Table 1.2 presents the statewide test results for the 2007 administration of the GEPA. This table shows the number and percentages of students in each of the Proficiency Levels – Partially Proficient, Proficient, and Advanced Proficient. The first column in Table 1.2 shows the total 108,474 enrolled students including 87,396 general education students, 18,197 special education students, and 2,999 limited English students. “General Education” excludes students coded as special education (SE) or limited English proficient (LEP) on their answer folders. “Special Education” includes students coded as SE. “Limited English Proficient” includes students coded as LEP. “Total Students” refers to all students tested (general education, special education, and current LEP students).

Following the Number Enrolled column are the columns for Number of APA Students, Number Not Present, and Number of Voids. Number enrolled represents total number of answer folders returned. The number of APA (Alternate Proficiency Assessment) students shows the number of answer folders marked for students taking the APA rather than GEPA for each content area.

**TABLE 1.2**  
**Total Student Group Testing in 2007**

TESTS	NUMBER OF STUDENTS ENROLLED	NUMBER OF APA STUDENTS	NUMBER NOT PRESENT	NUMBER OF VOIDS	NUMBER OF VALID SCALE SCORES	PROFICIENCY LEVELS						MEAN SCALE SCORE
						PARTIALLY PROFICIENT (100 – 199)		PROFICIENT (200 – 249)		ADVANCED PROFICIENT (250 – 300)		
						NO.	%	NO.	%	NO.	%	
<b>LANGUAGE ARTS LITERACY</b>												
General Education	87,396	0	311	466	86,619	14,672	16.9	60,218	69.5	11,729	13.5	221.8
Special Education	18,197	680	229	317	16,971	11,390	67.1	5,375	31.7	206	1.2	185.1
LEP Current & Former	4,020	6	132	486	3,396	2,467	72.6	900	26.5	29	0.9	180.6
LEP Current	2,999	3	129	483	2,384	1,939	81.3	436	18.3	9	0.4	173.4
LEP Former	1,021	3	3	3	1,012	528	52.2	464	45.8	20	2.0	197.6
<b>Total Students<sup>a</sup></b>	<b>108,474</b>	<b>680</b>	<b>667</b>	<b>1,262</b>	<b>105,865</b>	<b>27,901</b>	<b>26.4</b>	<b>66,020</b>	<b>62.4</b>	<b>11,944</b>	<b>11.3</b>	<b>214.9</b>
<b>MATHEMATICS</b>												
General Education	87,396	0	385	71	86,940	19,454	22.4	44,129	50.8	23,357	26.9	222.5
Special Education	18,197	681	250	76	17,190	12,234	71.2	4,339	25.2	617	3.6	185.2
LEP Current & Former	4,020	5	32	8	3,975	2,744	69.0	1,005	25.3	226	5.7	187.5
LEP Current	2,999	3	27	7	2,962	2,203	74.4	612	20.7	147	5.0	183.8
LEP Former	1,021	2	5	1	1,013	541	53.4	393	38.8	79	7.8	198.4
<b>Total Students<sup>b</sup></b>	<b>108,474</b>	<b>681</b>	<b>661</b>	<b>152</b>	<b>106,980</b>	<b>33,790</b>	<b>31.6</b>	<b>49,069</b>	<b>45.9</b>	<b>24,121</b>	<b>22.5</b>	<b>215.5</b>
<b>SCIENCE</b>												
General Education	87,396	0	430	71	86,895	12,023	13.8	49,619	57.1	25,253	29.1	228.9
Special Education	18,197	645	306	73	17,173	8,502	49.5	7,630	44.4	1,041	6.1	200.7
LEP Current & Former	4,020	4	41	5	3,970	2,626	66.1	1,263	31.8	81	2.0	190.9
LEP Current	2,999	2	34	5	2,958	2,143	72.4	774	26.2	41	1.4	187.2
LEP Former	1,021	2	7	0	1,012	483	47.7	489	48.3	40	4.0	201.7
<b>Total Students<sup>c</sup></b>	<b>108,474</b>	<b>645</b>	<b>767</b>	<b>149</b>	<b>106,913</b>	<b>22,576</b>	<b>21.1</b>	<b>58,003</b>	<b>54.3</b>	<b>26,334</b>	<b>24.6</b>	<b>223.3</b>

<sup>a</sup> The number of Valid Scale Scores includes 109 students who are both Special Education and Limited English Proficient.

<sup>b</sup> The number of Valid Scale Scores includes 112 students who are both Special Education and Limited English Proficient.

<sup>c</sup> The number of Valid Scale Scores includes 113 students who are both Special Education and Limited English Proficient.

Number not present indicates the number of answer folders returned that were totally blank excluding answer folders coded as APA. A student's answer folder can be voided at the time of testing due to illness, cheating or disruptive behavior, or some other reason. If a student's answer folder is voided, no total test score for that student is reported for the content area. A void code is printed in place of the total test score on the student's individual reports.

During the scoring process, a void code is given if a student's answer folder showed less than 20 percent of the items were attempted on the Mathematics or Science content area tests. During the 2007 administration, 152 Mathematics and 149 Science tests were voided due to the attempted criteria.

For Language Arts Literacy, if a student attempted less than 20 percent of the items on one or two testing days but attempted 20 percent or more on the other testing day, a Void code appeared instead of a total test score on the student's reports. However, cluster scores are provided for parts of the Language Arts Literacy that are attempted. During the 2007 administration, 268 Language Arts Literacy tests were voided due to the attempted criteria for Day 1 and 347 Language Arts Literacy tests were voided due to the attempted criteria for Day 2.

Table 1.2 shows that a total of 105,865 students had valid scale scores in Language Arts Literacy, 106,980 students had valid scale scores in Mathematics, and 106,913 students had valid scale scores in Science. The number of valid scale scores is the number enrolled excluding the number of APA students, number not present, and number of voids.

Performance data shown in the Proficiency Levels columns include students who received valid scale scores. The number of students who scored in each proficiency level excludes students coded as APA. Because each content area is independent, students may receive a scale score in one content area, but not in others.

The total GEPA Language Arts Literacy, Mathematics, and Science scores are reported as scale scores with a range of 100 to 300. Scale scores of 100 and 300 are a theoretical floor and ceiling, which may not actually be observed. The score ranges are as follows:

<i>Advanced Proficient</i>	<i>250–300</i>
<i>Proficient</i>	<i>200–249</i>
<i>Partially Proficient</i>	<i>100–199</i>

A series of tables summarizing the test results for the State (general education students, special education students, limited English proficient students, and total students), District Factor Groups, Special Needs Districts, and All Other (Non Special Needs) Districts appears in Appendix B. See <http://www.state.nj.us/education/finance/> for information about District Factor Groups and Special Needs Districts (Abbott Districts).

Note that the percentages shown in tables throughout this Technical Report may not total to 100 due to rounding.

## **1.2 Purpose of the GEPA**

The GEPA serves as a primary indicator for identifying those students who may need instructional intervention in the three content areas of Language Arts Literacy, Mathematics, and Science. The test also serves as an indicator for determining which local education programs may require revisions to ensure that instructional programs are aligned with the Core Curriculum Content Standards. The GEPA is designed to evaluate the progress students are making in mastering the knowledge and skills required by the end of eighth grade. Also, the GEPA provides an indication of students' progress in the skills required to pass the High School Proficiency Assessment.

Three proficiency levels have been determined for each of the content areas of the GEPA: Partially Proficient, Proficient, and Advanced Proficient. Students scoring in the lowest level, Partially Proficient, are considered below the state minimum level of proficiency. These students may need instructional intervention. Instructional decisions for all students are determined only after additional information is considered, e.g., classroom tests, teacher observations.

In 1996, the State Board of Education adopted Core Curriculum Content Standards to describe what all students should know and be able to do at the end of fourth grade, eighth grade, and upon completion of a New Jersey public school education. The Core Curriculum Standards delineate New Jersey's expectations for student learning. All New Jersey school districts are required to organize instruction and design curricula so that virtually all students achieve these content standards. The Core Curriculum Content Standards defined the development of three statewide assessments: the Elementary School Proficiency Assessment Program, which was administered from 1997-2002; the GEPA, which replaced the Early Warning Test (EWT) in 1998; and the High School Proficiency Assessment, which replaced the High School Proficiency Test as the state's graduation requirement for all students who entered the eleventh grade in the fall of 2001.

Previously, in 1988, the New Jersey Legislature passed a law that established the Early Warning Test. The Legislature moved the High School Proficiency Test from the ninth grade to the eleventh grade. The Grade 11 High School Proficiency Test assessed essential reading, mathematics, and writing skills. It served as a graduation requirement for all public school students in New Jersey who entered ninth grade on or after September 1, 1991, and prior to the fall of 2001.

The Early Warning Test was similar to the High School Proficiency Test in eleventh grade because it also measured basic skills in reading, mathematics, and writing. The Early Warning Test was administered to all eighth-grade students each spring to determine whether they were making satisfactory progress in mastering the skills they would need to pass the High School Proficiency Test in the eleventh grade. The Early Warning Test was first administered as an operational test in March 1994.

Following the adoption of the Core Curriculum Standards in 1996, the development of the GEPA was defined. The GEPA was initially administered as field tests in Language Arts Literacy and Mathematics. In March 1999, the GEPA was administered for the first time as an operational assessment. Additional field tests in Language Arts Literacy, Mathematics, and Science were also administered and the GEPA Speaking assessment was pilot tested. In March 2000, Science was included in GEPA as an operational test for the first time.

Because the State Board required that the Core Curriculum Content Standards be reviewed and revised every five years, a review process began in May 2001 involving teachers, school administrators, students, parents, and representatives from business, higher education, and the community.

The language arts literacy, mathematics, and science standards were adopted by the State Board of Education in July 2002. In April 2004, the language arts literacy standards were revised to comply with the requirements of the No Child Left Behind Act of 2001 (NCLB) and readopted by the Board.

The GEPA administration in 2007 included field test items that were aligned with the new Core Curriculum Content Standards for language arts literacy, mathematics, and science. The GEPA test development procedures are detailed in Chapter 2 of this *Technical Report*.

### **1.3 GEPA Organizational Support**

**New Jersey Department of Education (NJDOE)** The GEPA is administered by the Office of State Assessments within the Department of Education. The staff of the Office of State Assessments directs the implementation of the statewide assessment programs. In addition to planning, scheduling, and directing all GEPA activities, the staff is extensively involved in numerous test review, security, and quality control procedures.

**The Educational Measurement group of Pearson, a business of NCS Pearson, Inc. ("Pearson")** is the primary contractor working in partnership with Measurement Incorporated (MI) and Assessment and Evaluation Services (AES). In 1998, the contract for developing and administering the GEPA was awarded to Pearson. Major Pearson activities include the following:

- Supporting and monitoring the test development cycle and subcontractor efforts toward content development
- Printing test books and ancillary materials required for the GEPA
- Distributing assessment materials in a secure manner and in appropriate amounts based on the district quantity survey results
- Supporting the regional workshops that inform district test coordinators about the GEPA program
- Receiving, scanning, editing, and scoring the answer documents using clearly defined quality control procedures
- Packaging and transporting open-ended responses to be hand-scored
- Providing accurate reports of test results to New Jersey pupils, parents/guardians, schools, districts, and the state

**Measurement Incorporated (MI)** MI provides item development and scores all open-ended responses for the GEPA program. Items developed include multiple-choice and constructed-response items for Language Arts Literacy, Mathematics, and Science; and writing prompts for Language Arts Literacy. MI scoring directors, NJDOE Office of State Assessments content specialists, and New Jersey teachers use rangefinding procedures to prepare for scoring the GEPA open-ended items.

**Assessment and Evaluation Services (AES)** AES is responsible for GEPA technical activities such as specifying the item selection for the operational tests, equating the test forms, and developing the scale score conversion tables.

## **CHAPTER 2: TEST DEVELOPMENT**

---

The New Jersey Department of Education has developed a comprehensive set of assessments that measure student achievement of the Core Curriculum Content Standards. The validity of the GEPA is therefore based on the alignment of the GEPA, the Core Curriculum Content Standards, and the knowledge and skills expected of eighth-grade students.

This chapter presents validity evidence based on test content. A description of the test specification development is followed by the procedures for test item development. Details about item writing, as well as task, prompt, and passage selection, are included. The last section delineates the review work of the New Jersey Assessment Content Committees. Additionally, an external committee assisted the New Jersey Department of Education by reviewing the assessments to determine how well they measure the knowledge and skills stated in the standards, and by comparing the New Jersey standards with those in other states and countries.

### **2.1 Test Specifications**

The GEPA content areas of Language Arts Literacy, Mathematics, and Science were designed from their inception in 1997 to align with the original Core Curriculum Content Standards adopted by the New Jersey State Board of Education in 1996. The State Board required that the Core Curriculum Content Standards be reviewed every five years. New standards for the three content areas were adopted by the Board in July 2002. To comply with requirements of the federal No Child Left Behind Act of 2001 (NCLB), the Language Arts Literacy standards were also revised in April 2004.

The Core Curriculum Content Standards were developed by teachers and other educational professionals from New Jersey. The Core Curriculum Content Standards outline what students should know and be able to do at a certain grade level. The questions on the GEPA can contain items/concepts included in the grade eight standards as well as for those standards listed in the prior grade standards.

The GEPA was first administered as an operational assessment in 1999. Prior to that time, Language Arts Literacy and Mathematics was administered to all eighth-grade students as field tests and “due-notice” administrations. Science was initially field tested in 1999. The purpose of due-notice administrations was to help school districts identify potential gaps between their curriculum and the test objectives, and to allow schools time to modify their curriculum and instructional practices to meet the needs of students before the first operational assessment. Field test items for Language Arts Literacy, Mathematics, and Science continued to be included with the GEPA 2000 – 2007 test administrations.

Following adoption of the original Core Curriculum Content Standards in 1996, the New Jersey Assessment Content Committees met through 1997 to develop a directory of test specifications and sample items for each content area to provide content/skill outlines and sample items. These directories describe the test, item formats, and test item scoring. This test specification work done by New Jersey educators serves as the foundation for all test item development.

The committees of New Jersey educators rely upon their expertise and the Core Curriculum Content Standards to design a test that is universally accessible to all eighth-grade students and is composed of test questions that are age- and grade-appropriate. The material in the three directories of test specifications and sample items is designed for use by curriculum specialists and teachers to improve instruction at the district, school, and classroom levels. Figure 2.1 summarizes the steps of the test development process beginning with the development of the Core Curriculum Content Standards and ending with an operational GEPA test form. Brief descriptions of the test content measured in Language Arts Literacy, Mathematics, and Science are presented in the following sections.

### Language Arts Literacy

Language Arts Literacy measures students' achievements in reading and writing. Language Arts Literacy currently assesses knowledge and skills in two content clusters:

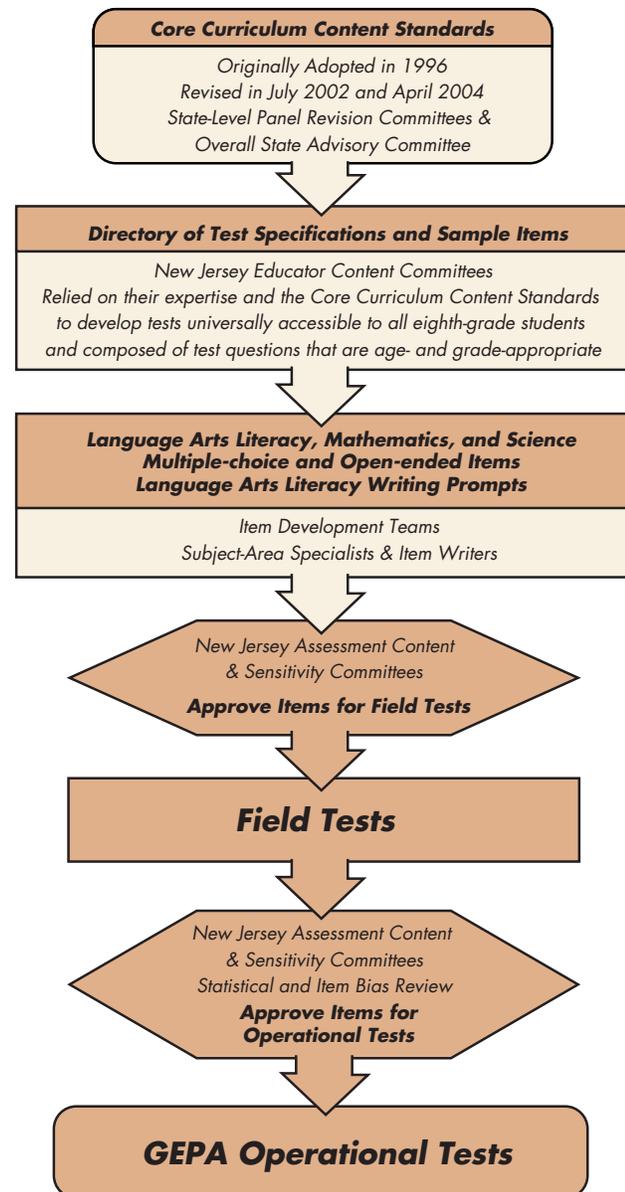
- Reading
- Writing

The Reading cluster consists of a narrative reading passage with ten multiple-choice and two open-ended items, and a persuasive reading passage with ten multiple-choice and two open-ended items. The passages are selected from published sources such as books, newspapers, magazines, and the Internet.

The Writing cluster for GEPA consists of two writing activities: a writing/persuade task in response to a prompt and a writing/speculate task in response to a picture.

For an in-depth description of the Language Arts Literacy assessment, refer to the *Directory of Test Specifications and Sample Items for the Elementary School Proficiency Assessment (ESPA)*, *Grade Eight Proficiency Assessment (GEPA)*, and *High School Proficiency Assessment (HSPA) in Language Arts Literacy* (February 1998). The directory is available online at <http://www.state.nj.us/education/njpep/assessment/TestSpecs/LangArts/TOC.html>, or by calling the New Jersey Department of Education, Publications Office, (609) 984-0549.

**FIGURE 2.1**  
**GEPA Test Development Process**



**The Technical Advisory Committee (TAC) advises and assists the Office of State Assessments in the development and implementation of the statewide testing program. TAC reviews and provides suggestions for each of the stages listed in the GEPA Test Development Process.**

### Mathematics

Mathematics measures students' ability to solve problems by applying mathematical concepts. The GEPA Mathematics assessment measures knowledge and skills in four content clusters:

- Number and Numerical Operations
- Geometry and Measurement
- Patterns and Algebra
- Data Analysis, Probability, and Discrete Mathematics

Mathematics items are also classified and reported as Knowledge (requiring conceptual understanding or procedural knowledge) and Problem Solving (applying mathematical concepts). For the operational test, there are a total of 30 multiple-choice and 6 open-ended items in Mathematics.

For an in-depth description of the GEPA Mathematics assessment, refer to the *Directory of Test Specifications and Sample Items for the Grade Eight Proficiency Assessment (GEPA) and the High School Proficiency Assessment (HSPA) in Mathematics* (February 1998). The directory is available online at <http://www.state.nj.us/education/njpep/assessment/TestSpecs/MathTestSpec/GEPAMath/MathIndex.html>, or by calling the New Jersey Department of Education, Publications Office, (609) 984-0549.

Additional information about the GEPA test specifications is included at [http://www.state.nj.us/education/njpep/assessment/TestSpecs/MathTestSpec/GEPAMath\\_sample\\_questions/worddocs/GEPA%20Math%202005%20presentation.ppt](http://www.state.nj.us/education/njpep/assessment/TestSpecs/MathTestSpec/GEPAMath_sample_questions/worddocs/GEPA%20Math%202005%20presentation.ppt)

### Science

Science measures knowledge and skills in three content clusters:

- Life Science
- Physical Science
- Earth Science

Science items are also classified and reported as Knowledge (Comprehension and Science, Society/Technology) and Application (Habits of Mind/Inquiry and Mathematics). For the operational test, there are a total of 45 multiple-choice and 3 open-ended items in Science.

For an in-depth description of the Science assessment, refer to the *Directory of Test Specifications and Sample Items for the Grade Eight Proficiency Assessment (GEPA) and the High School Proficiency Assessment (HSPA) in Science* (February 1998). The directory is available online at [http://www.state.nj.us/education/njpep/assessment/TestSpecs/science\\_test\\_specs/Science\\_GEPA\\_HSPA/](http://www.state.nj.us/education/njpep/assessment/TestSpecs/science_test_specs/Science_GEPA_HSPA/), or by calling the New Jersey Department of Education, Publications Office, (609) 984-0549.

Additional information about the GEPA test specifications is included at <http://www.state.nj.us/education/njpep/assessment/TestSpecs/ScienceGEPA/index.html>

Tables 2.1, 2.2, and 2.3 summarize the total points possible for Language Arts Literacy, Mathematics, and Science of the content areas of the operational GEPA administered in 2007.

## 2.2 Development of Test Items

The 2007 GEPA consists of two types of items:

- operational test items used to determine students' scores and
- field test items evaluated for use as future operational test items.

The 2007 operational test for Language Arts Literacy, Mathematics, and Science was composed of items field tested through 2006. The item development teams consisted of subject-area specialists and consulting item writers. These writers were teachers or former teachers with a great deal of specialized knowledge (e.g., education and training, years of classroom experience, familiarity with the student population, knowledge of the content area, and understanding of the pedagogy that defines the discipline) concerning their area of content expertise.

Each of the content areas consists of multiple-choice and open-ended items. The multiple-choice items are designed to measure those skills determined to be best measured by such item types, and the open-ended items are developed to measure those skills requiring students to do more than select a correct answer. That is, the open-ended items are designed to tap more complex and integrated skills. Language Arts Literacy includes a writing/persuade task and a writing/speculate task in response to a picture.

The Measurement Incorporated/Pearson item development process for each testing cycle begins with a formal review of the Core Curriculum Content Standards and the three directories of test specifications. Item-writing training sessions typically last from 8 to 16 hours over two days. The respective test development specialist for each content area conducts the training session. Between the first and second sessions, preliminary versions of test items developed in the first session are evaluated. At the second session, the training is focused on the items developed in the first session.

At the training, each consulting item writer is asked to sign a Letter of Agreement. This letter specifies the confidentiality and security regulations. The agreement also outlines the ownership regulations. No confidential materials related to the project are released without explicit approval of the NJDOE Office of State Assessments.

During the training, each item writer is given the following information:

- An overview of the GEPA
- Final test blueprint for each subject-area test and item specifications
- A description of the item formats used, including important characteristics of each format
- A description of the item writing process and measures to take to avoid writing biased items
- A listing of the security procedures followed during the item development process.

Important guidelines for the GEPA item development and test structure are outlined below.

1. Items are written to reflect what students know and understand based on classroom instruction and their mastery of skills included in the Core Curriculum Content Standards. Items are also designed to assess higher-order or critical thinking skills in varied contexts that students are likely to understand; yet, they are based upon solid theoretical frameworks.
2. For each content area, the multiple-choice items represent a range of difficulty. For example, approximately 25 percent of the items are relatively easy, 50 percent of the items are somewhat difficult,

**TABLE 2.1**

**Total Points Possible for the Language Arts Literacy Component of the GEPA**

<b>Language Arts Literacy</b>		
<b>Total</b>	<b>54 points</b>	
Reading	36 points*	
Writing	18 points*	
Writing/Speculate	6 points*	1 – 6 points, ratings averaged
Writing/Persuade	12 points*	1 – 6 points, ratings summed
-----		
Interpreting Text	15 points*	
Analyzing/Critiquing Text	21 points*	

\* Cluster-level results show how students perform on the sets of items that measure particular knowledge and skills (clusters above the dotted line) or particular processes (clusters below the dotted line). Though an item on the GEPA can contribute to a cluster above the line (for example, Reading) as well as a cluster below the line (for example, Interpreting Text), each item is counted only once in the total score.

**TABLE 2.2**

**Total Points Possible for the Mathematics Component of the GEPA**

<b>Mathematics</b>		
<b>Total</b>	<b>48 points</b>	
Number and Numerical Operations	12 points*	
Geometry and Measurement	12 points*	
Patterns and Algebra	12 points*	
Data Analysis, Probability, and Discrete Mathematics	12 points*	
-----		
Knowledge	48 points*	
Problem Solving	35 points*	

\* Cluster-level results show how students perform on the sets of items that measure particular knowledge and skills (clusters above the dotted line) or particular processes (clusters below the dotted line). All Mathematics items are classified as Knowledge because all items require conceptual understanding or procedural knowledge. Some items also measure Problem Solving. Each Mathematics item counts only once in the total score.

**TABLE 2.3**

**Total Points Possible for the Science Component of the GEPA**

<b>Science</b>		
<b>Total</b>	<b>54 points</b>	
Life	22 points*	
Physical	16 points*	
Earth	16 points*	
-----		
Knowledge	11 points*	
Application	43 points*	

\* Cluster-level results show how students perform on the sets of items that measure particular knowledge and skills (clusters above the dotted line) or particular processes (clusters below the dotted line). Though an item on the GEPA can contribute to a cluster above the line (for example, Life) as well as a cluster below the dotted line (for example, Knowledge), each item is counted only once in the total score.

and 25 percent of the items are difficult. This range of difficulty provides for a distribution of items with p-values from approximately 0.30 to 0.95. This distribution allows for a range of difficulty that supports the established proficiency levels, yet is not so difficult that low-achieving students cannot be assessed adequately.

3. Item content for all of the items, including the writing-task prompts, is carefully reviewed to ensure that the items are free from gender, ethnic and regional bias. Across all content areas of the GEPA and in any test material presented, there is a balance of gender and active/passive roles by gender.
4. Measurement Incorporated/Pearson construct initial rubrics for each open-ended item in Language Arts Literacy, Mathematics, and Science.
5. Writing task prompts for Language Arts Literacy are written in such a way that they focus on experiences that eighth-grade students may have every day. However, care must be taken to ensure that the writing task prompts are not intrusive in nature and do not elicit personal information of a biographical, religious, political, or affective nature. Topics must be chosen so that no group of eighth-grade students is put at a subject-related disadvantage. Instead, each writing task prompt is designed to sample the skills and abilities demanded of eighth-grade students. Each writing task is developmentally appropriate for students in both the academic and nonacademic environments.

As items are developed, Measurement Incorporated/Pearson documents each item's relevancy to the Core Curriculum Content Standards and to the directories of test specifications. During this process, each item is assigned a unique item ID number or coding system number. This unique number identifies the following: content area, skill measured, standard, and associated materials such as a reading passage, artwork, or display of data. The number is used to track the item throughout the development process and its eventual use on the operational test.

All items prepared by item writers are reviewed, revised, and edited by the subject area specialists and editors prior to review by the New Jersey Assessment Content Review Committees. Also, the New Jersey Assessment Sensitivity Review Committee approves passages used on the Language Arts Literacy section.

**2.3 Item Review Process**

The New Jersey Assessment Content Committee members provide expert judgments on the alignment of each test item with the Core Curriculum Content Standards and the content-specific test specifications. The committee members represent school districts across all District Factor Groups. Table 2.4 shows the District Factor Groups represented on each of the Content and Sensitivity Committees.

**TABLE 2.4**

**District Factor Groups (DFG) Represented on the GEPA Content and Sensitivity Committees**

<b>DFG</b>	<b>Language Arts Literacy</b>	<b>Mathematics</b>	<b>Science</b>	<b>Sensitivity</b>	<b>Total</b>
A	1	0	4	1	6
B	3	2	2	2	9
CD	1	2	0	1	4
DE	2	2	1	0	5
FG	3	3	1	2	9
GH	2	4	1	0	7
I	1	0	2	3	6
J	0	1	1	0	2
Retirees	3	7	3	3	16
Private School	0	0	1	0	1
Not in Districts	0	0	0	0	0
<b>Total</b>	<b>16</b>	<b>21</b>	<b>16</b>	<b>12</b>	<b>65</b>

Committee members sign a Confidentiality and Security Agreement noting they must maintain the security of the testing materials by not discussing and disclosing any confidential information related to the program.

**FIGURE 2.2**

**Item Approval Before Field Test**

<b>Sensitivity</b>	<b>Content</b>
<b>*Comments</b>	<b>*Comments</b>
Sensitivity Issue                      Yes              No	Meets Specifications                      Yes              No
If yes, identify category and explain*	Appropriate Difficulty                      Yes              No
	Accurate Coding                      Yes              No
Definitely Use	Definitely Use
Revise and Use With Approval	Revise and Use With Approval
Revise and Resubmit	Revise and Resubmit
Do Not Use*	Do Not Use*

Sensitivity Sign-off

Date

Content Chairperson's Signature

Date

Prior to field testing, all items are reviewed by the Office of State Assessments staff and committee members. Each test item is reviewed to determine if the item meets test specifications and addresses an appropriate level of difficulty. Committees also ensure that test questions are not offensive and do not reinforce negative stereotypes, and that test questions appropriately reflect multicultural society.

Figure 2.2 presents a sample of the form that must be marked “Definitely Use” or “Revise and Use With Approval” during review committee meetings before an item is included in a field test. The percentage of items accepted for field testing depends on the content area and the item type. The range of acceptance generally is 60-80% at this item review stage. During review, committee members approve items, amend or revise items, or reject items.

No new items were developed for the 2007 field test. Items field tested in 2007 were approved during item development review meetings in 2005 or earlier. Some items were previously field tested. During the statistical review of these previously field tested items, the items were judged revise for refield testing.

The 2007 field test items included 22 Language Arts Literacy, 14 Mathematics, and 35 Science items. The Writing component of Language Arts Literacy included one prompt for the writing/persuade task and four pictures for the writing/speculate task. Table 2.5 shows the number of multiple-choice and open-ended items specified for each content area.

The committees met in August 2007 to review item statistics from the March 2007 field testing. The statistical item review meetings are listed in Table 2.6. Because the Office of State Assessments requested no new item development for 2007 or 2008 field testing, no item development meetings were held during spring and summer 2006 and 2007.

TABLE 2.6

**GEPA 2007 Statistical  
Committee Meetings**

<b>Language Arts Literacy Committee</b>
Statistical Item Review Wednesday, August 22
<b>Mathematics Committee</b>
Statistical Item Review Friday, August 3
<b>Science Committee</b>
Statistical Item Review Thursday, August 23

**Because the Office of State Assessments requested no new item development for 2008 field testing, only statistical item review meetings were held during summer 2007.**

**TABLE 2.5  
Number of Items Field Tested in 2007**

	<b>Multiple-choice Items</b>	<b>Open-ended Items</b>	<b>Total Items</b>
<b>Language Arts Literacy</b>	18	4	22
<b>Mathematics</b>	10	4	14
<b>Science</b>	30	5	35
<b>TOTAL</b>	<b>58</b>	<b>13</b>	<b>71</b>

At the statistical review, committee members consider how well students did on each field test question in comparison to the other questions on the GEPA. If an item yields good statistics, it will become part of the operational pool for future GEPA tests. Otherwise, it will be eliminated or revised and re-field tested.

Prior to field test statistical review, the field-tested open-ended items and writing prompts must go through rangefinding to determine the scores on sample student responses. The field test rangefinding process involves scoring 30 student responses for each of the open-ended items and writing prompts. These 30 responses are selected to represent the wide range of responses to that item. The papers are scored by one or two content committee members, the NJDOE Content Coordinator, and representatives from Measurement Incorporated.

In Language Arts Literacy, the responses are scored according to the generic rubric for either reading or writing as appropriate. Use of these generic rubrics ensures that student responses are scored in the same way for the demonstration of the same level of knowledge and skills regardless of the prompt or the year.

For Mathematics and Science, each item has a unique scoring rubric, based on the generic one for each area. During rangefinding, the item specific rubric is refined, if necessary, to define each score point clearly. The rangefinding process aids in delineating between a 0 & 1, 1 & 2, and a 2 & 3 score point response. The holistic scoring guide is used quite often to refine the tenuous line between the score points.

For all content areas, the scored field test responses and the rubrics are used to create the holistic scoring guide, which is used to help refine the lines between the score points. This guide is then used to train the scorers of that item. If there is any problem or question with the scoring of a student's response, the NJDOE Content Coordinator is contacted and makes a final decision for the score of that paper. After the open-ended papers have been scored, the scorers discuss the types of responses and problems, if any, found during scoring of each item. The scoring director then writes a brief summary of these comments and sends it, along with a copy of each item, rubric, sample answer, and rangefinding paper to the statistics review. Other than this packet, the same field test review procedures are used for the open-ended and multiple-choice items.

Pearson computes item means, response frequencies, biserial correlations, and other descriptive statistics. Prior to the presentation of items and statistics to reviewers, the NJDOE Office of State Assessments defined boundaries within which item statistics should fall. In general, items with p-values below 0.30 or above 0.95 were considered usable only if a strong content argument could be made for their inclusion in the item bank. An item could be flagged for low or high p-value and/or low biserial correlation with operational test total scores.

For the statistical item review, the Mantel-Haenszel statistic is calculated to show whether or not students are responding to an item in a way that their overall ability would lead us to expect. This statistic takes into consideration both group membership (by race or by gender) and ability. The Mantel-Haenszel statistic is used for a classification determination of category A, B, or C. An item in Category A shows no or minor relationship between group membership and performance. Category B items are somewhat suspect. Category C items show a substantial relationship between group membership and item performance and must be examined carefully by the committees to make sure these items are not biased. The Mantel-Haenszel statistic is used at Educational Testing Service (ETS) as a classification determination of category A, B, and C as described by Zieky (1993):

- 
- Category A) MH D-DIF not significantly different from zero  
OR  
absolute value less than 1.0*
- Category B) MH D-DIF significantly different from zero and absolute value of at least 1.0  
AND EITHER  
(1) less than 1.5  
OR  
(2) not significantly greater than 1.0*
- Category C) MH D-DIF significantly greater than 1.0  
AND  
absolute value 1.5 or more. (p. 342)*
- 

For every open-ended item and writing prompt, the Sensitivity Committee reviews frequency distributions for the range of scores of the following student groups: total, white, African American, Hispanic, Asian, American Indian, male, and female.

For the multiple-choice items field tested during 2007, three items in Language Arts Literacy, no items in Mathematics, and one item in Science were flagged. All flagged items were approved during the sensitivity and content reviews. However, the Sensitivity Committee noted a concern about one of the other Language Arts Literacy multiple-choice items for the reading passage. For this item, the Sensitivity Committee marked “Do Not Use.”

**FIGURE 2.3**

**Item Approval Before Operational Test**

<b>Sensitivity</b>		<b>Content</b>	
*Comments		*Comments	
Sensitivity Issue <input type="checkbox"/> Yes <input type="checkbox"/> No		Appropriate Difficulty <input type="checkbox"/> Yes <input type="checkbox"/> No	
If yes, identify category and explain*		PVal = Biserial =	
Mantel-Haenszel Category C <input type="checkbox"/> W-AA <input type="checkbox"/> W-H <input type="checkbox"/> M-F			
<input type="checkbox"/> Yes <input type="checkbox"/> No	Definitely Use		<input type="checkbox"/> Yes <input type="checkbox"/> No
<input type="checkbox"/> Yes <input type="checkbox"/> No	Revise and Use With Approval**		<input type="checkbox"/> Yes <input type="checkbox"/> No
<input type="checkbox"/> Yes <input type="checkbox"/> No	Revise and Re-Field Test		<input type="checkbox"/> Yes <input type="checkbox"/> No
<input type="checkbox"/> Yes <input type="checkbox"/> No	Do Not Use*		<input type="checkbox"/> Yes <input type="checkbox"/> No

Sensitivity Sign-off

Date

Content Chairperson's Signature

Date

\*\*Requires director's approval

Figure 2.3 presents a sample of the form that must be marked “Definitely Use” or “Revise and Use With Approval” during review committee meetings of the field test statistics before an item is included on an operational base test.

Tables 2.7 – 2.10 present the number of items field tested during the administration.

Table 2.7 shows 18 multiple-choice items and 4 open-ended items were field tested for the Reading component of Language Arts Literacy. During the statistical review, the Language Arts Literacy committee approved 16 multiple-choice items and 4 open-ended items for the narrative passage for operational tests.

Table 2.8 shows the results of the Writing component of Language Arts Literacy from the field tested four pictures for the writing/speculate task and one prompt for the writing/persuade task. All speculative (picture) prompts and persuasive prompts were approved for operational tests.

Table 2.9 reports the results by content cluster for the 10 multiple-choice items and 4 open-ended Mathematics items field tested in 2007. Each content cluster is further divided into strands. Information about the test specifications, including the associated strands, is located at <http://state.nj.us/education/njpep/TestSpecs/MathTestSpec/GEPAMath/Macros.html>. Table 2.9 indicates that 100% Mathematics multiple-choice items and 25% Mathematics open-ended items were approved for an operational base test.

TABLE 2.7

**LANGUAGE ARTS LITERACY - READING**  
**Number of Field Test Items Approved During Statistical Review**

Passages	Field-Tested		Approved		Revise & Re-Field Test		Do Not Use	
	MC	OE	MC	OE	MC	OE	MC	OE
Narrative 1	18	4	16	4	0	0	2	0

TABLE 2.8

**LANGUAGE ARTS LITERACY - WRITING**  
**Number of Field Test Items Approved During Statistical Review**

Prompts	Field Tested	Approved	Revise & Refield Test	Do Not Use
Speculate (Picture)	4	4	0	0
Persuade	1	1	0	0
<b>TOTAL</b>	<b>5</b>	<b>5</b>	<b>0</b>	<b>0</b>

TABLE 2.9

**MATHEMATICS**  
**Number of Field Test Items Approved During Statistical Review**

Content Cluster	Strand	Field-Tested		Approved		Revise & Re-Field Test		Do Not Use	
		MC	OE	MC	OE	MC	OE	MC	OE
Number and Numerical Operations	A	1	0	1	0	0	0	0	0
	B	1	0	1	0	0	0	0	0
	C	0	1	0	1	0	0	0	0
Geometry and Measurement	A	1	0	1	0	0	0	0	0
	B	0	1	0	0	0	0	0	1
	C	0	1	0	0	0	0	0	1
	D	0	0	0	0	0	0	0	0
	E	2	0	2	0	0	0	0	0
Patterns and Algebra	A	1	0	1	0	0	0	0	0
	B	0	1	0	0	0	0	0	1
	C	1	0	1	0	0	0	0	0
	D	1	0	1	0	0	0	0	0
Data Analysis, Probability, and Discrete Mathematics	A	1	0	1	0	0	0	0	0
	B	1	0	1	0	0	0	0	0
	C	0	0	0	0	0	0	0	0
	D	0	0	0	0	0	0	0	0
<b>TOTAL</b>		<b>10</b>	<b>4</b>	<b>10</b>	<b>1</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>3</b>

**TABLE 2.10**  
**SCIENCE**  
**Number of Field Test Items Approved During Statistical Review**

Content and Skill Clusters	Field-Tested		Approved		Revise & Re-Field Test		Do Not Use	
	MC	OE	MC	OE	MC	OE	MC	OE
<b>Life</b>								
Knowledge	4	0	3	0	0	0	1	0
Application	8	3	8	2	0	0	0	1
<b>Physical</b>								
Knowledge	2	0	2	0	0	0	0	0
Application	10	1	8	1	0	0	2	0
<b>Earth</b>								
Knowledge	3	0	3	0	0	0	0	0
Application	3	1	1	1	0	0	2	0
<b>TOTAL</b>	<b>30</b>	<b>5</b>	<b>25</b>	<b>4</b>	<b>0</b>	<b>0</b>	<b>5</b>	<b>1</b>

Table 2.10 shows that 30 multiple-choice and 5 open-ended Science items were field tested in 2007. This indicates that 83.3% Science multiple-choice items and 80% Science open-ended items were approved for an operational test. The number of Science items field tested for each content cluster as well as by knowledge skill and application skill is shown in the table.

Information about the science test specifications is located at [http://www.state.nj.us/education/njpep/TestSpecs/ScienceGEPA/TestSpecsRev9\\_04.doc](http://www.state.nj.us/education/njpep/TestSpecs/ScienceGEPA/TestSpecsRev9_04.doc)

#### 2.4 Operational Test Development

Following the 1998 through 2001 administrations, GEPA examiners completed a feedback form seeking suggestions and concerns related to the testing procedures. Questions related to timing, directions, and answer documents were asked specifically for each content area tested. Also, examiners were asked to identify questions that arose on issues and topics not addressed in the test booklets, directions, or coordinator or examiner manuals.

A sample of the 2001 questions is provided below:

- Was the time allotted for students to complete the test sufficient?
  - too much time
  - time about right
  - too little time
- Were the directions clear?
  - yes, directions were clear
  - no, directions were somewhat confusing
- Was the space provided for student responses in the answer folder sufficient?
  - adequate space
  - not enough space

**TABLE 2.11**  
**Operational Test Specifications**

Content Areas	Cluster	Number of Items		
		MC	OE	Total
Language Arts Literacy	Reading	20	6	26
	Writing	20	4	24
	Writing/Speculate		1	1
	Writing/Persuade		1	1
Mathematics	Number and Numerical Operations	30	6	36
	Geometry and Measurement	9	1	10
	Patterns and Algebra	6	2	8
	Data Analysis, Probability, and Discrete Mathematics	6	2	8
Science	Data Analysis, Probability, and Discrete Mathematics	9	1	10
	Life	45	3	48
	Physical	19	1	20
	Earth	13	1	14

Information from the examiners' responses assisted the Office of State Assessments with determining the operational testing procedures.

The GEPA Content Committees assisted with recommending the emphases and priorities reflected in the number of items for each item type and cluster on the operational test. The operational test specifications appear in Table 2.11.

Following the approval of test items for use on operational tests by the Content and Sensitivity Review Committees, Assessment and Evaluation Services (AES) selected items for each GEPA administration to meet test specifications for Language Arts Literacy, Mathematics, and Science.

Relevant considerations for operational test development included content quality and scope, cluster representation, and appropriate item difficulty indices. The new operational test was parallel to the content, format, and statistical characteristics of the previous operational forms. Selecting test items for the operational tests is an iterative process to create test forms that are the perfect combination of content and statistical information. Through the iterative process, item content took precedence over statistical characteristics.

The operational test development used the Rasch model to pre-equate cluster and total test scores. Rasch item difficulty statistics were calibrated to the previous test administration. Common items were chosen to link the Mathematics and Science operational tests to previous forms for equating purposes. For Language Arts Literacy, the forward and backward items for equating purposes were specified. For each operational test, AES produces a spreadsheet that includes the following information for both the previous operational test and newly developed operational test.

- Item identifier with item type (multiple-choice or open-ended), content clusters, and skill clusters
- Common items for equating
- P-values and biserial correlations
- Item difficulties with sums and averages for clusters and total test

### **2.5 Review and Approve Operational Test Forms**

The Office of State Assessments approved the operational test forms for each GEPA administration. AES and Pearson assisted with quality control that included:

- Confirm that each test item appears on the operational test as it was approved by the Content and Sensitivity Review Committees.
- Confirm that all test specification requirements are met.
- Check adequacy of common item set (i.e., in terms of size, content and skill representation)
- Double-check that the item and mean difficulty levels are accurate and meet requirements.
- Take the test to be certain all content considerations including content/skill/topic balance, correct keys, no clueing, and correct graphics are met.

### **2.6 Test Materials for Visually Impaired Students**

The Office of State Assessments works with the New Jersey Commission for the Blind and Visually Impaired to identify items with graphs, charts, and illustrations that may not translate well into Braille or large-print versions of the test. For 2007, the Writing/Speculate prompt from Language Arts Literacy, four items from Mathematics, and five items from Science were removed from the Braille form.

---

## CHAPTER 3: TEST ADMINISTRATION

---

### 3.1 Participation

In 1988, the New Jersey State Legislature passed a law (18A:7C-6.2) requiring that a test be given to all eighth-grade students in public schools in New Jersey to assess their progress toward mastering the skills they will need to graduate from high school. All eighth-grade public school students must take the GEPA. This includes:

- General education students
- Limited-English Proficient (LEP) students
- Special Education (SE) students
- Students with Disabilities (Section 504)
- Retained eighth graders

In accordance with the Individuals with Disabilities Education Act (IDEA), students who are receiving special education services must participate in each subject area of the age-appropriate statewide assessment with the following exception:

---

*Students with disabilities shall participate in the Alternate Proficiency Assessment in each content area where the nature of the student's disability is so severe that the student is not receiving instruction in any of the knowledge and skills measured by the general statewide assessment and the student cannot complete any of the types of questions on the assessment in the content area(s) even with accommodations and modifications. (New Jersey Administrative Code Chapter 6A:14-4.11[a]2)*

---

The Alternate Proficiency Assessment (APA) is a portfolio-style assessment designed to measure progress toward achieving New Jersey's state educational standards for those students with severe disabilities who are unable to participate in the New Jersey Assessment of Skills and Knowledge (NJASK), the Grade Eight Proficiency Assessment (GEPA), or the High School Proficiency Assessment (HSPA).

### 3.2 Test Security Procedures

The test booklet and its contents are secure materials. They are not to be read or copied, either wholly or in part, for any purpose without express written permission from the New Jersey Department of Education. It is the responsibility of the school districts to guarantee the security of the test materials. Security breaches may have financial consequences for the district, professional consequences for staff, and disciplinary consequences for students.

The items and passages contained in the test booklet must remain confidential because some test items reappear in future versions of the tests. The answer folders (approximately 56 pages) contain grids for marking the answers to multiple-choice questions. Also, the answer folders are used by students for writing responses to the open-ended questions and the writing essay prompts. The security of test items and passages is required to maintain the stability of the test item pool over time from a technical perspective and to enable comparisons to be made from one year to the next. Examiners, proctors, and other school personnel are prohibited from discussing or disclosing any test items before, during, or after the test administration.

The following are secure materials for the GEPA administration:

- Test booklets
- Used answer folders
- All other answer folders until after testing
- Mathematics Reference Sheets until after testing

Pearson assigns a unique identification number to each secure test booklet and answer folder. The unique identification numbers are listed on security checklists. The unique identification number appears as a bar-code on test booklets. Following the test administration, Pearson compares bar-code scan files of returned test booklets with distribution files to determine if all secure materials have been returned from each school and district. Pearson contacts any district with missing secure test booklets or answer folders. For the 2007 administration, Pearson scanned more than 123,000 secure test booklets.

The NJDOE Office of State Assessments outlined the following security procedures in the *GEPA Test Manual*. District test coordinators were trained in these procedures during regional meetings held by the Office of State Assessments in January and February 2007.

1. The chief school administrator or designee must sign for the initial shipment of test materials after presenting the Authorization to Receive Secure Test Materials form to the agent of the delivery service when the materials are delivered.
2. When not being used during testing, test materials must be stored in a secure, locked place that is accessible only to individuals whose access has been authorized by the school test coordinator. During testing, secure materials must not be removed from the testing room for review or photocopying. **Security of test materials must be maintained at all times.**
3. Each test booklet and answer folder has a **unique identification number**. Students must use the same test booklet and the same answer folder for each day of testing. On the first day of testing, students should print their name on the front cover of the test booklet assigned to them, and record the number and form letter of that test booklet on their answer folder.

4. Teachers are NOT to be given their own test booklet. The shrink-wrapped packaging on the test booklets may be opened for distribution just prior to testing.
5. Each day's section of the test booklet is sealed on all open sides. There are separate seals for the Science section, the Mathematics section, and Day 1 and Day 2 of the Language Arts Literacy section of the test. These seals must not be broken until the student breaks them the day that test section is administered.
6. District and school test coordinators must use the District and School Security Checklists to maintain an accurate record of the chain of distribution and collection of all test booklets.
7. Answer folders must not be duplicated or handscored.
8. An answer folder must be gridded for every enrolled Grade 8 student regardless of APA status.
9. An Irregularity Report form is used to report irregularities involving test booklets, answer folders, or anything that could impact test takers.
10. The principal and the chief school administrator or his/her designee must review and sign the completed Header sheets before they are submitted for scoring. The signatures affirm that the number of answer folders returned is correct and that all GEPA test administration procedures outlined in the manuals have been followed.
11. The Office of State Assessments, in cooperation with county offices, monitors all aspects of testing and the implementation of security procedures at selected sites. Announcements of security visits are not made in advance.

The district test coordinators' training and the *Test Manual* include responsibility descriptions for the district test coordinator, school test coordinator, and examiner.

A security plan sample in the *Test Manual* delineated tasks and responsibilities for the following: turnkey training, storage of secure materials, delivery problems, missing test booklet, chain of command, sick child, disruptive student, fire drill/bomb scare, and inclement weather.

The Office of State Assessments staff members monitor the test administration with specific procedures such as:

- Prior to actual testing, observe initial instructions from the examiners and proctors to the students.
- Observe all testing sites, including rooms where special accommodations are provided.

Breach test forms and examiner's manuals were prepared in the event of a security breach. In schools with the security breaches, appropriate staff members completed each student's name, date of birth, and answer folder number so that the alternate scoring occurred properly for the students. Specialized scoring and reporting included developing alternate test score keys, conversion tables, and reports.

### **3.3 Test Administration Procedures**

The district test coordinators, school test coordinators, and examiners are responsible for the proper administration of the test. The district test coordinator is responsible for ensuring that examiners are selected and trained. All examiners must be certified teachers currently employed by the school. The district and school test coordinators, and examiners must read the *Test Manual* and *Examiner Manual* carefully to get an overview of all activities.

Student Rosters with appropriate Special Codes must be prepared to include each and every eighth-grade student in the district. Districts are required to have a student roster for EACH classroom in which testing is to occur. The roster should list all of the students that are testing in a particular classroom setting.

The Student Rosters must:

- List each eighth-grade student's name
- Identify students with SE classifications, IEP exemptions/accommodations, or Section 504 status
- Identify students who are Limited English Proficient and receiving accommodations
- Track students who need to make up a section of the test

Information from the Student Rosters is used to:

- ensure students are testing in the correct room
- verify correct gridding by students, and to
- keep attendee records.

Test booklets and answer folders are distributed to examiners only on the morning of each day of the test administration. Specific instructions for the test administration are contained in the *Examiner Manual*. The examiners' familiarity with the materials and the prescribed procedures is essential to the successful administration of the test. During the examiners' training, district and school test coordinators emphasize that students can be given no assistance or coaching beyond what is specified in the manual.

When more than 25 students are tested in one room, the examiner uses the assistance of proctors. The school test coordinator briefs the proctors on the test materials and procedures, and specifies their responsibilities before, during, and after test administration. Proctors help in distributing and collecting non-secure materials, in observing students from different points in the room during test administration, and in answering student questions when there is a problem related to the test directions.

Total testing time (including time for distributing and collecting materials, reading directions, and taking breaks) is approximately nine hours over four successive days. The GEPA test administration must be scheduled in the morning. The Science, Mathematics, and Language Arts Literacy content-area tests were administered on the specified dates during the regular and make-up testing weeks.

### **3.4 Test Accommodations**

To ensure that students are tested under appropriate conditions, the Department of Education has adopted test accommodations and modifications that may be used when testing special populations of students. The content of the test typically remains the same, but administration procedures, setting, and answer modes may be adapted. Students requiring accommodations must be tested in a separate location from general education students.

**General education students** receive no special testing accommodations other than the standard room setup and materials distribution described in the examiner's section of the *Test Manual*.

**Limited English Proficient (LEP) students** are tested with one or more of these accommodations:

- Additional time up to 150% of the administration times indicated
- Translation of directions only to the student's native language. Translations of passages, items, prompts, and tasks are NOT permitted
- Use of a bilingual dictionary, preferably one normally used by the student as part of the instructional program

**Students with Disabilities (SE/504)** must take the GEPA unless their Individualized Education Program (IEP) specifically states that they take the Alternate Proficiency Assessment (APA) and not the GEPA.

Students who are eligible under Section 504 of the Rehabilitation Act of 1973 may be tested using modified testing procedures that must be specified in the student's 504 accommodation plan.

**Visually impaired students** may take either a Braille or large-print version of the test. Specific instructions for administering the Braille and large-print versions of the test are provided in the supplementary instructions for examiners administering these forms.

Students using the Braille test booklets:

- are instructed to bring a Braille ruler and a talking calculator to the test session.
- are instructed to skip some items identified in the Braille instructions. The spaces for these items must be left blank on the student answer folder.
- have answer folders transcribed from Braille version by the examiner.
- dictate their answers to the examiner or use a device that produces Braille. For dictations and responses recorded in Braille:
  - Students must indicate all punctuation and must spell all key words.
  - Examiners must transcribe the Brailled responses into the regular answer folder.

Students using the large-print test booklets:

- mark their answers in the large-print answer folders.
- may be instructed to skip some questions. The spaces for these questions must be left blank in the student's large-print answer folder.
- who dictate responses on open-ended items and writing tasks indicate all punctuation and spell key words.

Accommodations and modifications of test administration procedures are listed in Appendix C of this report. Also, the accommodations and modifications are included in the *Test Manual*, the *Examiner Manual*, and at <http://www.state.nj.us/education/specialed/accom900.htm>

If a student requires an accommodation or modification that is not listed, district staff are instructed to contact the Office of State Assessments, GEPA Coordinator. Accommodations or modifications are classified as follows:

A = Setting Accommodations

B = Scheduling Accommodations

C = Test Materials/Modifications

D = Test Procedures Modifications

### **3.5 Results for Special Education Students and Section 504 Students Tested with Accommodations or Modifications**

The following tables show the proficiency level results for special education students and Section 504 students tested with accommodations and modifications. Also, the first row of each table includes the number of students and performance results for Special Education students as shown in Table 1.2 of this *Technical Report* and the state level Performance by Demographic Groups Report from Cycle II reporting.

Not every special education student or Section 504 student is tested with an accommodation or modification. Accommodations and modifications may be used separately or in combination. Table 3.1 shows the number of special education students with performance results and the number of Section 504 students with performance results tested with each of the accommodations and modifications.

Tables 3.2–3.5 show the numbers of students and proficiency results by special education disability category. Instructions to the examiners note that “...one and only one disability category for each special education student...” should be designated. The N category is used to indicate multiple grids. Also, the N category is a default code used when a school fails to provide the specific disability-category information listed for an APA student.

**TABLE 3.1**  
**Setting Accommodations and Modifications for Special Education Students and Section 504 Students**

	NUMBER OF STUDENTS ENROLLED	LANGUAGE ARTS LITERACY						MATHEMATICS						SCIENCE								
		PARTIALLY PROFICIENT (100 – 199)		PROFICIENT (200 – 249)		ADVANCED PROFICIENT (250 – 300)		PARTIALLY PROFICIENT (100 – 199)		PROFICIENT (200 – 249)		ADVANCED PROFICIENT (250 – 300)		PARTIALLY PROFICIENT (100 – 199)		PROFICIENT (200 – 249)		ADVANCED PROFICIENT (250 – 300)				
		NO.	%	NO.	%	NO.	%	NO.	%	NO.	%	NO.	%	NO.	%	NO.	%	NO.	%			
<b>Special Education</b>	<b>18,197</b>	<b>11,390</b>	<b>67.1</b>	<b>5,375</b>	<b>31.7</b>	<b>206</b>	<b>1.2</b>	<b>17,190</b>	<b>12,234</b>	<b>71.2</b>	<b>4,339</b>	<b>25.2</b>	<b>617</b>	<b>3.6</b>	<b>17,173</b>	<b>8,502</b>	<b>49.5</b>	<b>7,630</b>	<b>44.4</b>	<b>1,041</b>	<b>6.1</b>	
Special Education Accommodations		14,654	10,019	68.4	4,485	30.6	150	1.0	14,780	10,667	72.2	3,634	24.6	479	3.2	14,765	7,385	50.0	6,558	44.4	822	5.6
Setting Accommodations		14,819	10,064	67.9	4,603	31.1	152	1.0	14,942	10,740	71.9	3,707	24.8	495	3.3	14,927	7,442	49.9	6,636	44.5	849	5.7
Scheduling Accommodations		612	397	64.9	199	32.5	16	2.6	566	365	64.5	175	30.9	26	4.6	562	225	40.0	294	52.3	43	7.7
Test Materials Modifications		13,608	9,391	69.0	4,084	30.0	133	1.0	13,722	9,955	72.5	3,354	24.4	413	3.0	13,715	6,916	50.4	6,071	44.3	728	5.3
Test Procedures Modifications																						
Section 504 Accommodations		1,299	322	24.8	877	67.5	100	7.7	1,292	359	27.8	673	52.1	260	20.1	1,290	205	15.9	790	61.2	295	22.9
Setting Accommodations		1,415	349	24.7	948	67.0	118	8.3	1,411	387	27.4	730	51.7	294	20.8	1,409	213	15.1	873	62.0	323	22.9
Scheduling Accommodations		55	24	43.6	25	45.5	6	10.9	39	23	59.0	14	35.9	2	5.1	39	18	46.2	18	46.2	3	7.7
Test Materials Modifications		867	244	28.1	560	64.6	63	7.3	865	275	31.8	436	50.4	154	17.8	862	181	21.0	505	58.6	176	20.4
Test Procedures Modifications																						

**TABLE 3.2**  
**Setting Accommodations for Special Education Students and Section 504 Students**

	NUMBER OF STUDENTS ENROLLED	LANGUAGE ARTS LITERACY						MATHEMATICS						SCIENCE								
		PARTIALLY PROFICIENT (100 – 199)		PROFICIENT (200 – 249)		ADVANCED PROFICIENT (250 – 300)		PARTIALLY PROFICIENT (100 – 199)		PROFICIENT (200 – 249)		ADVANCED PROFICIENT (250 – 300)		PARTIALLY PROFICIENT (100 – 199)		PROFICIENT (200 – 249)		ADVANCED PROFICIENT (250 – 300)				
		NO.	%	NO.	%	NO.	%	NO.	%	NO.	%	NO.	%	NO.	%	NO.	%	NO.	%			
<b>Special Education</b>	<b>18,197</b>	<b>11,390</b>	<b>67.1</b>	<b>5,375</b>	<b>31.7</b>	<b>206</b>	<b>1.2</b>	<b>17,190</b>	<b>12,234</b>	<b>71.2</b>	<b>4,339</b>	<b>25.2</b>	<b>617</b>	<b>3.6</b>	<b>17,173</b>	<b>8,502</b>	<b>49.5</b>	<b>7,630</b>	<b>44.4</b>	<b>1,041</b>	<b>6.1</b>	
<b>Setting Accommodations</b>		<b>14,654</b>	<b>10,019</b>	<b>68.4</b>	<b>4,485</b>	<b>30.6</b>	<b>150</b>	<b>1.0</b>	<b>14,780</b>	<b>10,667</b>	<b>72.2</b>	<b>3,634</b>	<b>24.6</b>	<b>479</b>	<b>3.2</b>	<b>14,765</b>	<b>7,385</b>	<b>50.0</b>	<b>6,558</b>	<b>44.4</b>	<b>822</b>	<b>5.6</b>
A. Auditory Impaired	125	78	62.4	39	31.2	8	6.4	124	80	64.5	36	29.0	8	6.5	122	64	52.5	47	38.5	11	9.0	
B. Other Health Impaired	2,037	1,193	58.6	803	39.4	41	2.0	2,049	1,242	60.6	676	33.0	131	6.4	2,049	738	36.0	1,089	53.1	222	10.8	
C. Communication Impaired	925	738	79.8	184	19.9	3	0.3	927	767	82.7	150	16.2	10	1.1	931	593	63.7	321	34.5	17	1.8	
D. Emotionally Disturbed	953	704	73.9	245	25.7	4	0.4	1,002	745	74.4	230	23.0	27	2.7	1,004	517	51.5	433	43.1	54	5.4	
E. Cognitively Impaired	309	295	95.5	14	4.5	0	0.0	309	300	97.1	9	2.9	0	0.0	306	272	88.9	34	11.1	0	0.0	
F. Multiply Disabled	1,524	1,282	84.1	235	15.4	7	0.5	1,543	1,308	84.8	207	13.4	28	1.8	1,544	1,036	67.1	449	29.1	59	3.8	
G. Traumatic Brain Injury	62	49	79.0	12	19.4	1	1.6	62	53	85.5	9	14.5	0	0.0	62	37	59.7	23	37.1	2	3.2	
H. Orthopedically Impaired	25	7	28.0	14	56.0	4	16.0	24	12	50.0	9	37.5	3	12.5	24	4	16.7	14	58.3	6	25.0	
I. Specific Learning Disability	8,416	5,489	65.2	2,854	33.9	73	0.9	8,467	5,995	70.8	2,229	26.3	243	2.9	8,456	4,003	47.3	4,038	47.8	415	4.9	
J. Social Maladjustment	7	3	42.9	3	42.9	1	14.3	8	4	50.0	2	25.0	2	25.0	8	3	37.5	3	37.5	2	25.0	
K. Visually Impaired	16	7	43.8	6	37.5	3	18.8	4	3	75.0	1	25.0	0	0.0	3	2	66.7	0	0.0	1	33.3	
L. Speech-Language Services Only	40	15	37.5	21	52.5	4	10.0	41	18	43.9	17	41.5	6	14.6	41	12	29.3	25	61.0	4	9.8	
M. Autistic	213	157	73.7	55	25.8	1	0.5	219	139	63.5	59	26.9	21	9.6	214	103	48.1	82	38.3	29	13.6	
N. Multiple Grids and Default	2	2	100.0	0	0.0	0	0.0	1	1	100.0	0	0.0	0	0.0	1	1	100.0	0	0.0	0	0.0	
Section 504		1,299	322	24.8	877	67.5	100	7.7	1,292	359	27.8	673	52.1	260	20.1	1,290	205	15.9	790	61.2	295	22.9

TABLE 3.3

Scheduling Accommodations for Special Education Students and Section 504 Students

	NUMBER OF STUDENTS ENROLLED	LANGUAGE ARTS LITERACY						MATHEMATICS						SCIENCE							
		PARTIALLY PROFICIENT (100 - 199)		PROFICIENT (200 - 249)		ADVANCED PROFICIENT (250 - 300)		PARTIALLY PROFICIENT (100 - 199)		PROFICIENT (200 - 249)		ADVANCED PROFICIENT (250 - 300)		PARTIALLY PROFICIENT (100 - 199)		PROFICIENT (200 - 249)		ADVANCED PROFICIENT (250 - 300)			
		NO.	%	NO.	%	NO.	%	NO.	%	NO.	%	NO.	%	NO.	%	NO.	%	NO.	%		
<b>Special Education</b>	<b>18,197</b>	<b>11,390</b>	<b>67.1</b>	<b>5,375</b>	<b>31.7</b>	<b>206</b>	<b>1.2</b>	<b>17,190</b>	<b>12,234</b>	<b>71.2</b>	<b>4,339</b>	<b>25.2</b>	<b>617</b>	<b>3.6</b>	<b>8,502</b>	<b>49.5</b>	<b>7,630</b>	<b>44.4</b>	<b>1,041</b>	<b>6.1</b>	
<b>Scheduling Accommodations</b>	<b>14,819</b>	<b>10,064</b>	<b>67.9</b>	<b>4,603</b>	<b>31.1</b>	<b>152</b>	<b>1.0</b>	<b>14,942</b>	<b>10,740</b>	<b>71.9</b>	<b>3,707</b>	<b>24.8</b>	<b>495</b>	<b>3.3</b>	<b>7,442</b>	<b>49.9</b>	<b>6,636</b>	<b>44.5</b>	<b>849</b>	<b>5.7</b>	
A. Auditorily Impaired	126	78	61.9	40	31.7	8	6.3	125	80	64.0	37	29.6	8	6.4	64	52.0	48	39.0	11	8.9	
B. Other Health Impaired	2,083	1,221	58.6	818	39.3	44	2.1	2,095	1,262	60.2	696	33.2	137	6.5	759	36.2	1,106	52.8	229	10.9	
C. Communication Impaired	925	742	80.2	180	19.5	3	0.3	927	770	83.1	147	15.9	10	1.1	596	64.0	319	34.3	16	1.7	
D. Emotionally Disturbed	962	709	73.7	250	26.0	3	0.3	1,005	746	74.2	231	23.0	28	2.8	1,008	51.2	436	43.3	56	5.6	
E. Cognitively Impaired	305	289	94.8	16	5.2	0	0.0	307	297	96.7	10	3.3	0	0.0	304	269	88.5	35	11.5	0	0.0
F. Multiply Disabled	1,514	1,277	84.3	229	15.1	8	0.5	1,533	1,302	84.9	203	13.2	28	1.8	1,532	1,037	67.7	433	28.3	62	4.0
G. Traumatic Brain Injury	65	50	76.9	14	21.5	1	1.5	65	56	86.2	9	13.8	0	0.0	66	40	60.6	24	36.4	2	3.0
H. Orthopedically Impaired	27	8	29.6	15	55.6	4	14.8	26	13	50.0	10	38.5	3	11.5	26	6	23.1	14	53.8	6	23.1
I. Specific Learning Disability	8,531	5,507	64.6	2,952	34.6	72	0.8	8,586	6,051	70.5	2,285	26.6	250	2.9	8,576	4,036	47.1	4,109	47.9	431	5.0
J. Social Maladjustment	5	2	40.0	2	40.0	1	20.0	5	2	40.0	2	40.0	1	20.0	5	1	20.0	3	60.0	1	20.0
K. Visually Impaired	18	7	38.9	8	44.4	3	16.7	4	3	75.0	1	25.0	0	0.0	3	2	66.7	0	0.0	1	33.3
L. Speech-Language Services Only	41	15	36.6	22	53.7	4	9.8	42	18	42.9	18	42.9	6	14.3	42	12	28.6	26	61.9	4	9.5
M. Autistic	215	157	73.0	57	26.5	1	0.5	221	139	62.9	58	26.2	24	10.9	216	103	47.7	83	38.4	30	13.9
N. Multiple Grids and Default	2	2	100.0	0	0.0	0	0.0	1	1	100.0	0	0.0	0	0.0	1	1	100.0	0	0.0	0	0.0
Section 504	1,415	349	24.7	948	67.0	118	8.3	1,411	387	27.4	730	51.7	294	20.8	213	15.1	873	62.0	323	22.9	

TABLE 3.4

Test Materials Modifications for Special Education Students and Section 504 Students

	NUMBER OF STUDENTS ENROLLED	LANGUAGE ARTS LITERACY						MATHEMATICS						SCIENCE							
		PARTIALLY PROFICIENT (100 - 199)		PROFICIENT (200 - 249)		ADVANCED PROFICIENT (250 - 300)		PARTIALLY PROFICIENT (100 - 199)		PROFICIENT (200 - 249)		ADVANCED PROFICIENT (250 - 300)		PARTIALLY PROFICIENT (100 - 199)		PROFICIENT (200 - 249)		ADVANCED PROFICIENT (250 - 300)			
		NO.	%	NO.	%	NO.	%	NO.	%	NO.	%	NO.	%	NO.	%	NO.	%	NO.	%		
<b>Special Education</b>	<b>18,197</b>	<b>11,390</b>	<b>67.1</b>	<b>5,375</b>	<b>31.7</b>	<b>206</b>	<b>1.2</b>	<b>17,190</b>	<b>12,234</b>	<b>71.2</b>	<b>4,339</b>	<b>25.2</b>	<b>617</b>	<b>3.6</b>	<b>8,502</b>	<b>49.5</b>	<b>7,630</b>	<b>44.4</b>	<b>1,041</b>	<b>6.1</b>	
<b>Test Materials Modifications</b>	<b>612</b>	<b>397</b>	<b>64.9</b>	<b>199</b>	<b>32.5</b>	<b>16</b>	<b>2.6</b>	<b>566</b>	<b>365</b>	<b>64.5</b>	<b>175</b>	<b>30.9</b>	<b>26</b>	<b>4.6</b>	<b>225</b>	<b>40.0</b>	<b>294</b>	<b>52.3</b>	<b>43</b>	<b>7.7</b>	
A. Auditorily Impaired	4	3	75.0	0	0.0	1	25.0	3	2	66.7	1	33.3	0	0.0	3	66.7	1	33.3	0	0.0	
B. Other Health Impaired	101	57	56.4	41	40.6	3	3.0	99	61	61.6	28	28.3	10	10.1	97	37	38.1	51	52.6	9	9.3
C. Communication Impaired	39	28	71.8	11	28.2	0	0.0	37	27	73.0	9	24.3	1	2.7	36	17	47.2	17	47.2	2	5.6
D. Emotionally Disturbed	61	40	65.6	21	34.4	0	0.0	63	41	65.1	19	30.2	3	4.8	66	28	42.4	33	50.0	5	7.6
E. Cognitively Impaired	13	13	100.0	0	0.0	0	0.0	12	12	100.0	0	0.0	0	0.0	12	11	91.7	1	8.3	0	0.0
F. Multiply Disabled	85	69	81.2	16	18.8	0	0.0	66	54	81.8	11	16.7	1	1.5	64	36	56.3	25	39.1	3	4.7
G. Traumatic Brain Injury	3	3	100.0	0	0.0	0	0.0	2	2	100.0	0	0.0	0	0.0	2	1	50.0	1	50.0	0	0.0
H. Orthopedically Impaired	7	2	28.6	4	57.1	1	14.3	5	2	40.0	3	60.0	0	0.0	5	1	20.0	2	40.0	2	40.0
I. Specific Learning Disability	271	168	62.0	96	35.4	7	2.6	262	155	59.2	98	37.4	9	3.4	261	85	32.6	156	59.8	20	7.7
J. Social Maladjustment	0	0	0.0	0	0.0	0	0.0	0	0	0.0	0	0.0	0	0.0	0	0	0.0	0	0.0	0	0.0
K. Visually Impaired	14	5	35.7	7	50.0	2	14.3	3	3	100.0	0	0.0	0	0.0	2	2	100.0	0	0.0	0	0.0
L. Speech-Language Services Only	2	0	0.0	0	0.0	2	100.0	2	0	0.0	0	0.0	2	100.0	2	0	0.0	1	50.0	1	50.0
M. Autistic	12	9	75.0	3	25.0	0	0.0	12	6	50.0	6	50.0	0	0.0	12	5	41.7	6	50.0	1	8.3
N. Multiple Grids and Default	0	0	0.0	0	0.0	0	0.0	0	0	0.0	0	0.0	0	0.0	0	0	0.0	0	0.0	0	0.0
Section 504	55	24	43.6	25	45.5	6	10.9	39	23	59.0	14	35.9	2	5.1	18	46.2	18	46.2	3	7.7	

**TABLE 3.5**  
**Test Procedures Modifications for Special Education Students and Section 504 Students**

	NUMBER OF STUDENTS ENROLLED	LANGUAGE ARTS LITERACY						MATHEMATICS						SCIENCE								
		PARTIALLY PROFICIENT (100 - 199)		PROFICIENT (200 - 249)		ADVANCED PROFICIENT (250 - 300)		PARTIALLY PROFICIENT (100 - 199)		PROFICIENT (200 - 249)		ADVANCED PROFICIENT (250 - 300)		PARTIALLY PROFICIENT (100 - 199)		PROFICIENT (200 - 249)		ADVANCED PROFICIENT (250 - 300)				
		NO.	%	NO.	%	NO.	%	NO.	%	NO.	%	NO.	%	NO.	%	NO.	%	NO.	%			
<b>Special Education</b>	<b>18,197</b>	<b>16,971</b>	<b>11,390</b>	<b>67.1</b>	<b>5,375</b>	<b>31.7</b>	<b>206</b>	<b>1.2</b>	<b>17,190</b>	<b>12,234</b>	<b>71.2</b>	<b>4,339</b>	<b>25.2</b>	<b>617</b>	<b>3.6</b>	<b>17,173</b>	<b>8,502</b>	<b>49.5</b>	<b>7,630</b>	<b>44.4</b>	<b>1,041</b>	<b>6.1</b>
<b>Test Procedures Modifications</b>		<b>13,608</b>	<b>9,391</b>	<b>69.0</b>	<b>4,084</b>	<b>30.0</b>	<b>133</b>	<b>1.0</b>	<b>13,722</b>	<b>9,955</b>	<b>72.5</b>	<b>3,354</b>	<b>24.4</b>	<b>413</b>	<b>3.0</b>	<b>13,715</b>	<b>6,916</b>	<b>50.4</b>	<b>6,071</b>	<b>44.3</b>	<b>728</b>	<b>5.3</b>
A. Auditorily Impaired		122	77	63.1	37	30.3	8	6.6	121	77	63.6	37	30.6	7	5.8	119	61	51.3	48	40.3	10	8.4
B. Other Health Impaired		1,872	1,112	59.4	719	38.4	41	2.2	1,882	1,140	60.6	628	33.4	114	6.1	1,884	691	36.7	987	52.4	206	10.9
C. Communication Impaired		863	699	81.0	161	18.7	3	0.3	863	720	83.4	136	15.8	7	0.8	868	552	63.6	302	34.8	14	1.6
D. Emotionally Disturbed		877	669	76.3	207	23.6	1	0.1	914	692	75.7	201	22.0	21	2.3	920	482	52.4	393	42.7	45	4.9
E. Cognitively Impaired		282	269	95.4	13	4.6	0	0.0	286	277	96.9	9	3.1	0	0.0	283	253	89.4	30	10.6	0	0.0
F. Multiply Disabled		1,442	1,224	84.9	212	14.7	6	0.4	1,461	1,251	85.6	186	12.7	24	1.6	1,461	999	68.4	411	28.1	51	3.5
G. Traumatic Brain Injury		62	47	75.8	14	22.6	1	1.6	62	53	85.5	9	14.5	0	0.0	62	37	59.7	23	37.1	2	3.2
H. Orthopedically Impaired		23	7	30.4	13	56.5	3	13.0	22	10	45.5	10	45.5	2	9.1	22	5	22.7	11	50.0	6	27.3
I. Specific Learning Disability		7,814	5,116	65.5	2,634	33.7	64	0.8	7,866	5,578	70.9	2,073	26.4	215	2.7	7,856	3,721	47.4	3,768	48.0	367	4.7
J. Social Maladjustment		4	3	75.0	1	25.0	0	0.0	5	4	80.0	1	20.0	0	0.0	5	3	60.0	2	40.0	0	0.0
K. Visually Impaired		16	5	31.3	8	50.0	3	18.8	3	2	66.7	1	33.3	0	0.0	2	1	50.0	0	0.0	1	50.0
L. Speech-Language Services Only		31	11	35.5	17	54.8	3	9.7	32	15	46.9	12	37.5	5	15.6	32	10	31.3	19	59.4	3	9.4
M. Autistic		198	150	75.8	48	24.2	0	0.0	204	135	66.2	51	25.0	18	8.8	200	100	50.0	77	38.5	23	11.5
N. Multiple Grids and Default		2	2	100.0	0	0.0	0	0.0	1	1	100.0	0	0.0	0	0.0	1	1	100.0	0	0.0	0	0.0
Section 504		867	244	28.1	560	64.6	63	7.3	865	275	31.8	436	50.4	154	17.8	862	181	21.0	505	58.6	176	20.4

## CHAPTER 4: SCORING

### 4.1 Multiple-choice Items

Each multiple-choice item contributes one point to the total raw score for each content-area test. Responses for multiple-choice items are machine scored. The score points of multiple-choice items received for a content area are the total number of multiple-choice items answered correctly. For the Mathematics and Science content areas and the Language Arts reading component, the total score points of multiple-choice items are combined with the total number of points from the open-ended items for a student's score. For Language Arts Literacy, the reading component score points are added to score points received from the open-ended scoring of the two writing tasks which compose the writing component.

### 4.2 Open-ended Items

During April and May of 2007, Measurement Incorporated (MI) under subcontract to Pearson scored the student writing responses, and the reading, mathematics, and science open-ended items. MI has a staff of highly-trained scorers who must have at least a bachelor's degree and who must undergo rigorous and ongoing training and monitoring during the scoring process. Ten percent (10%) of the reading, mathematics, and science open-ended responses were read by a second rater. Each writing prompt was read independently by two scorers. If the two scorers disagreed by more than one point, a third scorer evaluated the response. Appendix A presents information about how the three scores are resolved for each of the content areas.

Table 4.1 shows the number of writing responses and open-ended items scored for the operational test.

**TABLE 4.1**

### **Number of Writing Prompts and Open-ended Items Scored**

<b>Content Area</b>	<b>Number of Writing Prompts and Open-ended Items Scored</b>
<i>Language Arts Literacy</i>	938,659
<i>Reading</i>	501,621
<i>Writing</i>	437,038
<i>Speculate</i>	218,587
<i>Persuade</i>	218,451
<i>Mathematics</i>	752,279
<i>Science</i>	377,650
<b>TOTAL</b>	<b>2,068,588</b>

### **Scorer Selection**

MI's senior project managers work closely with Content Coordinators in the Office of State Assessments. Current procedures for scoring the GEPA open-ended and writing responses are consistent with those used since the inception of a performance-based writing component in the New Jersey statewide assessment. Scoring of the open-ended and writing responses is monitored by trained, experienced personnel who have met the same rigorous standards established with the initial holistic scoring study conducted in 1986.

For selecting team leaders, MI's management staff and scoring directors reviewed the files of all returning staff who have previously scored the GEPA. The MI staff looked for people who were experienced team leaders with a record of good performance on previous projects and also considered scorers who have been recommended for promotion to the team leader position.

Many of the MI scorers have repeatedly scored the GEPA for previous test administrations. MI's procedures for selecting new scorers are very thorough. After advertising in local newspapers, with the job service, and elsewhere, and receiving applications, staff in MI's human resources department review applications and schedule interviews for qualified applicants. Qualified applicants are those with a four-year college degree in English, language arts, education, mathematics, science, or a related field. Each qualified applicant must pass an interview by experienced MI staff, write an acceptable essay, and receive good recommendations from references. All the information about each applicant is reviewed before offering employment.

MI is an equal opportunity employer that actively recruits minority staff. Historically, their temporary staff on major projects averages about 70 percent female, 30 percent male, 76 percent Caucasian, and 24 percent minority. MI strongly opposes illegal discrimination against any employee or applicant for employment with respect to hire, tenure, terms, conditions, or privileges of employment, or any matter directly or indirectly related to employment because of race, color, religion, sex, age, handicap, national origin, or ancestry.

### **Rangefinding**

Rangefinding is one of the most important elements of the scoring process. Rangefinding meetings provide an opportunity for finalizing scoring rubrics (in content areas with specific item rubrics) and making scoring decisions and interpretations regarding scoring issues before team leader and scorers' training begins. (See Appendix A for rubrics.) It is important that as many of the item-specific problems as possible be resolved prior to scorers' training so that scoring decisions can be made during scoring.

After consulting with Pearson to determine when the first “live” student responses would be available, MI scheduled a rangefinding meeting in Durham, other MI sites (operational test), and New Jersey (field test) to establish “true” scores for a representative sample of open-ended items. At this meeting, Office of State Assessments staff members, content committee members, and the MI project leaders read and scored 60-225 responses, which exemplified various points of the rubric and score scale. The number of responses varied according to the content area and score scale. The responses were selected from a broad range of New Jersey school districts in order to ensure that the sample was representative of overall student performance. Rangefinding took from two to six days per content area, depending on the number of items tested.

### **Development of Scoring Guides**

After the rangefinding responses were discussed and received a final score, MI used the selected responses to develop scoring guides, training sets (practice papers), and/or qualifying sets for each content area. Scoring guides consisted of three or more examples of each score point in score point order. In some content areas, the papers were annotated. Training and qualifying sets were clearly anchored papers in random score point order. Sufficient copies were made so that all scoring directors, team leaders, and scorers had their own copy during training and scoring.

### **Team Leader Training and Qualifying**

After the anchor papers, training, and/or qualifying papers were identified and finalized, team leader training began. The scoring director (for each content area or writing type) conducted training for the team leaders. Procedures were similar to those for training scorers (see below) but were more comprehensive, dealing with resolution of discrepant scores, identification of nonscorable responses, unusual prompt treatment, alert situation responses (e.g., child-in-danger), and other duties performed only by team leaders. The team leaders carefully prepared notes on the training papers in preparation for discussion with the scorers, and the scoring director counseled team leaders on training techniques and application of the rubric.

Team leaders assisted in training scorers in team discussions of training sets, and were responsible for distributing, collecting, and accounting for training packets and sample papers during each scoring session. During scoring, team leaders responded to questions, spot-checked reader packets, and counseled scorers having difficulty with the criteria.

Team leaders also administered the quality control (validity) set, monitored the scoring patterns of each reader throughout the project, conducted retraining as necessary, performed some resolution readings, and maintained a professional working environment. The validity sets were generally selected by the team leaders and scoring director for each content area prior to reader training.

Team leader training lasted from two to four days. Team leaders generally worked 7.75 hours per day, excluding breaks. They set up the room prior to reader arrival each day and met with scoring directors after scoring each day.

### **Scorer Training and Qualifying**

All scorers were trained using the scoring guides and rubrics, training papers, and/or qualifying papers selected during the rangefinding meetings. Scorers were assigned to a scoring group consisting of one team leader and 10-12 scorers. Each scorer was assigned an individual number for easy identification of their scoring work throughout the scoring session.

After the contracts and nondisclosure forms were signed and the introductory remarks given, training began. Scorer training followed the same format as team leader training except that scorers were not required to annotate each paper in the training sets, although they were encouraged to take notes. The scoring director presented the writing or open-ended item task and introduced the guide, then discussed, room-wide, each score point. This presentation was followed by practice scoring on the training sets. Each scorer was given a clean copy of the scoring guide and training sets, as well as a monitor sheet on which to record training set scores. Because it is easy in a large group to overlook a shy scorer who may be having difficulty, scorers did break into teams to score and discuss the papers in the training sets. This arrangement provided scorers an opportunity to discuss any possible points of confusion or problems in understanding the criteria.

Team leaders collected the monitor sheets after the scoring of each training set and recorded results on a customized log, which was examined by the scoring director to determine which papers were giving scorers difficulty. The scoring director also “floated” from team to team, listening to the team leaders’ explanations and adding additional information when necessary. If a particular paper or type of paper seemed to cause difficulty across teams, the problem was discussed room-wide to ensure that everyone heard the same explanation.

Like team leaders, scorers must demonstrate their ability to score accurately by attaining the agreement percentage established by the New Jersey Department of Education before they may score packets of “live” papers. Any scorer unable to meet these standards was dismissed. All scorers understand this stipulation when they are hired.

Training was carefully orchestrated so that scorers understood how to apply the rubric in scoring the papers, learned how to reference the scoring guide, developed the flexibility needed to deal with a variety of responses, and retained the consistency needed to score all papers accurately.

Scorers were trained to recognize and flag nonscorable responses (fragment, off-topic, not English, no response) and “alert” papers (e.g., suspicion of child abuse) so that these papers could be handled in the correct manner. Alert papers were scored, but then forwarded to the scoring director for review. If the scoring director agreed that the student’s own words specifically stated a situation that qualified as an alert or reflected a potential risk situation for a child, the paper was copied and sent to the Office of State Assessments for follow-up with school district personnel. Alert papers are flagged if they reflect potential abuse, emotional or psychological difficulty, dangerous thoughts, or possible plagiarism.

In addition to completing all of the initial training and qualifying, a significant amount of time was allotted for demonstrations of paper flow, explanations of “alerts” and “flagging,” and instructions about other procedures which were necessary for the conduct of a smooth project. Scorer training lasted from two to five days. Scorers generally worked 7.0 hours per day, excluding breaks.

### **Scoring Procedures and Paper Flow**

Each student response was scored by two independent scorers using the scoring scale developed and approved for those items. If the two assigned scores differed by more than one point, the paper was returned for a third “resolution” reading by team leaders or scoring directors. Information about how the three scores were resolved appears in Appendix A.

Before opening a packet, scorers began by writing their assigned reader numbers, as well as the date, on the front of their packet envelope. The stapled packet of papers and the appropriate monitor sheet (first or second reading) was then removed from the envelope. Scorers checked the packet number on the header sheet against the number on the monitor sheet for agreement, and then recorded their scorer identification numbers in the designated space on the scannable monitor sheet. The scorer decided on the score, and the assigned scores are recorded in the appropriate spaces provided on the monitor. As scorers progressed through a packet, they checked each paper’s student ID number against the number printed on the monitor sheet. If there was a discrepancy, the packet was flagged for the scoring director to check.

As a scorer completed a packet of papers, he or she returned it to the envelope and gave it to the team leader, along with the monitor sheet. The clerical aide picked up completed packets and monitor sheets, and redistributed the packets for second readings.

The packet proceeded to the second reading stage while the first reading scores were being scanned. The procedure for the second reading was the same as that for the first reading, except that the second scorer used the second scoring monitor sheet in the envelope. At no time does the second scorer have access to the scores given by the first scorer. As with the first scoring monitors, the second monitors were scanned and the scores merged into the database.

After the second scores were entered, they were matched with the first scores already in the database. When scores differed by more than one point on any response, the response was classified as “discrepant,” a third scoring list by packet and response number was printed, and the response was returned for a third independent reading. After the clerical aide returned the packet to the scoring room, the scoring director located the papers needing a third reading and followed the normal scoring procedures. The third score was scanned in the same manner as the first two scores. The packet was returned to the warehouse and refiled.

### Scorer Monitoring

Scorers were monitored in several ways. Team leaders answered scorers' questions, using the guide and training papers as examples. They also read behind their team members by reviewing packets after they were turned in, looking for papers that might merit discussion with the scorer. In addition, every day the scoring director and team leaders received the printout of the scorer statistics—including the scorers' perfect, adjacent, and resolution agreement with other scorers, and the scorers' score point distribution. In this way, the scoring director and team leader can look at any one scorer, team, or the room as a whole and rollover items can be compared to previous years.

### Agreement Between Scorers for the Writing Tasks and Open-Ended Items

Table 4.2 shows the percentages of writing tasks and open-ended items scored with exact agreement, adjacent agreement, and resolution needed.

The Writing cluster within Language Arts Literacy consists of two writing activities:

- writing/speculate task in response to a picture —  
1-6 points, scorer ratings averaged
- writing/persuade task —  
1-6 points, scorer ratings summed

Each writing task is rated by two independent scorers. Of the approximately 220,000 task responses scored for the 2007 administration, 65.1% received exactly the same scores by the scorers and 33.8% received scores that were adjacent. Thus, approximately 98.9% of the task responses required only two scorers. The remaining 1.2% received scores on the writing tasks that differed by more than one point and, therefore, required a third scorer.

All content areas included open-ended items. Ten percent (10%) of these open-ended responses were read by a second scorer. The purpose of a second-rating is to investigate the consistency between scorers. For the Reading open-ended items, the rubric used by the scorers had score points that ranged from 0 to 4. Two Reading open-ended items are presented for each of two reading passages. For reading, over 76% of the responses were assigned a score by a second scorer that was in exact agreement with the first scorer. About 23% of the second ratings were assigned an adjacent score by a second scorer. An adjacent score is a score assigned by the second scorer that is no more than one score point above or below the score assigned by the first scorer.

Six open-ended items were presented for Mathematics. These six items had percents at perfect agreement ranging from 85.9% to 94.1%. Over 90% of the total Mathematics responses were assigned a score by a second scorer that was in exact agreement with the first scorer. More than 9% of the second ratings were assigned an adjacent score by a second scorer.

Three open-ended items were included for Science. These items had a perfect agreement rate ranging from 71.3% to 79.7%. Over 76% of the total Science responses were assigned a score by a second scorer that was in exact agreement with the first scorer. About 20% of the second ratings were assigned an adjacent score by a second scorer.

**TABLE 4.2**  
**Consistency Between Raters Scoring**  
**GEPA Writing Tasks and Open-Ended Items**  
**March 2007**

GEPA Writing Tasks and Open-Ended Items	Percent Raters In Exact Agreement	Percent Raters In Adjacent Agreement	Percent Resolution Needed
<b>Language Arts Literacy</b>			
<b>Writing Total</b>	<b>65.1</b>	<b>33.8</b>	<b>1.2</b>
Writing/Speculate	65.0	33.8	1.2
Writing/Persuade	65.1	33.8	1.1
<b>Reading Total</b>	<b>76.6</b>	<b>22.9</b>	<b>0.6</b>
Open-Ended Item 1	74.0	25.2	0.8
Open-Ended Item 2	74.0	25.3	0.7
Open-Ended Item 3	80.5	19.3	0.2
Open-Ended Item 4	77.8	21.6	0.5
<b>Mathematics</b>			
<b>Mathematics Total</b>	<b>90.1</b>	<b>9.5</b>	<b>0.4</b>
Open-Ended Item 1	87.7	12.2	0.1
Open-Ended Item 2	91.3	8.0	0.7
Open-Ended Item 3	88.4	11.0	0.6
Open-Ended Item 4	93.2	6.7	0.1
Open-Ended Item 5	85.9	13.5	0.6
Open-Ended Item 6	94.1	5.5	0.4
<b>Science</b>			
<b>Science Total</b>	<b>76.6</b>	<b>20.0</b>	<b>3.4</b>
Open-Ended Item 1	78.7	18.5	2.9
Open-Ended Item 2	79.7	19.2	1.0
Open-Ended Item 3	71.3	22.3	6.4

### 4.3 Quality Control Procedures in Data Preparation

Quality control procedures at Pearson begin with the use of the Capability Maturity Model (CMM), a software development management tool. Key process areas of CMM are requirements management, software project planning, software project tracking and oversight, software quality assurance, and software configuration management. Pearson examples of CMM documents include a customer requirements allocation document, a project schedule, functional specifications, a software development project plan, unit test plans, and verification and validation plans. Pearson is certified by an external auditor for CMM Level 4, the second highest level of certification.

After software requirements have been identified, the Pearson software development team prepares project schedules, project plans, functional specifications, and design documents. Pearson begins by creating detailed test plans at both the unit and systems level. A unit test plan is a list of code-unit test cases that are executed and recorded by the software developer. The purpose of the code-unit test process is to ensure that software is developed, maintained, documented, and verified to meet the project requirements for coding and unit testing. As such, the process provides the mechanisms that are necessary to implement the software requirements and design as well as provides code-units quality assurance prior to system test.

After all modules (units) are tested within a system, the CMM process requires a system test. The system test ensures that all the units work together and that outputs from one module match up to the proper inputs for the next module in the system. It also uses expected results to ensure that all requirements have been met. It is important that the system test be performed by a group that is independent of the software development team. This process allows independent verification and interpretation of the requirements. Once the independent testing group has completed the test and given its approval, the system is moved into production mode. It is ready for processing the quality-checking answer documents and files submitted by a quality-checking team.

#### **Scanning and Scoring**

Before actual answer documents are machine-scanned, a comprehensive check of the scanning and scoring system is performed. The software development tester creates test decks of gridded answer documents with specific test criteria. The test decks are designed and gridded to cover all response ranges, ID ranges, blanks, and double grids as well as any other responses used by the GEPA. A file containing the scanned responses is then compared to the expected test results for each document to ensure the scanner is operating correctly. The test decks are processed through the programs for scanning and editing answer documents, and packetizing and printing scoring monitors.

The second check involves processing and quality-checking the first actual answer documents received. The NJDOE Office of State Assessments and Pearson asked approximately 60 districts to return their answer documents early following the test administration so that all test forms could be processed and quality-checked. Also, these early return districts provided the actual student papers for determining score ranges for the writing tasks and open-ended items. Districts were selected to be representative for size and DFG. All information on approximately 60 answer documents was hand checked against the scanned file. In addition, periodically, throughout the processing of the documents, individual answer documents were checked by hand to ensure that scanning was continuing to perform correctly.

### **NJDOE Quality Control of Score Reporting**

NJDOE Office of State Assessments conducted the first round of quality control of multiple-choice items scoring on April 30-May 4, 2007. Pearson printed score sheets for each of approximately 700 students from more than 20 districts selected by the Office of State Assessments for quality control.

Original answer folders for all students in the quality control sample were shipped to the meeting site. Pearson maintained a copy of all answer folders in the quality control sample. Pearson provided the following materials to the Office of State Assessments for the quality control:

1. Scoring masks (punched index and transparency sheets) for all versions of the tests
2. Answer keys for the multiple-choice items
3. Double-grid documentation included a sample of edits for students who marked more than one answer for a multiple-choice item
4. Irregularity reports included all reports dealing with multiple answer folders for students and provided documentation about how these answer folders were merged
5. List of removed items from the Braille and large-print forms
6. List of names of all students taking a Braille or large-print form
7. County-district-school master files with district test coordinators' names and phone numbers
8. Frequency distributions for the student groups, including total, general, LEP, SE, IEP exempts by content area, void counts by reporting category, and Title 1 counts by reporting category

In the two weeks following the first round of quality control, Measurement Incorporated completed scoring the open-ended and essay responses. Assessment and Evaluation Services equated the test forms after which the NJDOE Office of State Assessments and independent reviewers approved the equating procedures and raw score to scale score conversion tables. Pearson staff loaded the conversion tables and produced Cycle I score reports for the quality control sample for review.

The second round of the Office of State Assessments quality control on the Cycle I score reports occurred over three weeks beginning May 22, 2007. At this time the open-ended and essay scores were available.

The multiple-choice, open-ended, and essay item scores for each cluster and total for the three content areas were systematically checked on all Cycle I score reports. Individual Student Reports for all large-print, Braille, and breach students were produced and reviewed.

Calculations for the Total Scale Score Means and the Just Proficient Means (the mean score for all students across the state whose scale scores were 200 on a particular content area) were verified for each cluster in the content areas by the Office of State Assessments staff. Summary statistics included on the School and District Summary Statistics reports were reviewed and approved.

## CHAPTER 5: STANDARD SETTING

### 5.1 Overview of the Process

A proficiency level setting (standard setting) was conducted June 8-11, 1999, to describe and delineate the thresholds of performance that are indicative of Partially Proficient, Proficient, and Advanced Proficient performance for the GEPA Language Arts Literacy and Mathematics. A standard setting study for Science was conducted July 10-12, 2000. Results of these studies were used to formulate recommendations to the Commissioner of Education and the New Jersey State Board of Education for the adoption of the cut scores (i.e., proficiency levels).

The standard setting studies in 1999 and 2000 were conducted by staff from the New Jersey Department of Education, Office of Assessment; Assessment and Evaluation Services; and NCS Pearson. The document, *GEPA Standard Setting Report*, outlines the studies and presents the resulting documentation.

Participants in the standard setting study were chosen because of their qualifications as judges of student performance and content expertise. The judges represented the general population of New Jersey educators. Special care was taken to ensure adequate professional, gender, racial/ethnic, regional, and District Factor Group (DFG) representation on all panels.

A holistic classification method was used for the GEPA standard settings. The judges reviewed student papers sampled to represent the full range of student scores for the March 1999 GEPA administration of the Language Arts Literacy and Mathematics. The judges were asked to classify student work into three categories: Partially Proficient, Proficient, and Advanced Proficient. The judges had the opportunity to review, discuss, and modify their proficiency classifications. Using a logistic regression method, two cut scores were calculated based on judges' classifications. These two cut scores yielded three proficiency levels. Before they finalized their recommended cut scores, the judges examined how their recommended cut scores affected all New Jersey eighth-grade students who took these tests during the first operational administration in 1999.

The methodology and procedures for the Science standard setting study mirrored those used for the Language Arts Literacy and Mathematics standard setting studies. During the Science standard setting in July 2000, judges examined how their recommended cut scores affected all New Jersey eighth-grade students who took the first operational administration of the Science test in 2000.

### 5.2 Procedures

Prior to the standard setting studies, descriptions for Proficient and Advanced Proficient performance were developed by independent panels of eighth-grade language arts, mathematics, and science teachers. The proficiency level descriptors were developed to reflect actual test content. Proficiency level descriptors that are anchored in test content allow for more accurate decisions to be made by the judges. The committees developed the following proficiency level descriptors:

**Language Arts Literacy–Proficiency Level Descriptors**

**Proficient**

Eighth-grade students performing at the proficient level are able to construct meaning as they generate their own texts and work with texts generated by others. Proficient students show an overall understanding of the text at literal and inferential levels. They are able to connect with prior knowledge while interacting with, interpreting, and analyzing text.

In reading exercises, students are able to identify and discuss central themes, supporting details, and organizational structures of text. They can extrapolate and synthesize information, monitor their understanding of text, and identify a purpose for reading. Students at this level are able to identify support for and discuss opinions and conclusions as well as to explain textual conventions and literary elements.

Eighth-grade students proficient in their writing are able to develop a central theme, supporting details, and an organizational structure. They establish and sustain a purpose for writing and elaborate on information as they monitor development of text. Students at this level are able to provide support for opinions and conclusions and to use textual and literary elements appropriately.

**Advanced Proficient**

Eighth-grade students performing at the advanced level are able to construct and extend meaning as they generate their own texts and work with texts generated by others. Advanced students show a sophisticated understanding of abstract themes and ideas that build a text and extend information. They are able to connect with prior knowledge while interacting with, interpreting, analyzing, and critiquing text.

In addition to consistently demonstrating the qualities outlined for a proficient student, the advanced student will demonstrate the ability to synthesize, analyze, and evaluate written text. Students at this level are able to manipulate understanding and will show a high degree of sustained control over textual conventions and literary elements.

***Mathematics–Proficiency Level Descriptors******Proficient***

The student performing at the proficient level demonstrates evidence of conceptual understanding and of procedural and analytic skills. The student demonstrates the ability to apply mathematical skills and knowledge to theoretical and real-world situations. In addition, the student communicates the required skills and makes connections within and among the mathematical content areas.

The student at this level demonstrates a thorough understanding of basic arithmetic operations—an understanding sufficient for problem solving in practical situations. The student understands the connections between fractions, decimals, percents, and other mathematic topics.

The student understands and applies geometric properties and spatial relationships; applies the principles of similarity, symmetry, and coordinate geometry: interprets data and graphs; determines probabilities; applies the concepts and methods of discrete mathematics, and uses algebraic concepts and processes.

***Advanced Proficient***

The student performing at the advanced level demonstrates clear and consistent evidence of thorough conceptual understanding, and of procedural and analytic skills. The student consistently demonstrates the qualities outlined for proficient performance. In addition, the student at the advanced level demonstrates the use of abstract thinking and provides explanations that are consistently clear and thorough.

**Science–Proficiency Level Descriptors**

**Proficient**

The proficient student can recognize the structural levels of living things. This student knows that some traits of organisms are beneficial and some detrimental. This student can interpret visual and textual data to understand the relationship within a food web and the interdependence of living and nonliving systems.

The proficient student can recognize the effect force has on an object, trace the flow of energy through a system, and use the properties of matter to identify and separate materials. This student can understand different types of energy and use information from data charts to interpret relationships and predict outcomes.

The proficient student can recognize the existence of a relationship between the moon and tides, recognize the different characteristics of the planets in the solar system, and understand the natural forces that change the surface of the Earth, including chemical and physical weathering.

**Advanced Proficient**

The advanced proficient student can support scientific conclusions with valid contextual and visual data and make predictions based on the interactions of living things. This student is able to use interpretive skills to analyze visual and textual data in order to solve problems dealing with the application of force and energy.

The advanced proficient student understands the difference between types of energy waves and can recognize and apply experimental principles and empirical data.

The advanced proficient student can recognize the nature of the tides' relationship to Earth, Sun, and moon; interpret topographical maps; and identify the steps in the process of weathering and erosion.

***Judge Selection Process and Criteria***

The standard setting process relied on expert judgments. Therefore, nominations were solicited from school districts for teachers or administrators representing excellence in the teaching profession in terms of knowledge of content area, knowledge of eighth-grade students' skills and abilities, and some understanding of assessment procedures. It was considered critical that these judges represent the more general body of expert New Jersey public school educators. Special care was taken to select judges who were representative of the various District Factor Groups (DFGs) within the state. Additionally, districts were specifically asked to include special education, ESL, and bilingual teachers among their nominees. Districts were also encouraged to nominate members of underrepresented minority populations, e.g., African American or Hispanic, in order to ensure an appropriate diverse representation of statewide populations. Other criteria used in the selection process included number of years teaching experience, the level of content knowledge and student understanding possessed by the nominees, and active participation in content-area professional associations.

Teachers, educators, and content-area experts selected as judges exemplified the required content-area knowledge, teaching experience, and/or understanding of students necessary for an appropriate and comprehensive standard setting study. Each panelist participating in the process represented the knowledge and understanding of his or her peers throughout the course of the process, lending a balance between diverse opinion and consensus.

A concerted effort was made to balance each content-area panel on the basis of county representation, urban representation, representation of schools serving various sizes of populations, gender, and race/ethnicity. The overarching goal of consensus in this forum was not the unanimous agreement of all parties, but the bringing together of individual divergent experiences to form a common understanding of student performance in a content-area that is truly larger, and broader, than its individual parts. The judges selected for the standard setting study represented the same diversity of people and demographics as the students being assessed.

***Holistic/Paper Sorting Methodology***

The judges' task was to classify student work into one of three performance categories defined to capture levels of performance as expressed by the Partially Proficient, Proficient, and Advanced Proficient categories. The method was holistic in that the judges considered the whole of an individual student's open-ended and multiple-choice responses, i.e., all the items of a particular student for a content area. With the holistic sorting method, the judges reviewed folders of student papers sampled to represent the full range of scores and were asked to sort these folders into three performance levels as represented by the quality of the students' work. An outline of the standard setting procedures follows:

### Overview of the 8-Step Plan

#### *Large-Group Session*

The standard setting study began with a large-group session. All judges and participants listened to introductory comments and directions for the three-day meeting. The definitions of the standards, their purpose, and ultimate use were discussed. This session was designed to provide a common orientation to judges across content areas.

#### **Step 1 – Description of the Standard Setting Process**

*Judges worked in their own content area and in separate rooms for the remainder of the process. Step 1 provided the judges with an introduction to the process, their role in the process, and a review of the purpose of the standards.*

- *Introductions*
- *Judge Selection Process and Criteria*
- *Purpose of the Standards*
- *Standard Setting Process*
- *Review of the Agenda*
- *Administrative Tasks*

#### **Step 2 – Review of the Assessment Material**

*Judges became familiar with the assessment at this point. They took the assessment under standardized conditions to get a feel for the experience and content. Judges were also introduced to the content validity evidence for the assessment and the open-ended scoring procedures.*

- *Review of Test Content*
- *Brief Description of the Assessment Development Process*
- *Administration of the Assessment to Judges*
- *Scoring the Assessment*

#### **Step 3 – Defining the Standards**

*Step 3 introduced judges to the definitions of the standards. Judges used exercises to brainstorm student work which typified the definitions for each standard. Judges did not write or re-write the definitions at this time. This step only served to familiarize judges with the definitions, which were previously determined, and to help the judges think about students who are at each standard.*

- *Definitions of Student Performance Standards*
- *Interpretation of Proficient Performance*
- *Interpretation of Advanced Proficient Performance*
- *Summary of Student Performance Levels*

#### **Step 4 – Introduction of the Standard Setting Process**

*Step 4 introduced the specific process to the judges. They practiced reviewing student work and sorting student work into three levels of performance – poor, medium, and high. Judges were provided with information about which multiple-choice items were answered correctly on each sample. In addition, scoring rubrics for the open-ended items were reviewed to facilitate the judgment process for the open-ended items.*

- *Description of the Holistic Sorting Method*
- *Summary of the Standard Setting Process*
- *Process Check-off*

**Step 5 – Round 1: Holistic Classification of a Wide Range of Student Papers**

Judges were instructed in the process of completing the rating sheets. Then, judges were given a set of 33 student papers to classify.

The 33 papers were selected to represent the complete range of test scores for each content area. The raw score distribution for a content area was divided into 11 equal intervals. For each interval, three papers were selected to represent a high score, middle score, and low score within the interval. Judges classified each student work sample as representing an Advanced Proficient, Proficient, or Partially Proficient student by the definitions. Judges recorded their classifications on their rating sheets.

Rating sheets were collected and tabulated with results presented to the judges. Classification frequencies for each paper number were shown to the judges. Judges met in small groups to discuss their classifications. Following the discussions, judges were allowed to make changes to their classifications of the student work on their rating sheets.

- Distribution of Rating Sheets and Instructions
- Classification of Papers (Round 1.1)
- Discussion of Judges' Ratings
- Review of Classifications (Round 1.2)

**Step 6 – Round 2: Holistic Classification of a Targeted Range of Student Papers**

Based on the judges' ratings from Step 5, preliminary cut scores for Advanced Proficient and Proficient were determined using a logistic response model regression of paper scores upon classification decisions. Two papers from each score point at the preliminary cut score and in a range of 5 score points above and below that cut score were selected. Approximately 22 papers were selected to target the borderline between Advanced Proficient and Proficient and approximately 22 papers were selected to target the borderline between Proficient and Partially Proficient.

Judges were then given the 44 student papers targeted at the preliminary cut scores. Judges classified each of these 44 papers as typical of an Advanced Proficient, Proficient, or Proficient/Partially Proficient student by the definitions. Like Step 5, rating sheets were collected and tabulated with results presented to the judges. Classification frequencies for each paper number were shown to the judges. Judges met in small groups to discuss their classifications. Following the discussions, judges were allowed to make changes to their classifications of the student work on their rating sheets before these were collected.

- Distribution of Rating Sheets and Instructions
- Classification of Papers (Round 2.1)
- Discussion of Judges' Ratings
- Review of Classifications (Round 2.2)

**Step 7 – Review of Impact Data**

Judges received reports summarizing their individual ratings and the group cut scores after Step 6. They were provided the statewide performance data to judge the impact of group standards. Judges were allowed, if they desired, to change the raw score value of their cut score according to this new information.

- Introduction of Individual Judgments and Group Cut Scores
- Introduction of Impact Data
- Final Standard Determinations

**Step 8 – Evaluation of the Standard Setting Process**

Judges were encouraged to rate the process using a five-point scale (five being the highest and one being the lowest). Judges were asked to rate the defining and understanding process of Proficient Performance, Advanced Proficient Performance, and Standard Setting Procedures. Finally, they were asked to rate their confidence in the standard setting results. Additionally, open-ended comments were encouraged.

### 5.3 Results

Judges were provided with graphical data depicting the impact of the resulting cut scores on the actual score distributions of New Jersey eighth-grade students. In other words, if the Proficient cut score is X and the Advanced Proficient cut score is Y, then A percent of the students would be Partially Proficient, B percent of the students would be Proficient, and C percent of the students would be Advanced Proficient. The data were based on more than 88,000 students for each of the content areas.

Judges had an opportunity to review the implications of their standards in the form of impact data. Judges received cumulative frequency distributions of student scores that allowed them to see the percent and number of students in each category given the standards the judges had set.

Table 5.1 presents the cut scores determined by the judges at each round of the standard setting. The numbers in the table indicate the Proficient/Advanced Proficient cut scores in raw score points. The judges' ratings were quite stable from Round 1.1 to the final recommended cut score. Table 5.2 shows the percentage of students achieving at each proficiency level for the total population with the final cut scores.

The final cut score recommendations shown in Table 5.1 were approved and adopted by the New Jersey State Board of Education.

TABLE 5.1

**Proficiency-Level Cut Scores**

<b>Cut Scores Proficient/Advanced Proficient</b>	<b>Language Arts Literacy</b>	<b>Mathematics</b>	<b>Science</b>
<b>Total Possible Points</b>	<b>62</b>	<b>56</b>	<b>52</b>
Round 1.1	28.6/45.2	24.4/43.5	24.2/40.1
Round 1.2	28.6/44.7	24.2/43.1	23.7/39.3
Round 2.1	28.2/44.7	24.3/42.8	23.0/39.0
Round 2.2	28.5/45.0	24.5/42.7	24.3/40.2
<b>Final</b>	<b>29.5/44.5</b>	<b>24.0/43.0</b>	<b>24.0/40.0</b>

TABLE 5.2

**Percentage of Students Achieving Each Performance Level**

	<b>Partially Proficient</b>	<b>Proficient</b>	<b>Advanced Proficient</b>
Language Arts Literacy	24.9%	68.8%	6.3%
Mathematics	40.2%	42.7%	17.0%
Science	26.3%	54.5%	19.2%

## CHAPTER 6: SCALING AND EQUATING

### 6.1 Scaling

The individual student scores are reported as scale scores with a range of 100 to 300. The scores 100 and 300 are a theoretical floor and ceiling and may not actually be observed. The scale score of 250 is the cut score between Proficient students and Advanced Proficient students. The scale score of 200 is the cut score between Proficient students and Partially Proficient students. The score ranges are as follows:

<i>Advanced Proficient</i>	<i>250-300</i>
<i>Proficient</i>	<i>200-249</i>
<i>Partially Proficient</i>	<i>100-199</i>

The Partial Credit Model (PCM) is used for scaling and equating the GEPA operational tests. Masters and Wright (1997) provide this description of the Partial Credit Model:

The Partial Credit Model (PCM) is a unidimensional model for the analysis of responses recorded in two or more ordered categories.... it belongs to the Rasch family of models and so shares the distinguishing characteristics of that family: separable person and item parameters, sufficient statistics, and, hence, conjoint additivity. These features enable "specifically objective" comparisons of persons and items (Rasch, 1977) and allow each set of model parameters to be conditioned out of the estimation procedure for the other.

The PCM (Masters, 1982, 1987, 1988a, 1988b) is the simplest of all item response models for ordered categories. It contains only *two* sets of parameters: one for persons and one for items. All parameters in the model are *locations* on an underlying variable. (p. 101)

WINSTEPS was used to provide the Rasch analyses used for generating the item and student statistics.

Raw score to scale score conversion tables are shown in Appendix D. Appendix E shows Language Arts Literacy, Mathematics, and Science scale score frequency distributions.

### 6.2 Equating

Equating designs must take into account the form of the assessment. Two equating designs are used. Mathematics and Science are equated using a common anchor item, non-equivalent group, design in which all students take common items. These common items are selected to be representative of the total test form in terms of content, difficulty, and format.

The structure of the Language Arts Literacy does not allow for a subset of common exercises to be selected for use across test administrations because the smallest item exercises are unique and singular.

Reading Comprehension is divided into two passage types. These two types cannot be thought of as representative of each other. The Language Arts Literacy equating is accomplished using an embedded equating/field test section that is used for common-item equating.

### Mathematics and Science Equating Design

Common-item equating is used to determine form equivalence from one form, or test administration year, to the next. A set of common (anchor) operational items from the 2006 Mathematics and Science tests was embedded in the 2007 tests. The anchor items include both multiple-choice and open-ended items. Each student participating in the Mathematics and Science testing took the set of common items, and these items contributed to the student's total score. To the maximum extent possible, these items were selected to be proportionally representative of the content and statistics of the total test forms. In addition, the anchor items occupied similar locations in the 2006 and 2007 test forms. These sets of anchor items (14 items with a total of 18 points in Mathematics and 13 items with a total of 15 points in Science) represent approximately one-third of the Mathematics and Science operational tests in terms of number of items and number of points.

The following were applied:

*Calibrate the 2007 test items using the Partial Credit Model and fix the item difficulties to their estimated values based on the 2006 calibration. The common set of items is used.* The item difficulties for the common anchor items on the spring 2007 test were fixed to the estimated item difficulties from the calibration of the 2006 operational test. This placed all parameter estimates for the 2007 calibration on the 2006 scale. This also produced the new raw score to ability (theta) table for the 2007 test.

*Develop a raw score to scale score table for the 2007 assessments.* Using the ability to scale score relationship found in the 2006 test calibrations, scale scores were assigned to the raw scores from the 2007 assessments. This was possible because each ability in the ability to scale score table corresponds to a single raw score; therefore, the scale score assigned to that ability can also be assigned to the raw score.

Checks during the equating process were necessary to establish the stability of the common items and determine model fit. One such check was accomplished through the use of the common anchor items from the 2006 operational test embedded in the 2007 operational test. The following is a summary of the steps used for the anchor item analysis.

1. Identify anchor item difficulties from the item bank,
2. Calibrate 2007 form without fixing anchor item difficulties with WINSTEPS,
3. Calculate mean of the bank anchor items difficulties,
4. Calculate mean of 2007 anchor items,
5. Add constant to 2007 anchor item difficulties so the mean equals that found in the bank values,
6. Subtract 2007 and the bank anchor difficulties after adding the constant,
7. Drop item with largest absolute difference greater than or equal to 0.30 for consideration as anchor item, and
8. Repeat steps 1-7 using remaining anchor items.

The final product from the equating procedure was the raw score to scale score table developed in Step 2. When equating was completed, raw score to scale score conversion tables were available for scoring. These two steps can be applied for future assessments.

### **Language Arts Literacy Equating Design**

Scaling and equating for Language Arts Literacy was accomplished through a different design. Each assessment has an embedded equating/field test section that is used for either common-item equating or new-item field testing. Language Arts Literacy was equated using a design in which operational items appeared in a section designated for equating or field testing.

The test included the operational items and four equating sections. Students across the state took one of the equating sections or a field test section. Sampling was done by school and stratified by District Factor Grouping to approximate equivalent groups between equating sets. Sample sizes for each equating/field test form were more than 17,000 students or approximately 17 percent of the student examinee population.

The Language Arts Literacy was equated using a common item design with a combined run. Two forms of the 2007 assessment contained two of the operational passages from 2006 in the field test section. This design allowed for the development of a matrix design in the data, with a combination of data records from 2006 and 2007. All data was analyzed in a combined run with the 2006 item parameters fixed to their 2006 values. This places the 2007 item parameters onto the 2006 scale. Using those 2007 item parameters, a raw score to theta relationship was calculated. This was then used to develop the raw score to scale score table.

### **Summary of Equating Statistics**

Table 6.1 shows a summary of the statistics used to evaluate the psychometric quality of the assessments. All three assessments had a high degree of reliability ranging from 0.88 to 0.90. The standard errors in terms of raw scores ranged from 2.46 to 3.43.

Examination of the fit statistics shows that the Partial Credit Model fits the data reasonably well. The INFIT statistic is a measure of the model fit weighted by the placement of the person locations. The OUTFIT statistic does not apply this weighting and is more sensitive to misfit. It is generally accepted that items with statistics between 0.7 and 1.3 have good fit. On average all assessments demonstrated fit within these limits. On an individual bases, all items had INFIT statistics within this range, but some of the items had OUTFIT statistics which fell outside this range. The number of items with OUTFIT statistics falling outside the range of 0.7 and 1.3 is consisted with past analysis.

Table 6.2 lists the cut scores resulting from the current equating results. Also, those derived from 2002, 2003, 2004, 2005, and 2006 are provided for comparison.

**TABLE 6.1**  
**Summary of Equating Statistics**

	<b>Language Arts Literacy</b>	<b>Mathematics</b>	<b>Science</b>
Number of items	26	36	48
Raw Score Range	0 to 54	0 to 48	0 to 54
Coefficient Alpha	.88	.90	.88
Count of negative biserials	None	None	None
Raw Score (Population)			
Mean	32.2	28.0	30.5
SD	7.1	9.8	9.9
SEM	2.46	3.10	3.43
Rasch Person Measures*			
Mean	0.95	0.33	0.57
SD	1.35	1.20	0.94
SEM	0.47	0.38	0.33
Item Infit MNSQ			
Mean	0.99	1.00	1.00
SD	0.12	0.08	0.07
# Between 0.7 and 1.3	26 of 26	36 of 36	48 of 48
Item Outfit MNSQ			
Mean	1.04	1.03	1.01
SD	0.23	0.15	0.12
Between 0.7 and 1.3	22 of 26	35 of 36	47 of 48

**TABLE 6.2**  
**Cut Scores and Associated Thetas for Proficiency Levels**

	<b>Raw Score Cuts</b>		<b>Rasch Theta Score Cuts</b>	
	<b>Proficient</b>	<b>Advanced</b>	<b>Proficient</b>	<b>Advanced</b>
<b>Language Arts Literacy</b>				
2002	26.5	44.0	0.253	2.780
2003	29.5	45.0	0.238	2.715
2004	31.0	46.5	0.244	2.773
2005	29.0	41.0	0.253	2.664
2006	29.0	41.0	0.223	2.715
2007	29.5	40.0	0.276	2.656
<b>Mathematics</b>				
2002	24.0	39.0	-0.074	1.297
2003	24.0	38.5	-0.061	1.323
2004	24.0	38.0	-0.079	1.278
2005	25.0	39.0	-0.062	1.333
2006	25.0	38.5	-0.079	1.278
2007	24.0	37.0	-0.134	1.244
<b>Science</b>				
2002	22.0	39.5	-0.132	1.344
2003	22.0	39.5	-0.157	1.319
2004	21.0	38.0	-0.132	1.352
2005	20.5	37.5	-0.174	1.344
2006	22.0	39.0	-0.169	1.340
2007	22.0	39.0	-0.179	1.262

## CHAPTER 7: TEST STATISTICS

---

### 7.1 Reliability of the Test Scores

Table 7.1 summarizes reliability estimates for the content areas and clusters. The reliability coefficients given in this table are based on Cronbach's coefficient alpha measure of internal consistency. Cronbach's alpha is used on tests containing items that can be scored along a range of values. The standard errors of measurement (SEMs) for the major content areas are expressed in terms of the raw score metric and the scale score metric. The scale scores range from 100 to 300.

Reliabilities and SEMs for the dichotomously scored items in each cluster are reported using the Kuder-Richardson Formula 20 (KR-20) in Table 7.2.

When evaluating these results, it is important to recall that reliability is partially a function of test length. Therefore, the reliability of a content area is likely to be greater than the reliability of a cluster simply because the content area has more items. Similarly, clusters with more items are likely to be more reliable than clusters with fewer items. The data provided in Tables 7.1 and 7.2 reflect the expected positive relationship between test length and reliability.

The SEMs are useful when interpreting students' scores. Measurement error occurs in every test. A student's true score is a hypothetical average score that the student would obtain if a test were repeatedly administered to the student without the effects of instruction, practice, or fatigue. Mehrens and Lehmann (1991) suggest this use of the SEM:

---

The standard error measurement is often used for what is called band interpretation. Band interpretation helps convey the idea of imprecision of measurement...If we assume that the errors are random, an individual's observed scores will be normally distributed about his true score over repeated testing. Thus, one can say that a person's observed scores will lie between  $\pm 1Se$  of his true score approximately 68 percent of the time, or  $\pm 2Se$  of his true score about 95 percent of the time. Of course, we do not know the true score, but one can infer with about 68% (or 95%) certainty that a person's true score is within  $\pm 1Se$  (or  $\pm 2Se$ ) of his observed score. (p. 252)

---

TABLE 7.1

**Reliability Estimates and Standard Errors of Measurement (SEMs)  
for Content Areas and Clusters - 2007**

<b>GEPA Test Section</b>	<b>Number of Points</b>	<b>Reliability Cronbach's Alpha</b>	<b>SEM Raw Score</b>	<b>SEM Scale Score</b>
<b>Language Arts Literacy</b>	<b>54</b>	<b>0.87</b>	<b>2.55</b>	<b>13.17</b>
Reading	36	0.85	2.06	
Writing	18	0.68	1.18	
Interpreting Text	15	0.70	2.58	
Analyzing/Critiquing Text	21	0.76	1.51	
<b>Mathematics</b>	<b>48</b>	<b>0.89</b>	<b>3.18</b>	<b>13.23</b>
Number and Numerical Operations	12	0.68	1.60	
Geometry and Measurement	12	0.65	1.58	
Patterns and Algebra	12	0.68	1.63	
Data Analysis, Probability, and Discrete Mathematics	12	0.66	1.56	
Knowledge	48	0.89	3.18	
Problem Solving	35	0.86	2.78	
<b>Science</b>	<b>54</b>	<b>0.88</b>	<b>3.38</b>	<b>10.69</b>
Life	22	0.75	2.17	
Physical	16	0.68	1.86	
Earth	16	0.71	1.81	
Knowledge	11	0.58	1.49	
Application	43	0.86	3.03	

TABLE 7.2

**Reliability Estimates and Standard Errors of Measurement (SEMs)  
for Dichotomously Scored Items Within Content Clusters - 2007**

<b>GEPA Content Area</b>	<b>Number of Items</b>	<b>Reliability (KR-20)</b>	<b>SEM Raw Score</b>
<b>Language Arts Literacy</b>	<b>20</b>	<b>0.79</b>	<b>1.65</b>
Reading	20	0.79	1.65
Writing	–	–	–
Writing/Speculate	–	–	–
Writing/Persuade	–	–	–
Interpreting Text	11	0.67	1.26
Analyzing/Critiquing Text	9	0.64	1.06
<b>Mathematics</b>	<b>30</b>	<b>0.84</b>	<b>2.33</b>
Number and Numerical Operations	9	0.62	1.29
Geometry and Measurement	6	0.54	1.05
Patterns and Algebra	6	0.54	1.08
Data Analysis, Probability, and Discrete Mathematics	9	0.60	1.21
Knowledge	30	0.84	2.33
Problem Solving	17	0.75	1.77
<b>Science</b>	<b>45</b>	<b>0.86</b>	<b>2.99</b>
Life	19	0.72	1.95
Physical	13	0.64	1.61
Earth	13	0.67	1.59
Knowledge	11	0.58	1.49
Application	34	0.83	2.59

\* There were no dichotomously scored writing items.

## CHAPTER 8: ITEM-LEVEL STATISTICS

The GEPA test specifications are aligned with the Core Curriculum Content Standards. Please refer to the *Technical Manual* and Part 2 of this *Technical Report* for information about the test specifications and test development.

### 8.1 Classical Item Statistics

In Table 8.1, summary statistics are given that describe the difficulty and discrimination of the items comprising each cluster. For dichotomously scored items, means and standard deviations of proportion-correct values (p-values) and point-biserials are given. For the open-ended items, the index of item difficulty is calculated by dividing students' average score on an item by the maximum possible score on the item. Item discrimination for each open-ended item is the correlation between students' item score and their total score on the test section. For both the item-test correlation and the point-biserial correlation, students' total test scores are expressed in terms of the raw score metric.

**TABLE 8.1**

**Item Difficulty and Discrimination Summary Statistics for Dichotomously Scored and Open-Ended Items by Test Section and Cluster - 2007**

GEPA Test Section/Cluster	Dichotomous			Open-Ended		
	Item Difficulty		Item Discrimination	Item Difficulty		Item Discrimination
	Mean	S.D.	Mean	Mean	S.D.	Mean
<b>Language Arts Literacy</b>	<b>0.77</b>	<b>0.12</b>	<b>0.44</b>	<b>0.50</b>	<b>0.12</b>	<b>0.92</b>
Reading	0.77	0.12	0.44	0.44	0.14	0.85
Writing	–	–	–	0.55	0.12	0.85
Picture	–	–	–	0.57	0.13	0.75
Persuasive	–	–	–	0.54	0.13	0.79
Interpret Text	0.76	0.13	0.43	0.40	0.17	0.68
Analyze/Critique Text	0.80	0.11	0.44	0.45	0.14	0.83
<b>Mathematics</b>	<b>0.64</b>	<b>0.13</b>	<b>0.41</b>	<b>0.49</b>	<b>0.24</b>	<b>0.93</b>
Number and Numerical Operations	0.64	0.11	0.41	0.66	0.35	0.72
Geometry and Measurement	0.58	0.17	0.43	0.38	0.25	0.74
Patterns and Algebra	0.58	0.11	0.42	0.46	0.28	0.80
Data Analysis	0.72	0.10	0.39	0.62	0.36	0.70
Knowledge	0.64	0.13	0.41	0.49	0.24	0.93
Problem Solving	0.63	0.13	0.41	0.49	0.24	0.93
<b>Science</b>	<b>0.60</b>	<b>0.12</b>	<b>0.37</b>	<b>0.42</b>	<b>0.26</b>	<b>0.82</b>
Life	0.59	0.13	0.37	0.48	0.35	0.64
Physical	0.62	0.10	0.36	0.33	0.34	0.65
Earth	0.59	0.13	0.39	0.44	0.32	0.62
Knowledge	0.58	0.13	0.36	–	–	–
Application	0.60	0.12	0.38	0.42	0.26	0.82

Tables 8.2, 8.3, and 8.4 present frequency distributions of item difficulty (p-values) and item discrimination indices by content cluster. The top section of each table shows the distribution of item difficulty values; the bottom section shows the distribution of point-biserial correlations.

Point-biserial indices are produced to evaluate operational test items. Millman and Greene (1989) note that the point-biserial index gives a true reflection of the item's contribution to the functioning of the test. For field test item review (described in Test Development) biserial correlations are computed. The biserial indices tend to be more stable across samples.

TABLE 8.2

**Frequency Distributions of Item Difficulty  
and Item Discrimination by Content Cluster**

**2007 Language Arts Literacy**

<b>Item Statistics</b>	<b>Interpreting Text</b>	<b>Analyzing Text</b>	<b>Total</b>
<b>ITEM DIFFICULTY: P-VALUES</b>			
.90+	2	1	3
.80 – .89	2	4	6
.70 – .79	3	2	5
.60 – .69	2	1	3
.50 – .59	2	1	3
<.40 – .49	0	0	0
<b>MEAN P-VALUE</b>	<b>.76</b>	<b>.80</b>	<b>.77</b>
<b>MEDIAN P-VALUE</b>	<b>.79</b>	<b>.83</b>	<b>.79</b>
<b>ITEM DISCRIMINATION: POINT-BISERIAL CORRELATIONS</b>			
.50+	1	1	2
.40 – .49	6	7	13
.30 – .39	4	1	5
<.30	0	0	0
<b>MEAN POINT-BISERIAL</b>	<b>.43</b>	<b>.44</b>	<b>.44</b>
<b>MEDIAN POINT-BISERIAL</b>	<b>.43</b>	<b>.43</b>	<b>.43</b>
<b>TOTAL NUMBER OF ITEMS</b>	<b>11</b>	<b>9</b>	<b>20</b>

TABLE 8.3

**Frequency Distributions of Item Difficulty and Item Discrimination by Content Cluster**  
**2007 Mathematics**

<b>Item Statistics</b>	<b>Number and Numerical Operations</b>	<b>Geometry and Measurement</b>	<b>Patterns and Algebra</b>	<b>Data Analysis, Probability, and Discrete Mathematics</b>	<b>Knowledge</b>	<b>Problem Solving</b>	<b>Test Total</b>
<b>ITEM DIFFICULTY: P-VALUES</b>							
.90+	0	0	0	0	0	0	0
.80 – .89	0	1	0	2	3	2	3
.70 – .79	3	1	1	4	9	4	9
.60 – .69	2	0	2	2	6	3	6
.50 – .59	3	1	0	1	5	5	5
.40 – .49	1	3	3	0	7	3	7
.30 – .39	0	0	0	0	0	0	0
<.30	0	0	0	0	0	0	0
<b>MEAN P-VALUE</b>	<b>.64</b>	<b>.58</b>	<b>.58</b>	<b>.72</b>	<b>.64</b>	<b>.63</b>	<b>.64</b>
<b>MEDIAN P-VALUE</b>	<b>.65</b>	<b>.52</b>	<b>.57</b>	<b>.73</b>	<b>.65</b>	<b>.64</b>	<b>.65</b>
<b>ITEM DISCRIMINATION: POINT-BISERIAL CORRELATIONS</b>							
.50+	1	1	1	1	4	4	4
.40 – .49	4	3	2	3	12	4	12
.30 – .39	4	2	3	3	12	7	12
<.30	0	0	0	2	2	2	2
<b>MEAN POINT-BISERIAL</b>	<b>.41</b>	<b>.43</b>	<b>.42</b>	<b>.39</b>	<b>.41</b>	<b>.41</b>	<b>.41</b>
<b>MEDIAN POINT-BISERIAL</b>	<b>.40</b>	<b>.43</b>	<b>.41</b>	<b>.34</b>	<b>.41</b>	<b>.39</b>	<b>.41</b>
<b>TOTAL NUMBER OF ITEMS</b>	<b>9</b>	<b>6</b>	<b>6</b>	<b>9</b>	<b>30</b>	<b>17</b>	<b>30</b>

TABLE 8.4

**Frequency Distributions of Item Difficulty  
and Item Discrimination by Content Cluster**

**2007 Science**

<b>Item Statistics</b>	<b>Life</b>	<b>Physical</b>	<b>Earth</b>	<b>Knowledge</b>	<b>Application</b>	<b>Total Test</b>
<b>ITEM DIFFICULTY: P-VALUES</b>						
.80 +	2	0	0	1	1	2
.70 – .79	1	1	3	0	5	5
.60 – .69	5	9	3	5	12	17
.50 – .59	6	1	4	2	9	11
.40 – .49	3	1	2	2	4	6
<.40	2	1	1	1	3	4
<b>MEAN P-VALUE</b>	<b>.59</b>	<b>.62</b>	<b>.59</b>	<b>.58</b>	<b>.60</b>	<b>.60</b>
<b>MEDIAN P-VALUE</b>	<b>.58</b>	<b>.65</b>	<b>.59</b>	<b>.60</b>	<b>.61</b>	<b>.60</b>
<b>ITEM DISCRIMINATION: POINT-BISERIAL CORRELATIONS</b>						
.50 +	1	0	0	1	0	1
.40 – .49	7	7	6	4	16	20
.30 – .39	6	3	7	2	14	16
.20 – .29	5	2	0	4	3	7
<.20	0	1	0	0	1	1
<b>MEAN POINT-BISERIAL</b>	<b>.37</b>	<b>.36</b>	<b>.39</b>	<b>.36</b>	<b>.38</b>	<b>.37</b>
<b>MEDIAN POINT-BISERIAL</b>	<b>.37</b>	<b>.40</b>	<b>.39</b>	<b>.36</b>	<b>.39</b>	<b>.39</b>
<b>TOTAL NUMBER OF ITEMS</b>	<b>19</b>	<b>13</b>	<b>13</b>	<b>11</b>	<b>34</b>	<b>45</b>

## 8.2 Speededness

The amount of time allotted for students to complete the test is intended to provide nearly all students with sufficient time to answer all the questions. Table 8.5 presents data concerning the extent to which this intent was met. Open-ended items appear at the end of each part. For this reason, Table 8.5 shows the percentage of students omitting each of the last three multiple-choice items in each part and all open-ended items.

The percent of students omitting the Reading multiple-choice items is very small at about 0.1% to 0.3%. The percent of students omitting the open-ended items varies from 1.3% to 2.8%.

TABLE 8.5

**Percentage of Students Omitting the  
Last Items of Each Test Part - 2007**

Test Section	Multiple-Choice		Open-Ended	
	Item Number	Percentage Omitting	Item Number	Percentage Omitting
<b>Reading</b>				
Part A	Item 8	0.3%	Item 11	1.3%
	Item 9	0.2%	Item 12	2.8%
	Item 10	0.3%		
Part A	Item 8	0.1%	Item 11	1.3%
	Item 9	0.1%	Item 12	2.1%
	Item 10	0.2%		
<b>Mathematics</b>				
Part A	Item 8	0.2%	Item 11	1.5%
	Item 9	0.8%	Item 12	4.3%
	Item 10	0.3%		
Part B	Item 8	0.4%	Item 11	1.5%
	Item 9	0.2%	Item 12	3.8%
	Item 10	0.3%		
Part C	Item 8	0.2%	Item 11	2.7%
	Item 9	0.5%	Item 12	3.7%
	Item 10	0.4%		
<b>Science</b>				
Part A	Item 13	0.5%	Item 16	3.2%
	Item 14	0.8%		
	Item 15	1.0%		
Part B	Item 13	0.2%	Item 16	3.1%
	Item 14	0.5%		
	Item 15	0.4%		
Part C	Item 13	0.3%	Item 16	5.7%
	Item 14	0.4%		
	Item 15	0.6%		

The percent of students omitting the Mathematics multiple-choice items ranges from 0.2% to 0.8%. The percent of students omitting the Mathematics open-ended items varies from 1.5% to 4.3%.

The percent of students omitting the Science multiple-choice items ranges from 0.2% to 1.0%. The percent of students omitting the Science open-ended items varies from 3.1% to 5.7%.

Overall, these data indicate that the amount of time provided for completing the test is appropriate and that speed of response is not a factor that affects students' performances or detracts from the validity of scores.

### 8.3 Intercorrelations

The Pearson product-moment correlation between student scores on Language Arts Literacy and Mathematics was .73, Language Arts Literacy and Science was .71, and Mathematics and Science was .80. Table 8.6 shows the correlations between students' scores in the major content clusters and item types. Table 8.7 shows the correlations between student scores on the content clusters. The scores used for all correlations were expressed in the raw score metric.

Note that correlations between a content area and cluster within that content area are partially a function of the proportion of the content area that is made up of items from the given cluster. Clusters with many items that make up a large proportion of the content area score increase the cluster with content area correlation.

For example, the correlation between Reading and Language Arts Literacy in Table 8.6 is quite high (.98) because 36 Reading points are part of the total Language Arts Literacy 54 points.

In addition, correlations are partially a function of the number of items in the measures being correlated. Therefore, the number of items in the content areas and clusters being correlated must be considered when their correlations are evaluated. In Table 8.7, the L3 Writing/Speculate cluster has only six points, so this cluster may not correlate as highly with other clusters due to this small number of points.

**TABLE 8.6**

#### **Intercorrelations Among Major Content Clusters and Item Types - 2007**

Major Content and Item Types	Major Content Clusters and Item Types										
	Language Arts Literacy					Mathematics			Science		
	LAT	R	R MC	R OE	W	M T	M MC	M OE	ST	S MC	S OE
<b>LAT Language Arts Literacy (54)</b>											
R Reading (36)	.98										
R MC Reading Multiple-Choice (20)	.91	.95									
R OE Reading Open-ended (16)	.85	.85	.63								
W Writing (18)	.85	.72	.62	.69							
<b>MT Mathematics (48)</b>	<b>.73</b>	<b>.72</b>	<b>.67</b>	<b>.61</b>	<b>.63</b>						
M MC Mathematics Multiple-Choice (30)	.68	.66	.63	.55	.57	.96					
M OE Mathematics Open-ended (18)	.72	.70	.65	.61	.62	.93	.79				
<b>ST Science (54)</b>	<b>.71</b>	<b>.71</b>	<b>.68</b>	<b>.58</b>	<b>.58</b>	<b>.80</b>	<b>.76</b>	<b>.76</b>			
S MC Science Multiple-Choice (45)	.69	.69	.67	.55	.55	.78	.75	.73	.99		
S OE Science Open-ended (9)	.64	.63	.58	.55	.54	.68	.62	.67	.82	.72	

Number in parentheses is the number of score points.

Language Arts Literacy N = 105,824; Mathematics N = 106,899; Science N = 106,832.

**TABLE 8.7**  
**Intercorrelations Among Content Areas and Clusters - 2007**

Test Section/Cluster	Test Section/Cluster																				
	Language Arts Literacy						Mathematics						Science								
	LAT	L1	L2	L3	L4	L5	L6	MT	M1	M2	M3	M4	M5	M6	ST	S1	S2	S3	S4	S5	
<b>LAT Language Arts Literacy (54)</b>																					
L1 Reading (36)	.98																				
L2 Writing (18)	.85	.72																			
L3 Writing/Speculate (6)	.75	.66	.83																		
L4 Writing/Persuade (12)	.79	.65	.96	.64																	
L5 Interpreting Text (15)	.90	.93	.64	.59	.58																
L6 Analyzing/Critiquing Text (21)	.93	.95	.70	.64	.64	.76															
<b>MT Mathematics (48)</b>	<b>.73</b>	<b>.72</b>	<b>.63</b>	<b>.57</b>	<b>.57</b>	<b>.68</b>	<b>.67</b>														
M1 Number and Numerical Operations (12)	.64	.63	.55	.50	.51	.59	.58	.88													
M2 Geometry and Measurement (12)	.61	.59	.53	.49	.49	.56	.55	.87	.68												
M3 Patterns and Algebra (12)	.65	.63	.55	.51	.50	.60	.59	.89	.71	.70											
M4 Data Analysis, Probability, and Discrete Mathematics (12)	.66	.65	.55	.50	.50	.62	.60	.86	.69	.66	.68										
M5 Knowledge (48)	.73	.72	.63	.57	.57	.68	.67	1.00	.88	.87	.89	.86									
M6 Problem Solving (35)	.73	.71	.62	.57	.57	.67	.66	.98	.87	.86	.87	.85	.98								
<b>ST Science (54)</b>	<b>.71</b>	<b>.71</b>	<b>.58</b>	<b>.55</b>	<b>.52</b>	<b>.68</b>	<b>.65</b>	<b>.80</b>	<b>.69</b>	<b>.70</b>	<b>.71</b>	<b>.69</b>	<b>.80</b>	<b>.79</b>							
S1 Life (22)	.67	.67	.54	.51	.49	.64	.61	.74	.63	.64	.66	.64	.74	.73	.92						
S2 Physical (16)	.63	.63	.52	.49	.47	.60	.58	.72	.61	.63	.64	.62	.72	.71	.89	.73					
S3 Earth (16)	.62	.62	.51	.48	.46	.59	.57	.71	.61	.62	.63	.61	.71	.69	.89	.73	.70				
S4 Knowledge (11)	.55	.55	.44	.42	.40	.53	.50	.63	.54	.56	.56	.54	.63	.62	.82	.74	.75	.73			
S5 Process Skills (43)	.71	.71	.58	.55	.52	.68	.65	.79	.68	.70	.71	.69	.79	.78	.99	.92	.87	.88	.71		

Number in parentheses is the number of score points.  
Language Arts Literacy N = 105,824; Mathematics N = 106,899; Science N = 106,832.

## CHAPTER 9: TEST VALIDITY

---

The validity chapter in the *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 1999, p. 9) begins:

---

Validity refers to the degree to which evidence and theory support the interpretations of test scores entailed by proposed uses of tests. Validity is, therefore, the most fundamental consideration in developing and evaluating tests. The process of validation involves accumulating evidence to provide a sound scientific basis for the proposed score interpretations. It is the proposed uses that are evaluated, not the test itself. When test scores are used or interpreted in more than one way, each intended interpretation must be validated.

Validity logically begins with an explicit statement of the proposed interpretation of test scores, along with a rationale for the relevance of the interpretation to the proposed use.

---

The purposes served by the GEPA scores are noted in the following paragraph from page 7 of the manual, *Score Interpretation Manual*:

---

The GEPA is intended to evaluate the progress students are making in mastering the knowledge and skills required by the end of the eighth grade and in mastering the knowledge and skills they will need to pass the HSPA. The GEPA should serve as a primary indicator for identifying those students who may need instructional intervention. The test should also serve as an indicator for determining which local education programs may need revisions to ensure that instructional programs are aligned with the Core Curriculum Content Standards.

---

What represents a sufficient collection of evidence in the demonstration of test validity has been the subject of considerable research, thought, and debate in the measurement community over the years. Several different conceptions of validity and approaches to test validation have been proposed, and as a result the field has evolved. In 1995, Messick clarified:

---

The validity issues of score meaning, relevance, utility, and social consequences are many-faceted and intertwined. They are difficult if not impossible to disentangle, which is why validity has come to be viewed as a unified concept (APA, AERA, & NCME, 1985; Messick, 1989). However, to speak of validity as a unified concept does not imply that validity cannot be usefully differentiated into distinct aspects to underscore issues and nuances that might otherwise be downplayed or overlooked, such as the social consequences of performance assessments or the role of score meaning in applied use. The intent of these distinctions is to provide a means of addressing functional aspects of validity that help disentangle some of the complexities inherent in appraising the appropriateness, meaningfulness, and usefulness of score inferences.

In particular, six distinguishable validity aspects are delineated emphasizing content, substantive, structural, generalizability, external, and consequential aspects of construct validity (Messick, 1994, in press). (pp. 5 and 6)

---

The fifth edition of the *Standards* (1999) recommends establishing the validity of a test through use of a validity argument. The *Standards* (1999) defines a validity argument as “An explicit scientific justification of the degree to which accumulated evidence and theory support the proposed interpretation(s) of test scores.”

The *Standards* (1999) recognized the following possible sources of validity evidence:

- Evidence based on test content
- Evidence based on response processes
- Evidence based on internal structure
- Evidence based on relations to other variables
- Evidence based on consequences of testing

The present chapter of this report concerning sources of GEPA validity evidence is organized in sections according to the following traditional validity terms: content and curricular validity, construct validity, criterion-related validity, and consequential validity evidence.

For each of the GEPA content areas, New Jersey educators defined the content and skill test specifications. Content area committees assisted with developing the Directory of Test Specifications and Sample Items which delineate specifications used to create the assessments and to measure student proficiency in the knowledge and skills outlined in the Core Curriculum Content Standards.

Test specifications for the GEPA content areas were designed to align with the Core Curriculum Content Standards. The GEPA Content Committees recommended the emphases and priorities reflected in the number of items for each item type and cluster on the test. The 2007 test specifications are based on the standards adopted in 2002 and 2004.

Curriculum developers and teachers use the specifications, along with curriculum frameworks, the standards themselves, and the score reports, to improve instruction at the district, school, and classroom levels. A number of reports have been designed to assist educators with focusing on pertinent information. Report forms designed to meet specific needs extend the effectiveness of a testing program by making it easier to use test results for educational planning. Chapter 10 of this *Technical Report* includes descriptions and examples of the reports.

Beginning with the 1991 EWT due notice testing, the students' essays also have been returned to the districts for distribution to appropriate district staff members for analysis and use in classroom instruction. A manual, *Cycle II Criterion-Based Holistic Scoring: A Writing Handbook* included with the essays, presents the scoring method and criteria used to evaluate student writing and offers suggestions for using the New Jersey's scoring rubrics and student test data to improve classroom instruction. Teachers are encouraged to review the sample responses in the handbook, the annotations on each of the sample responses, and the features of the respective score scales.

The State Department of Education releases a State Summary Report for each content area tested, which contains summary results at the state, district, and school levels as well as statewide results by District Factor Groups (DFG) and special needs districts. Districts are required to report test results to their boards of education and to the public within 30 days after receiving test reports. Analysis and interpretation of the school and district reports is required by the New Jersey Administrative Code (N.J.A.C. 6A:8-4.3(a), (b)).

Further information about the legal and historical background for the GEPA is available at:

<http://www.state.nj.us/education/code/current/title6a/chap8.pdf>

<http://www.state.nj.us/education/assessment/history.shtml>

### 9.1 Content and Curricular Validity (Evidence Based on Test Content)

Content validity is the most relevant and important source of evidence for the GEPA. The validity of the GEPA scores is based on the alignment of the GEPA to the Core Curriculum Content Standards and the knowledge and skills expected of eighth-grade students.

The Core Curriculum Content Standards were developed by teachers and other educational professionals from New Jersey. The Core Curriculum Content Standards outline what students should know and be able to do at a certain grade level. The questions on the GEPA can contain items/concepts included in the grade eight standards as well as for those standards listed for the prior grades.

The content area committees assisted with developing the Directory of Test Specifications and Sample Items for each of the assessed areas. Attributes of New Jersey educators serving on the committees include:

- strong knowledge of the content area,
- familiarity with New Jersey’s Core Curriculum Content Standards for the specific content area,
- understanding of student’s skills and abilities at the eighth-grade benchmark level,
- some understanding of assessment procedures,
- the ability to work effectively in teams,
- a commitment to educational excellence, and
- sensitivity to students’ needs.

The three content area directories are available online at:

<http://www.state.nj.us/education/njpep/assessment/TestSpecs/LangArts/TOC.html>

<http://www.state.nj.us/education/njpep/assessment/TestSpecs/MathTestSpec/GEPAMath/MathIndex.html>

<http://www.state.nj.us/education/njpep/assessment/TestSpecs/ScienceGEPA/index.html>

Sequential procedures of test specification development through operational test approval described in Chapter 2 of this report ensure the content validity of the tests. The item development teams at Measurement Incorporated begin each item development cycle with a review of the Core Curriculum Content Standards and the three directories of test specifications. Using their years of experience with New Jersey item writing and reviews, item writers understand how to develop multiple-choice and open-ended items that tap the appropriate range of skills. They understand the cognitive complexity required within their content area. Items are designed to assess higher-order or critical thinking skills in varied contexts that are familiar to students. Item content for all items, including the writing-task prompts, is carefully reviewed to ensure that items are free from gender, racial, ethnic, and regional bias.

Prior to field testing, all test items are reviewed by the New Jersey Assessment Content and Sensitivity Review Committees as well as the Office of State Assessments staff to ensure that items meet GEPA test specifications including appropriate difficulty and skill requirements. Item approval forms used by the Content Review Committees include two categories that address the cognitive complexity of items:

- match to the test specifications
- appropriate difficulty

The Sensitivity Review Committee reviews to ensure that test questions are not offensive and do not reinforce negative stereotypes, and that test questions appropriately reflect multicultural society. Item approval forms used by the Sensitivity Review Committee require each item to be identified as “Definitely Use” or “Revise and Use With Approval” before the item can be included on a field test.

### **9.2 Construct Validity (Evidence Based on Response and Evidence Based on Internal Structure)**

The glossary of *Standards for Educational and Psychological Testing* (1999) presents this definition of construct validity:

---

A term used to indicate that the test scores are to be interpreted as indicating the test taker’s standing on the psychological construct measured by the test. A construct is a theoretical variable inferred from multiple types of evidence, which might include the interrelations of the test scores with other variables, internal test structure, observations of response processes, as well as the content of the test. In the current standards, all test scores are viewed as measures of some construct, so the phrase is redundant with validity. The validity argument establishes the construct validity of a test. (p. 174)

---

Item statistics and intercorrelations provide validity evidence related to internal structure. A large percentage of the GEPA score points for each content area come from open-ended and essay test questions. Beginning with the rangefinding process and continuing through statistical review, many of the responses to these questions are scored, reviewed, and discussed by the Content Review Committees members, the NJDOE Content Coordinators, and the Measurement Incorporated staff. These processes have been repeated annually since 1993. Information obtained from students’ responses to these questions provides insight used for test item acceptance, modification, and rejection as well as for future test item development.

Open-ended questions and essays compose about 63% (34/54) of the Language Arts Literacy points, 38% (18/48) of the Mathematics points, and 17% (9/54) of the Science points. Many open-ended items are field tested each year. During 2007, 4 Reading open-ended items, 5 writing prompts, 4 Mathematics open-ended items, and 5 Science open-ended items were field tested. For each open-ended item, the Measurement Incorporated Project Director prepared a brief summary discussing the types of responses with notes about any issues and concerns. This summary was included with a copy of each item, rubric, sample answer, and rangefinding papers for reference during the statistical review.

For all field test items, Pearson computed item means, response frequencies, biserial correlations (the field test item with the base test total score), and other descriptive statistics. Content Review Committee members used these statistics, their classroom experiences, and the open-ended responses to discuss and explain the processes they believed students were using to provide the correct and incorrect responses to items. Committee members reviewed for concerns related to ambiguity, irrelevant clues, and inaccuracy. Each item must be classified as “Definitely Use” or “Revise and Use with Approval” before it could appear on an operational test.

In addition, several statistics including item difficulty, item discrimination, and item omits are produced for the operational test and printed in each *Technical Report*. Other operational statistics calculated include Pearson product-moment correlations between students’ scores on the operational test content clusters and item types.

### 9.3 Criterion-Related Validity (Evidence Based on Relations to Other Variables)

The *Standards for Educational and Psychological Testing* (1974) presents this definition of criterion validity:

---

Criterion-related validities apply when one wishes to infer from a test score an individual's most probable standing on some other variable called a criterion. Statements of predictive validity indicate the extent to which an individual's future level on the criterion can be predicted from a knowledge of prior test performance; statements of concurrent validity indicate the extent to which the test may be used to estimate an individual's present standing on the criterion. The distinction is important. (p. 26)

---

Sources of evidence related to concurrent and predictive validity for GEPA score interpretations are linked to the purposes that score report information serves for districts, schools, and teachers. The *Score Interpretation Manual* provides procedures for disseminating score reports and using test score information. A section using reports for student-level evaluation notes:

---

Further examination of a student's knowledge and skill deficiencies should include the analysis of the student's whole profile. Decisions about appropriate instructional programs should be based on examination of a student's classroom test results, grades, anecdotal records, portfolios, checklists, school-level results, and other measures of performance. (p. 38)

---

An important purpose of the GEPA is its predictive relationship to the High School Proficiency Assessment (HSPA). A study by Zhao, Robinson, and Guo (2007) provides evidence of the predictive relationship between GEPA scores and HSPA scores. The study considered two cohort samples:

- Cohort 1 (n=37,161) includes students who took the GEPA as eighth graders in 2000 and took the HSPA as eleventh graders in 2003.
- Cohort 2 (n=38,653) includes students who took the GEPA in 2001 and the HSPA in 2004.

Because the GEPA and HSPA programs have no common student identifier, GEPA students' names, gender, and date of birth within school districts were used to match to HSPA students' names, gender, and date of birth to identify students' records to use for the study. The authors noted they matched students within school district to reduce mobility impact and data merge concerns. They did not include Limited English Proficient (LEP) and Special Education (SE) students in the study because these students typically show greater score variation across years.

Zhao, Robinson, and Guo found the correlation coefficient 0.72 for the GEPA and HSPA Language Arts Literacy total scores for both Cohort 1 and Cohort 2. The 0.72 correlation coefficient indicates that the GEPA Language Arts Literacy total score explains 51.8% of the variance in the HSPA Language Arts Literacy total score. Similarly, the correlation coefficient 0.85 was determined for the GEPA and HSPA Mathematics total scores for Cohort 1 and Cohort 2 which indicates 72.3% of the variance in the HSPA mathematics total score is explained by the GEPA mathematics score.

Zhao, Robinson, and Guo calculated the number and percentage of students in the cohorts whose GEPA and HSPA Language Arts Literacy and Mathematics total scores were partially proficient and those students whose GEPA and HSPA Language Arts Literacy and Mathematics total scores were proficient or advanced proficient. In addition to determining the number and percentage for all students in the Cohort 1 and Cohort 2 groups in Language Arts Literacy and Mathematics, the percentages for the Special Needs districts as well as the DFG I and DFG J districts were also calculated.

Language Arts Literacy results are included in Table 9.1 and the Mathematics results are included in Table 9.2.

**TABLE 9.1**  
**LANGUAGE ARTS LITERACY**  
**Percentages of Students Across GEPA/HSPA Proficiency Levels**

<b>All</b>
<p><b>Cohort 1</b></p> <p>7.6% of the students received GEPA Partially Proficient scores and HSPA Proficient or Advanced Proficient scores.</p> <p>2.91% of the students received GEPA Proficient or Advanced Proficient scores and HSPA Partially Proficient scores.</p> <p><b>Cohort 2</b></p> <p>9.25% of the students received GEPA Partially Proficient scores and HSPA Proficient or Advanced Proficient scores.</p> <p>2.07% of the students received GEPA Proficient or Advanced Proficient scores and HSPA Partially Proficient scores.</p>
<b>Special Needs</b>
<p><b>Cohort 1</b></p> <p>14.9% of the students received GEPA Partially Proficient scores and HSPA Proficient or Advanced Proficient scores.</p> <p>5.82% of the students received GEPA Proficient or Advanced Proficient scores and HSPA Partially Proficient scores.</p> <p><b>Cohort 2</b></p> <p>19.7% of the students received GEPA Partially Proficient scores and HSPA Proficient or Advanced Proficient scores.</p> <p>3.91% of the students received GEPA Proficient or Advanced Proficient scores and HSPA Partially Proficient scores.</p>
<b>DFG I and DFG J</b>
<p><b>Cohort 1</b></p> <p>2.94% of the students received GEPA Partially Proficient scores and HSPA Proficient or Advanced Proficient scores.</p> <p>1.12% of the students received GEPA Proficient or Advanced Proficient scores and HSPA Partially Proficient scores.</p> <p><b>Cohort 2</b></p> <p>3.09% of the students received GEPA Partially Proficient scores and HSPA Proficient or Advanced Proficient scores.</p> <p>0.75% of the students received GEPA Proficient or Advanced Proficient scores and HSPA Partially Proficient scores.</p>

**TABLE 9.2**  
**MATHEMATICS**  
**Percentages of Students Across GEPA/HSPA Proficiency Levels**

<b>All</b>
<p><b>Cohort 1</b></p> <p>10.43% of the students received GEPA Partially Proficient scores and HSPA Proficient or Advanced Proficient scores.</p> <p>4.75% of the students received GEPA Proficient or Advanced Proficient scores and HSPA Partially Proficient scores.</p> <p><b>Cohort 2</b></p> <p>10.75% of the students received GEPA Partially Proficient scores and HSPA Proficient or Advanced Proficient scores.</p> <p>3.79% of the students received GEPA Proficient or Advanced Proficient scores and HSPA Partially Proficient scores.</p>
<b>Special Needs</b>
<p><b>Cohort 1</b></p> <p>10.6% of the students received GEPA Partially Proficient scores and HSPA Proficient or Advanced Proficient scores.</p> <p>9.04% of the students received GEPA Proficient or Advanced Proficient scores and HSPA Partially Proficient scores.</p> <p><b>Cohort 2</b></p> <p>12.8% of the students received GEPA Partially Proficient scores and HSPA Proficient or Advanced Proficient scores.</p> <p>7.09% of the students received GEPA Proficient or Advanced Proficient scores and HSPA Partially Proficient scores.</p>
<b>DFG I and DFG J</b>
<p><b>Cohort 1</b></p> <p>6.11% of the students received GEPA Partially Proficient scores and HSPA Proficient or Advanced Proficient scores.</p> <p>2.07% of the students received GEPA Proficient or Advanced Proficient scores and HSPA Partially Proficient scores.</p> <p><b>Cohort 2 DFG I and DFG J</b></p> <p>5.93% of the students received GEPA Partially Proficient scores and HSPA Proficient or Advanced Proficient scores.</p> <p>1.70% of the students received GEPA Proficient or Advanced Proficient scores and HSPA Partially Proficient scores.</p>

A possible source of criterion-related validity is the relationship of the GEPA scores to those received on the National Assessment of Educational Progress (NAEP).

The New Jersey assessments and NAEP have several similarities and major differences. The New Jersey assessments and the NAEP are based on content standards and frameworks that are revised or replaced on a regular basis to keep them in line with current instructional practices. Likewise, both the NAEP and New Jersey assessments create test specifications based on their respective frameworks that provide guidelines for developing the test items.

However, the New Jersey assessments and NAEP are distinctly different assessments because of:

- context and purpose,
- content and skills measured,
- item difficulty and formats, and
- method used for setting performance standards (i.e. cut points or achievement levels).

For these reasons, the New Jersey assessments and the NAEP, even in the same content area, may not yield comparable test results.

New Jersey results for the 2007 NAEP Reading and Mathematics tests for grade eight students included the following:

**TABLE 9.3**  
**NAEP Results for New Jersey Grade 8 Students**

	<b>Percentage of Students Below NAEP Basic</b>	<b>Percentage of Students At or Above NAEP Proficient</b>	<b>Average Scale Score</b>
<b>Reading</b>			
2007	19%	39%	270 (0 to 500 point scale)
2005	20%	38%	269 (0 to 500 point scale)
2003	21%	37%	268 (0 to 500 point scale)
<b>Mathematics</b>			
2007	23%	40%	289 (0 to 500 point scale)
2005	26%	36%	284 (0 to 500 point scale)
2003	28%	33%	281 (0 to 500 point scale)
1992	38%	24%	272 (0 to 500 point scale)
1990	42%	21%	270 (0 to 500 point scale)
<b>Science</b>			
2005	35%	33%	153 (0 to 300 point scale)

Further information about the NAEP and the New Jersey assessments is available online at <http://www.state.nj.us/education/assessment/naep/nj.shtml>

### 9.4 Consequential Validity Evidence (Evidence Based on Consequences of Testing)

Standard 13.1 in Chapter 13: Educational Testing and Assessment in Part 3: “Fairness in Testing,” of the *Standards* (1999) addresses intended and unintended consequences. A very similar standard appears as Standard 15.7 in Chapter 15: Testing in Program Evaluation and Public Policy of Part 3. Standard 13.1 is listed below:

---

When educational testing programs are mandated by school, district, state, or other authorities, the ways in which test results are intended to be used should be clearly described. It is the responsibility of those who mandate the use of tests to monitor their impact and to identify and minimize potential negative consequences. Consequences resulting from the uses of the test, both intended and unintended, should also be examined by the test user. (p. 145)

---

Beginning with the EWT due notice testing in 1991, the EWT and GEPA scores have provided districts information to help align their curriculum and instruction with the content and skills tested. The *Score Interpretation Manual* was developed to assist in the analysis and interpretation of GEPA score reports. The manual gives examples of uses of test results, discusses the various test scores, provides information about the appropriate score uses, and cautions against inappropriate score use.

Reports such as the District-Designed Reports were developed to provide districts with tools for organizing data to assist with instructional planning. Students’ score information is arranged on District Design Reports according to a school-developed plan to aggregate their students’ performance. School personnel code students’ answer folders following the school’s plan for grouping and organizing reports. For 2007, 53 districts requested District-Designed Reports for selected groups of students.

The return of student essays for instructional purposes has been an important aspect of Cycle II reporting. The *Cycle II Criterion-Based Holistic Scoring: A Writing Handbook* presents information about the scoring method and criteria used to evaluate student writing. The handbook offers suggestions for using New Jersey’s scoring rubrics and student test data to improve classroom instruction.

A number of materials including the *Cycle II Criterion-Based Holistic Scoring: Mathematics and Science Handbook*, *Cycle II Criterion-Based Holistic Scoring: A Reading Handbook*, *Cycle II Criterion-Based Holistic Scoring: A Writing Handbook*, and the Directory of Test Specifications and Sample Items for each of the GEPA content areas give guidance to teachers and curriculum developers for both instructional improvement and alignment.

Longitudinal graphs from 1999-2007 for Language Arts Literacy and Mathematics and from 2000-2007 for Science are available for the following groups:

- Subgroups – Total, General Education, Special Education, Limited English Proficient
- Gender – Female, Male
- Ethnicity – White, Black, Asian, Hispanic

The longitudinal graphs for the percent proficient and above by economic status appear in Figure 9.1 for Language Arts Literacy and Figure 9.2 for Mathematics for 1999–2006. The Language Arts Literacy graphs show that the proficient and above scores hovered between 46.2% and 50.6% for the economically disadvantaged students, and between 78.3% and 83.1% for the

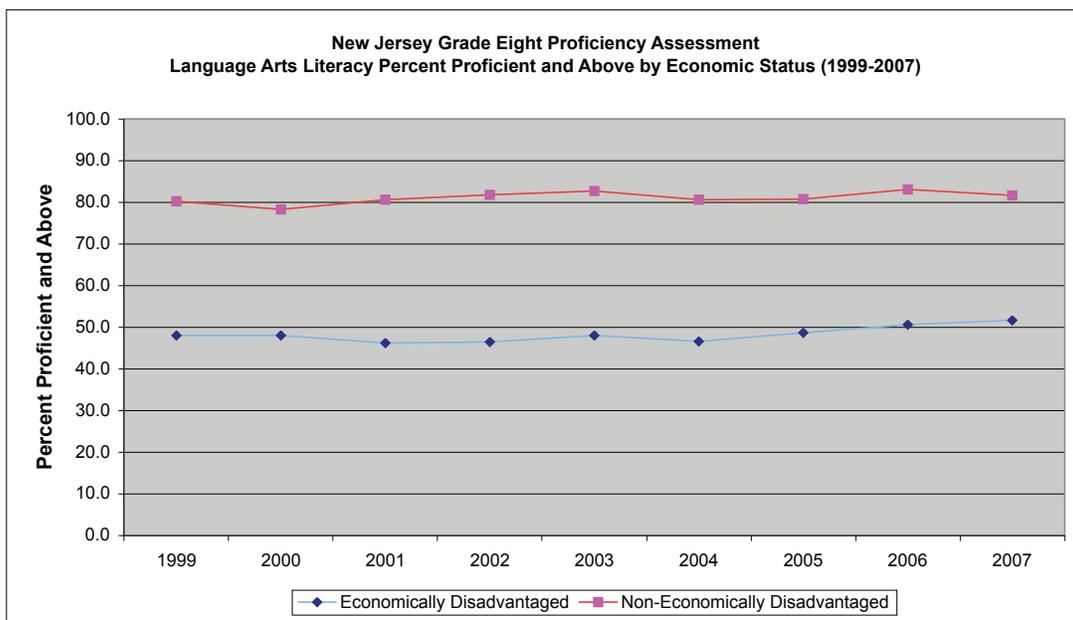
non-economically disadvantaged students. In 2007, 51.7% of the economically disadvantaged students had Language Arts Literacy proficient and above scores while 81.7% of the non-economically disadvantaged students had Language Arts Literacy proficient and above scores.

The graphs for Mathematics show generally increasing percents of students with proficient and above scores for both the economically disadvantaged and non-economically disadvantaged groups. The economically disadvantaged students with Mathematics proficient and above scores ranged from 25.4% on the 1999 Mathematics test administration to 38.4% on the 2006 test administration. In 2007, 44.6% of the economically disadvantaged students received Mathematics proficient and above scores. The non-economically disadvantaged students with Mathematics proficient and above scores ranged from 64.8% in the 1999 test administration and 63.7% in the 2000 test administration, to 74.2% on the 2006 test administration. In 2007, 77.4% of the non-economically disadvantaged students received Mathematics proficient and above scores.

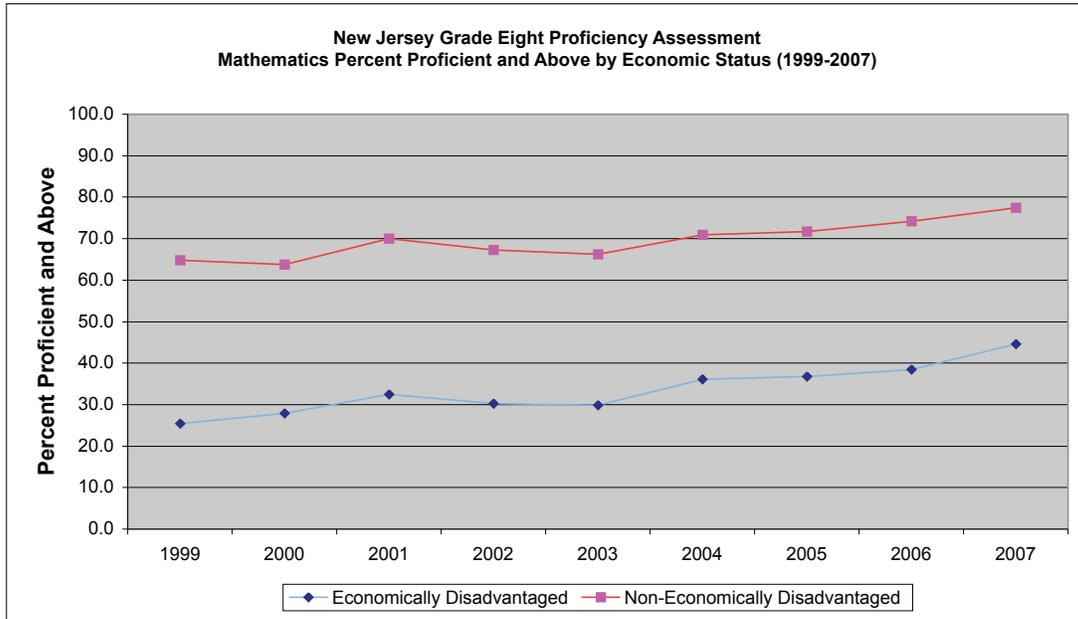
Similarly, Figure 9.3 shows Science 2000-2006 graphs with generally increasing percents for proficient and above scores for the economically disadvantaged and non-economically disadvantaged student groups. The economically disadvantaged students with Science proficient and above scores ranged from 36.9% on the 2000 Science test administration to 57.1% in the 2006 test administration. In 2007, 56.7% of the economically disadvantaged students received Science proficient and above scores. The non-economically disadvantaged students with Science proficient and above scores ranged from 78.2% in the 2000 Science test administration to 87.2% on the 2006 test administration. Also in 2007, 87.2% of the non-economically disadvantaged students received Science proficient and above scores.

The complete group of longitudinal graphs are available online at: <http://www.state.nj.us/education/schools/achievement/2008/gepa/graphs.pdf>

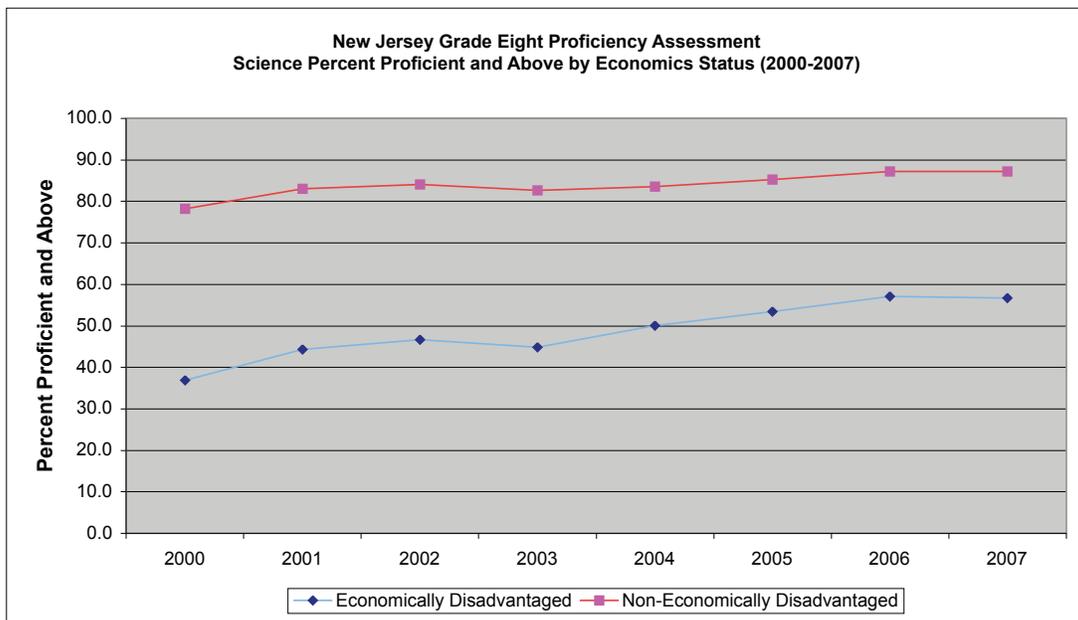
**FIGURE 9.1**  
**LANGUAGE ARTS LITERACY**  
**Longitudinal Graph by Economic Status**



**FIGURE 9.2**  
**MATHEMATICS**  
**Longitudinal Graph by Economic Status**



**FIGURE 9.3**  
**SCIENCE**  
**Longitudinal Graph by Economic Status**



## CHAPTER 10: REPORTING

To help school personnel identify the needs of eighth-grade students tested and to assist in the evaluation of school and district programs, a variety of reports are produced and distributed.

The GEPA reports were produced in two cycles:

- Cycle I reports, including Individual Student Reports and preliminary school and district aggregate reports, were received in the districts in mid-June.
- Cycle II reports, including cluster means reports and performance reports for demographic groups, were received in the districts in mid-July.

Cycle II data is used by the Office of Title I Program and Planning and Accountability for Adequate Yearly Progress (AYP) calculations. The State Summary is completed and posted on the NJDOE website in January.

Beginning in 2007, reports were provided electronically from Pearson's NCS SchoolHouse via controlled secure web access. Using district specific passwords, district offices downloaded and distributed their district and school reports. However, paper reporting continued for Individual Student Reports and Student Stickers as well as district summaries, school summaries, and All Sections Rosters for the counties.

### 10.1 Information on the Reports

The Cycle I and Cycle II score reports are designed to show a range of student identification and score information to assist school personnel with identifying the needs of their students and recognizing weaknesses in instructional programs.

**Student Identification** - Score reports display student demographic information gridded on the answer documents or submitted on a pre-ID label files. Prior to reporting, a roster showing the students' demographic information was distributed to school districts to provide an opportunity for corrections.

In addition to the student's name and the Test ID Number assigned to the student, the following information is collected:

- Date of Birth (DOB)
- Gender is indicated by M (male) or F (female).
- Ethnic codes
- Y (for yes) is indicated in the TIS<1 column if a student was coded as being enrolled in the school for less than a year.
- Y (for yes) is indicated in the TID<1 column if a student was coded as being enrolled in the district for less than a year.
- <, 1, 2, 3, F1 or F2 is indicated in the LEP column if a student was coded as limited English proficient (see LEP in Appendix F). If multiple bubbles were colored, a Y will appear in this column.
- A through M (see SE codes in Appendix F) is indicated in the SE column if a student was coded as a special education student.
- The first letter of a content area (L, M, and S) is indicated in the APA column if a student was coded as taking the APA.
- The first letter of a content area (L, M, and S) is indicated in the T-I column if a student was coded as receiving Title I services for any of the three content areas.

- Y (for yes) is indicated in the ED column if a student was coded as Economically Disadvantaged.
- Y (for yes) is indicated for students coded as having Migrant status.
- Y (for yes) is indicated for students coded by their receiving school [public or private] as being an Out of District placement student.
- Y (for yes) is indicated for students coded as being an Out of Residence Placement student.

**Void Codes** – Immediately following testing, examiners mark if a student’s answer document should be voided due to illness, disruptive behavior, or some other reason. The answer folder is not scored and a void code is printed in place of the total test score on the student’s reports. These void codes are as follows:

V1 (voided due to illness)

V2 (voided due to cheating or disruptive behavior)

V3 (voided due to the student not being an eighth grader)

V5 (voided due to breach of security by a school or district).

Also, a student’s answer document may be voided at the time of scoring. For Mathematics and Science, if a student attempted less than 20 percent of the items, no cluster data will appear and, instead of the content area score, the report will list a V4. For Language Arts Literacy, if a student attempted less than 20 percent of the items on one of the testing days but did attempt 20 percent or more on the other testing day, a V4 will appear instead of the Language Arts Literacy score, but cluster data will be provided on the report.

During the 2007 administration, 152 Mathematics and 149 Science tests were voided due to the attempted criteria. For Language Arts Literacy, 268 tests were voided due to the attempted criteria for Day 1 and 347 tests were voided due to the attempted criteria for Day 2.

**Score Information** – The total GEPA Language Arts Literacy, Mathematics, and Science scores are reported as scale scores with a range of 100 to 300. The scores of 100 and 300 are a theoretical floor and ceiling which may not actually be observed. The scale score of 250 is the cut point between Proficient students and Advanced Proficient students. The scale score of 200 is the cut point between Partially Proficient students and Proficient students. The score ranges are as follows:

<i>Advanced Proficient</i>	<i>250 – 300</i>
<i>Proficient</i>	<i>200 – 249</i>
<i>Partially Proficient</i>	<i>100 – 199</i>

The scores of students who are included in the Partially Proficient level are considered to be below the state minimum level of proficiency. These students may need additional instructional support, which could be in the form of individual and programmatic intervention. District staff should consider multiple measures for all students before making decisions about students’ instructional placement.

In addition to the total GEPA scores in Language Arts Literacy, Mathematics, and Science, various score reports contain the following information for each cluster (scores at the cluster level are raw scores):

- **Points Earned** – This number represents the number of score points a student received for a given cluster. On the Student Roster for Language Arts Literacy, the “Points Earned” is provided for Reading and Writing as well as for each of the writing tasks.
- **Just Proficient Mean** – This number represents the average (mean) number of score points received for each cluster by all students in the state whose scale scores are 200 for a particular content. Students who took Large-Print or Braille forms are excluded from calculating just proficient means.

**Automatic Rescores** – The scoring process entails an automatic adjudication of scoring on open-ended items for students whose scores are close to, but not over, the proficiency level. In 2003, GEPA adopted automatic rescoring of all open-ended responses for all students who received a scale score ranging from 197 to 199. This process was replaced for the 2007 administration. Beginning with the 2007 administration, automatic rescoring is conducted for any student whose raw score total falls within two raw score points of a proficiency level (Proficient or Advanced Proficient).

### 10.2 Types of Reports

#### Cycle I Reports

##### Individual Student Report (ISR) and Student Sticker

The Individual Student Report (ISR) is a two-sided report showing specific student score information on the front of the ISR. A description of the GEPA and an interpretation of the ISR scores are printed on the back. Figure 10.1 presents the front of a student’s sample report with demographic information, scale scores, proficiency levels, and cluster raw scores and Just Proficient Means. Figure 10.2 shows the GEPA description and ISR interpretation printed for all students.

Two copies of the ISR are produced for every student tested. After educators and school staff analyze the score information on the front of the ISR, one copy is placed in the student’s permanent folder and the other copy is shared with the student’s parent/guardian in a manner determined by the local district. When a student attends a private school as an Out of District Placement student, a third copy of the ISR is produced and sent to the private school.

A student’s scale scores and proficiency levels with the student’s identification information are printed on a peel-off label for attaching to a student’s permanent folder.

##### All Sections Roster

The All Sections Roster, an alphabetical listing of students’ names, provides students’ identification and score information. Each student’s scale scores with proficiency levels are listed for the three content areas. Users of this report can quickly determine how a particular student performed in each of the three content areas. The All Sections Roster provides the most complete listing of the student identification information with codes.

**FIGURE 10.1**  
**Individual Student Report (ISR Front)**



Test Date: March 2007

County: 99 MIDSTATES  
District: 9999 MIDSTATES  
School: 99 MIDSTATES M.S.

**Student Name: MARLL, LZLF**

Test ID No: 002797000  
Answer Folder No: 02466  
District/School Student ID No:

Date of Birth: 01/11/YY  
Gender: M

LEP:  
SE:  
APA:  
Title I:

Content Area	Your Scale Score	Proficiency Level
Language Arts Literacy	229	Proficient
Mathematics	204	Proficient
Science	256	Advanced Proficient

Partially Proficient: Scale Score BELOW 200  
Proficient: Scale Score AT OR ABOVE 200 but BELOW 250  
Advanced Proficient: Scale Score AT OR ABOVE 250

Language Arts Literacy	Mathematics	Science																																																															
<p>The Language Arts Literacy section assesses a student's abilities in the following clusters. A check mark indicates areas of possible strength.</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Cluster</th> <th>Your Points</th> <th>Just Proficient Mean</th> </tr> </thead> <tbody> <tr> <td>Writing</td> <td>10.0 out of 18</td> <td>9.2</td> </tr> <tr> <td>Reading</td> <td>26.0 out of 36</td> <td>20.3</td> </tr> <tr> <td colspan="3"><hr/></td> </tr> <tr> <td>Interpreting Text</td> <td>13.0 out of 15</td> <td>8.8</td> </tr> <tr> <td>Analyzing/Critiquing Text</td> <td>13.0 out of 21</td> <td>11.5</td> </tr> </tbody> </table>	Cluster	Your Points	Just Proficient Mean	Writing	10.0 out of 18	9.2	Reading	26.0 out of 36	20.3	<hr/>			Interpreting Text	13.0 out of 15	8.8	Analyzing/Critiquing Text	13.0 out of 21	11.5	<p>The Mathematics section assesses a student's abilities in the following clusters. A check mark indicates areas of possible strength.</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Cluster</th> <th>Your Points</th> <th>Just Proficient Mean</th> </tr> </thead> <tbody> <tr> <td>Number &amp; Numerical Operations</td> <td>8.0 out of 12</td> <td>6.7</td> </tr> <tr> <td>Geometry &amp; Measurement</td> <td>5.0 out of 12</td> <td>4.6</td> </tr> <tr> <td>Patterns &amp; Algebra</td> <td>4.0 out of 12</td> <td>5.0</td> </tr> <tr> <td colspan="3"><hr/></td> </tr> <tr> <td>Data Analysis, Probability, &amp; Discrete Mathematics</td> <td>9.0 out of 12</td> <td>7.7</td> </tr> <tr> <td>Knowledge</td> <td>26.0 out of 48</td> <td>24.0</td> </tr> <tr> <td>Problem Solving Skills</td> <td>19.0 out of 35</td> <td>16.5</td> </tr> </tbody> </table>	Cluster	Your Points	Just Proficient Mean	Number & Numerical Operations	8.0 out of 12	6.7	Geometry & Measurement	5.0 out of 12	4.6	Patterns & Algebra	4.0 out of 12	5.0	<hr/>			Data Analysis, Probability, & Discrete Mathematics	9.0 out of 12	7.7	Knowledge	26.0 out of 48	24.0	Problem Solving Skills	19.0 out of 35	16.5	<p>The Science section assesses a student's abilities in the following clusters. A check mark indicates areas of possible strength.</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Cluster</th> <th>Your Points</th> <th>Just Proficient Mean</th> </tr> </thead> <tbody> <tr> <td>Life Science</td> <td>16.0 out of 22</td> <td>9.2</td> </tr> <tr> <td>Physical Science</td> <td>12.0 out of 16</td> <td>6.4</td> </tr> <tr> <td>Earth Science</td> <td>14.0 out of 16</td> <td>6.3</td> </tr> <tr> <td colspan="3"><hr/></td> </tr> <tr> <td>Knowledge</td> <td>9.0 out of 11</td> <td>4.8</td> </tr> <tr> <td>Application</td> <td>33.0 out of 43</td> <td>17.2</td> </tr> </tbody> </table>	Cluster	Your Points	Just Proficient Mean	Life Science	16.0 out of 22	9.2	Physical Science	12.0 out of 16	6.4	Earth Science	14.0 out of 16	6.3	<hr/>			Knowledge	9.0 out of 11	4.8	Application	33.0 out of 43	17.2
Cluster	Your Points	Just Proficient Mean																																																															
Writing	10.0 out of 18	9.2																																																															
Reading	26.0 out of 36	20.3																																																															
<hr/>																																																																	
Interpreting Text	13.0 out of 15	8.8																																																															
Analyzing/Critiquing Text	13.0 out of 21	11.5																																																															
Cluster	Your Points	Just Proficient Mean																																																															
Number & Numerical Operations	8.0 out of 12	6.7																																																															
Geometry & Measurement	5.0 out of 12	4.6																																																															
Patterns & Algebra	4.0 out of 12	5.0																																																															
<hr/>																																																																	
Data Analysis, Probability, & Discrete Mathematics	9.0 out of 12	7.7																																																															
Knowledge	26.0 out of 48	24.0																																																															
Problem Solving Skills	19.0 out of 35	16.5																																																															
Cluster	Your Points	Just Proficient Mean																																																															
Life Science	16.0 out of 22	9.2																																																															
Physical Science	12.0 out of 16	6.4																																																															
Earth Science	14.0 out of 16	6.3																																																															
<hr/>																																																																	
Knowledge	9.0 out of 11	4.8																																																															
Application	33.0 out of 43	17.2																																																															

**Note: The scores in this report are for illustrative purposes only.** For example, the raw score to scale score conversions are not necessarily the same as those used in the actual reports.

## FIGURE 10.2

### Individual Student Report (ISR Back)

#### ABOUT THE GRADE EIGHT PROFICIENCY ASSESSMENT

The Grade Eight Proficiency Assessment (GEPA) was administered in March 2007 to approximately 108,000 eighth-graders throughout New Jersey. The 2007 GEPA measured students' knowledge and skills in three important content areas — Language Arts Literacy, Mathematics, and Science. The GEPA focuses on the knowledge and skills needed for future success in high school and beyond high school.

The GEPA Language Arts Literacy section measures both reading and writing. The Reading component asks students to read real-world texts and to answer related questions. It contains both multiple-choice and open-ended items. The open-ended items require students to write a few sentences or a few paragraphs to answer a question about the text. The Writing component asks students to write two essays. Both the tasks in the Writing component require students to write a response.

The GEPA Mathematics section measures students' abilities to solve problems using mathematical concepts. The components in this section measure: Number and Numerical Operations; Geometry and Measurement. Patterns and Algebra; Data Analysis, Probability, and Discrete Mathematics. The Mathematics section, like the Reading component of the Language Arts Literacy section, contains both multiple-choice and open-ended items. The open-ended items require students to solve a problem as well as explain their solution.

The GEPA Science section measures students' knowledge and skills in Life Science, Physical Science, and Earth Science. The Science section contains both multiple-choice and open-ended items. The open-ended items require students to respond to a question as well as explain their answer.

Carefully trained readers score the open-ended items and essays. Two readers independently score each student's response to ensure the accuracy of each student's results.

The GEPA serves as an early warning for possible difficulties on the High School Proficiency Assessment (HSPA). Beginning with students who entered their junior year in 2001–2002, the HSPA is the "test of record" for receiving a high school diploma. All students are required to receive a score of "proficient" or higher on each section of the HSPA to earn a New Jersey-endorsed high school diploma. (The HSPA replaced the High School Proficiency Test or HSPT.)

#### HOW TO READ THIS REPORT

This **Individual Student Report** presents your child's Language Arts Literacy, Mathematics, and Science scores on the 2007 GEPA. The report is available only to parents, guardians, students, and authorized school personnel. If you have any questions about the report or how to interpret the scores, you should contact your child's teacher or principal.

Your child's total GEPA Language Arts Literacy, Mathematics, and Science scores are presented in the box on the top half of the report. In the column labeled **Your Score**, your child's scale score for each GEPA section is printed. To the right of the scale score is a column labeled **Proficiency Level**. If the scale score is below 200, your child is "Partially Proficient" in that content area. If the scale score is at or above 200 but below 250, your child is "Proficient" in that content area. Finally, if the scale score is at or above 250, your child is "Advanced Proficient" in that content area. Scores below 200 may indicate a need for additional instructional assistance. However, like any single test score, GEPA results should not be used as the sole basis for instructional decisions.

Additional information to assist in identifying your child's strengths and weaknesses is presented on the bottom half of the report. Cluster-level results show how your child performed on the sets of items that measure particular knowledge and skills (clusters above the line) or particular processes (clusters below the line). Though an item on the GEPA can contribute to a cluster above the line (for example, Reading) as well as a cluster below the line (for example, Interpreting Text), each item is counted only once in the total score.

For each cluster, the column to the right of the cluster name, labeled, **Your Score** shows the number of points your child earned on the clusters. The right-most column for each content area, labeled **Just Proficient Mean**, is a yardstick against which to measure your child's cluster score. It is calculated by taking the average of the raw scores of all the students across the state whose scale scores are 200 for a particular content area. For example, among all students who scored at 200 for the GEPA Language Arts Literacy section, the reading cluster score was 20.3. If your child scored at or above 20.3 for this cluster, a check mark (✓) indicates this area is a possible strength for your child. If your child scored below 20.3 on this cluster, he or she may need additional help in this area.

At the top of the report, a notation may appear if, for some reason, your child's test booklet was not scored. These reasons include illness (V1), cheating or disruptive behavior (V2), or not an eighth grader (V3). For Mathematics and Science, if a student attempted less than 20% of the items, no cluster data will appear and instead of **Your Score** the report will indicate a V4. For Language Arts Literacy, if a student attempted less than 20% of the items on each of the two testing days, no cluster data will appear and instead of **Your Score** for Language Arts Literacy the report will indicate a V4. If a student attempted at least 20% of the items on one of the two testing days, but did not attempt 20% on the other testing day, a V4 will appear in **Your Score** for Language Arts Literacy but cluster scores will be provided on the report. If a student did not take a section of the test, no cluster data will appear and instead of your score, the report will indicate **Not Present**.

### **Student Roster – Language Arts Literacy, Mathematics, and Science**

Three Student Rosters are produced – one for each content area. Students’ names are listed in descending order of the content area scores. Figure 10.3 shows an example of the Student Roster – Mathematics listing the student with the highest score mathematics score first followed with the other students in this school. A dashed line is printed across the roster after the last student in each proficiency level.

No students in the example shown in Figure 10.3 had scores at or above 250, the Advanced Proficient cut point. A dashed line appears across the roster under 200, the Proficient cut point. Students whose answer documents were voided and students who were coded indicating they were taking the Alternate Proficiency Assessment (APA) are listed alphabetically at the end of each content area roster.

### **Summary of School Performance and Summary of District Performance**

A Summary of School Performance is produced for each of the three content areas and a Summary of District Performance is produced for each of the three content areas. The report for each content area provides the number and percent of students in each proficiency level as well as the number of general education students, special education students (including students coded as taking the APA), and limited English proficient students tested for the content area.

The total test information includes the school or district mean for the reported content area. In addition, the means are provided for each of the clusters. The total test and cluster means are printed for these student groups: total, general education, special education, limited English proficient, current LEP, and former LEP.

The following summary information is provided for each subgroup shown on the report:

- Number Enrolled: total number of answer folders returned
- Number Not Present: number of answer folders returned that were totally blank excluding answer folders coded as APA
- Number of Voids: number of answer folders coded void by the school [V1, V2, and V3] AND coded void due to less than 20% of the test items being taken, including answer folders coded as APA [V4] AND coded void due to a security breach [V5].
- Number of Valid Score Scores: total number of students tested excluding not present and voids
- Total number of students who scored in each proficiency level, excluding students coded as APA who did not take the GEPA
- Percent of students who scored in each proficiency level, excluding students coded as APA who did not take the GEPA

### **Preliminary Performance by Demographic Groups – State Report**

This preliminary report is produced with the Cycle I reports prior to the completion of the automatic rescoring. The one-page report presents the results for the total, general education, special education, and limited English proficient student groups, and by gender, migrant status, ethnicity, and economic status. Data are based on scale scores and the percentage of students that fall into each of the three proficiency levels. The report does not show cluster level data.

**FIGURE 10.3**  
**Student Roster (Mathematics)**



**New Jersey Statewide Testing System**  
**Grade Eight Proficiency Assessment**  
**Student Roster - Mathematics**

TEST DATE: MARCH 2007  
REPORT PRINTED:

COUNTY: 99 MIDSTATES  
DISTRICT: 9999 MIDSTATES  
SCHOOL: 99 MIDSTATES M.S.

PAGE: 19

ANSWER FOLDERS PROCESSED: 36

POINTS EARNED

STUDENT NAME/ TEST ID NO.	DATE OF BIRTH	GENDER	LEP	SE	APA	T-I	TOTAL SCORE						
							NUMBER & NUMERICAL OPERATIONS	GEOMETRY & MEASUREMENT	PATTERNS & ALGEBRA	DATA ANALYSIS, PROBABILITY, & DISCRETE MATHEMATICS	KNOWLEDGE	PROBLEM SOLVING SKILLS	
LINBRLIN 002797198	08/16/92	F					12 <sup>1</sup> 6.7 <sup>2</sup>	12 <sup>1</sup> 4.6 <sup>2</sup>	12 <sup>1</sup> 5.0 <sup>2</sup>	12 <sup>1</sup> 7.7 <sup>2</sup>	48 <sup>1</sup> 24.0 <sup>2</sup>	35 <sup>1</sup> 16.5 <sup>2</sup>	16.0
MERDJB 002797147	06/29/92	M					8.0	4.0	3.0	9.0	24.0	17.0	17.0
NARTH 002797227	08/08/93	M		I			8.0	5.0	7.0	4.0	24.0	17.0	15.0
TALJBB 002797315	02/20/92	M					7.0	5.0	6.0	6.0	24.0	15.0	18.0
TKPBPDB 002797335	08/20/92	M		B			8.0	5.0	3.0	8.0	24.0	18.0	19.0
TQBDZ 002797353	02/28/92	M					6.0	5.0	4.0	9.0	24.0	19.0	16.0
XINTFRT 002797381	06/24/92	F					9.0	4.0	4.0	7.0	24.0	16.0	14.0
ZUPPBRDO 002797388	10/16/93	F		E			7.0	4.0	6.0	7.0	24.0	14.0	---
-----													
DANBVITO 002797045	12/20/93	M		A			10.0	3.0	4.0	7.0	24.0	22.0	16.0
BKLFN 002797029	10/30/93	M					6.0	5.0	4.0	8.0	23.0	16.0	15.0
BORDFN 002797048	08/31/93	F		I			7.0	5.0	4.0	7.0	23.0	15.0	15.0
DORRJFRE 002797092	10/24/93	M					5.0	5.0	6.0	7.0	23.0	15.0	14.0
NYFRT 002797253	08/04/93	F		I			9.0	6.0	2.0	6.0	23.0	14.0	17.0
SEDLINO 002797274	08/12/93	F					8.0	5.0	5.0	5.0	23.0	17.0	16.0
TOBRBNDERA 002797340	02/24/92	F					6.0	3.0	7.0	7.0	23.0	16.0	15.0
BERNBDDHI 002797041	10/10/92	F					6.0	4.0	4.0	8.0	22.0	15.0	---

<sup>1</sup> THE NUMBERS IN THIS ROW ARE THE NUMBER OF POINTS POSSIBLE.  
<sup>2</sup> THE NUMBERS IN THIS ROW ARE THE STATEWIDE RAW SCORE MEANS FOR THE STUDENTS WHOSE SCALE SCORE IS 200.

Note: The scores in this report are for illustrative purposes only. For example, the raw score to scale score conversions are not necessarily the same as those used in the actual reports.

## **Preliminary Performance by Demographic Groups – School and District Reports**

This report is produced before the rescore is completed. This report does not break the data out at the cluster level. Data are based on scale scores and the percentage of students who fall into each of the three proficiency levels.

### **Cycle II Reports**

The Cycle II reports include a final Performance by Demographic Groups report that reflects any changes that may have occurred during the processing of automatic rescoring.

## **School and District Cluster Means Reports**

Figure 10.4 shows an example of the School Cluster Means Report – Language Arts Literacy. The School and District Cluster Means reports consist of three reports – one for each content area.

The first column on the report presents the mean cluster scores for students in the state whose scale score is 200, i.e., students who are “just proficient.” Data include raw score means of all students (total, general education, special education, limited English proficient, and Title I student groups) at the cluster level for each content area. A similar format is used for both the School Reports and District Reports. The District Reports present aggregated data for the district, DFG, and the state. Additionally, the School Reports show school level data.

## **District-Designed Reports**

The District-Designed Reports are similar to the School Cluster Means Reports except schools create the reports for selected groups of students. Schools used a “special” code category on the GEPA answer documents to obtain cluster means for selected student groups. Like the School Cluster Means Reports, a District-Designed Report is produced for each content area.

Student answer documents may be coded in any of the four two-column “Special Codes” grids labeled A, B, C, or D. These special codes were assigned by the school during the test administrations. The special code, as coded on the students’ answer folders, is printed in the report title. Student groups must contain ten or more students.

**FIGURE 10.4**  
**Cluster Means Report**

**New Jersey Statewide Testing System  
Grade Eight Proficiency Assessment  
Cluster Means<sup>1</sup>**



Test Date: March 2007

**CYCLE II**

COUNTY: 99 MIDSTATES  
DISTRICT: 9999 MIDSTATES  
SCHOOL: 99 MIDSTATES M.S.

School

	JUST PROFICIENT MEAN <sup>2</sup>	SCHOOL MEAN	DISTRICT MEAN	DFG FG MEAN	STATE MEAN
<b>Language Arts Literacy</b>					
<b>Clusters</b>					
1. Writing <sup>(18)</sup> Total Students <sup>3</sup> General Education Special Education Limited English Proficient <sup>4</sup> Current LEP Former LEP Title I	9.2	10.1 10.4 8.6	10.2 10.5 8.7 8.5 8.5	10.1 10.5 8.1 8.0 7.8 8.8 9.4	9.9 10.4 7.8 7.8 7.2 9.1 8.8
2. Reading <sup>(36)</sup> Total Students <sup>3</sup> General Education Special Education Limited English Proficient <sup>4</sup> Current LEP Former LEP Title I	20.3	24.3 25.2 19.8	24.0 24.7 19.6 19.3 19.3	23.3 24.2 18.2 16.9 16.4 18.9 20.0	22.5 23.8 17.2 16.1 14.7 19.3 18.9
3. Interpreting Text <sup>(15)</sup> Total Students <sup>3</sup> General Education Special Education Limited English Proficient <sup>4</sup> Current LEP Former LEP Title I	8.8	10.7 11.1 8.9	10.7 11.0 8.7 7.7 7.7	10.3 10.8 8.1 7.3 7.1 8.1 8.7	9.9 10.5 7.6 7.0 6.4 8.3 8.2
4. Analyzing/Critiquing Text <sup>(21)</sup> Total Students <sup>3</sup> General Education Special Education Limited English Proficient <sup>4</sup> Current LEP Former LEP Title I	11.5	13.5 14.1 10.9	13.3 13.7 10.9 11.6 11.6	13.0 13.5 10.2 9.6 9.3 10.8 11.3	12.6 13.3 9.7 9.1 8.3 10.9 10.7

\* Refers to total raw score points in each cluster.  
<sup>1</sup> Cluster means exclude students who took Large print, Braille, and alternate forms as well as students coded as voids.  
<sup>2</sup> Just Proficient Means are statewide raw score means for students whose scale score is 200.  
<sup>3</sup> Students appear in each applicable category, but they are included in Total Students only once.  
<sup>4</sup> Includes students coded as LEP or former LEP.

**Note: The scores on this report are for illustrative purposes only.**

**Performance by Demographic Groups – State Report**

Performance by Demographic Groups – State Report summarizes statewide total population data collected from districts regarding total, general education (GE), special education (SE), Limited English Proficiency (LEP), gender, migrant status, ethnicity, and economic status (disadvantaged vs. not disadvantaged). This report includes data from all three content areas. Data are based on scale scores and the percentage of students who fall into each of the three proficiency levels. The report does not break out the data at the cluster level.

The Cycle II Test Results in Appendix B include the Performance by Demographic Groups – State Report.

**Performance by Demographic Groups – School and District Reports**

Performance by Demographic Groups – School and District Reports present results by total, general education, special education, Limited English Proficiency, gender, migrant status, ethnicity, and economic status (disadvantaged vs. not disadvantaged) for all three content areas. These group reports provide additional achievement information that can be used to make adjustments to curricula that may better serve these student subgroups. Figure 10.5 shows an example of the school level Performance by Demographic Groups.

Similar to the Performance by Demographic Groups – State Report, data included are based on scale scores and the percentage of students who fall into each of the three proficiency levels. The reports do not break out the data at the cluster level.

**FIGURE 10.5**  
**Performance by Demographic Groups**



**New Jersey Statewide Testing System  
Grade Eight Proficiency Assessment  
Preliminary Performance by Demographic Groups  
School**

Test Date: March 2007

COUNTY: 99 MIDSTATES  
DISTRICT: 9999 MIDSTATES  
SCHOOL: 99 MIDSTATES M.S.

	Language Arts Literacy						Mathematics						Science												
	Students Enrolled	APA Students	Not Present	Voids <sup>1</sup>	Valid Scale Scores	% Partially Proficient	% Proficient	Advanced Proficient	Scale Score Mean	APA Students	Not Present	Voids	Valid Scale Scores	% Partially Proficient	% Proficient	Advanced Proficient	Scale Score Mean	APA Students	Not Present	Voids	Valid Scale Scores	% Partially Proficient	% Proficient	Advanced Proficient	Scale Score Mean
<b>Total Students<sup>2</sup></b>	368	4	0	2	362	14.6	74.6	10.8	222.4	7	0	1	360	13.1	63.1	23.9	226.0	4	1	0	363	5.2	66.4	28.4	232.5
General Education <sup>3</sup>	300	0	0	0	300	7.0	80.0	13.0	227.6	0	0	0	300	6.7	66.3	27.0	230.7	0	0	0	300	3.0	65.7	31.3	235.5
Special Education	67	4	0	1	62	51.6	48.4	0.0	197.0	7	0	1	59	44.1	47.5	8.5	202.9	4	1	0	62	14.5	71.0	14.5	218.7
Limited English Proficient <sup>4</sup>	1	0	0	1	0	0.0	0.0	0.0	0.0	0	0	0	1	100.0	0.0	0.0	175.0	0	0	0	1	100.0	0.0	0.0	167.0
Current LEP	1	0	0	1	0	0.0	0.0	0.0	0.0	0	0	0	1	100.0	0.0	0.0	175.0	0	0	0	1	100.0	0.0	0.0	167.0
Former LEP	0	0	0	0	0	0.0	0.0	0.0	0.0	0	0	0	0	0.0	0.0	0.0	0.0	0	0	0	0	0.0	0.0	0.0	0.0
<b>Gender<sup>5</sup></b>																									
Female	182	3	0	2	177	8.5	75.7	15.8	229.1	4	0	1	177	12.4	67.8	19.8	224.1	3	0	0	179	7.8	70.4	21.8	229.4
Male	186	1	0	0	185	20.5	73.5	5.9	215.9	3	0	0	183	13.7	58.5	27.9	227.7	1	1	0	184	2.7	62.5	34.8	235.5
<b>Migrant Status</b>																									
Migrant	0	0	0	0	0	0.0	0.0	0.0	0.0	0	0	0	0	0.0	0.0	0.0	0.0	0	0	0	0	0.0	0.0	0.0	0.0
Non-Migrant	368	4	0	2	362	14.6	74.6	10.8	222.4	7	0	1	360	13.1	63.1	23.9	226.0	4	1	0	363	5.2	66.4	28.4	232.5
<b>Ethnicity</b>																									
White	329	4	0	1	324	13.0	76.2	10.8	223.0	7	0	1	321	11.8	64.2	24.0	226.2	4	0	0	325	5.2	67.4	27.4	232.5
Black	15	0	0	0	15	33.3	66.7	0.0	213.7	0	0	0	15	20.0	53.3	26.7	231.0	0	0	0	15	6.7	46.7	239.8	
Asian	9	0	0	0	9	22.2	55.6	22.2	222.9	0	0	0	9	22.2	33.3	44.4	230.2	0	0	0	9	0.0	44.4	55.6	237.3
Pacific Islander	0	0	0	0	0	0.0	0.0	0.0	0.0	0	0	0	0	0.0	0.0	0.0	0.0	0	0	0	0	0.0	0.0	0.0	0.0
Hispanic	15	0	0	1	14	28.6	57.1	14.3	216.5	0	0	0	15	26.7	66.7	6.7	213.5	0	1	0	14	7.1	78.6	14.3	220.9
Amer Indian/AK Native	0	0	0	0	0	0.0	0.0	0.0	0.0	0	0	0	0	0.0	0.0	0.0	0.0	0	0	0	0	0.0	0.0	0.0	0.0
Others	0	0	0	0	0	0.0	0.0	0.0	0.0	0	0	0	0	0.0	0.0	0.0	0.0	0	0	0	0	0.0	0.0	0.0	0.0
<b>Economic Status</b>																									
Econ Disadvantaged	21	1	0	1	19	15.8	73.7	10.5	221.9	1	0	0	20	15.0	65.0	20.0	222.5	1	0	0	20	10.0	65.0	25.0	230.7
Non-Econ Disadvantaged	347	3	0	1	343	14.6	74.6	10.8	222.4	6	0	1	340	12.9	62.9	24.1	226.2	3	1	0	343	5.0	66.5	28.6	232.6

<sup>1</sup> Includes students coded as LEP exempt (LAL only).  
<sup>2</sup> Students appear in each applicable category, but are included in Total Students only once.  
<sup>3</sup> Includes students coded as former LEP.  
<sup>4</sup> Excludes students who did not have gender coded.  
<sup>5</sup> Excludes students who did not have ethnicity coded.  
<sup>6</sup> Includes students who did not have an ethnicity coded and students who had more than one ethnicity coded.

**Note: The scores on this report are for illustrative purposes only.**

# ***Appendix A***

## ***Scoring Rubrics and 3rd Reader Score Calculation Charts***

# NEW JERSEY REGISTERED HOLISTIC SCORING RUBRIC

In scoring, consider the grid of written language	Inadequate Command	Limited Command	Partial Command	Adequate Command	Strong Command	Superior Command
Score	1	2	3	4	5	6
<b>Content and Organization</b>	<ul style="list-style-type: none"> <li>May lack opening and/or closing</li> <li>Minimal response to topic; uncertain focus</li> <li>No planning evident; disorganized</li> <li>Details random, inappropriate, or barely apparent</li> <li>No apparent control</li> <li>Severe/numerous errors</li> </ul>	<ul style="list-style-type: none"> <li>May lack opening and/or closing</li> <li>Attempts to focus</li> <li>May drift or shift focus</li> <li>Attempts organization</li> <li>Few, if any, transitions between ideas</li> <li>Details lack elaboration, i.e., highlight paper</li> <li>Numerous errors</li> <li>Excessive monotony/same structure</li> <li>Numerous errors</li> <li>Numerous serious errors</li> </ul>	<ul style="list-style-type: none"> <li>May lack opening and/or closing</li> <li>Usually has single focus</li> <li>Some lapses or flaws in organization</li> <li>May lack some transitions between ideas</li> <li>Repetitious details</li> <li>Several unelaborated details</li> <li>Errors/patterns of errors may be evident</li> <li>Little variety in syntax</li> <li>Some errors</li> <li>Patterns of errors evident</li> </ul>	<ul style="list-style-type: none"> <li>Generally has opening and/or closing</li> <li>Single focus</li> <li>Ideas loosely connected</li> <li>Transitions evident</li> <li>Uneven development of details</li> <li>Some errors that do not interfere with meaning</li> <li>Some variety</li> <li>Generally correct</li> <li>No consistent pattern of errors</li> <li>Some errors that do not interfere with meaning</li> </ul>	<ul style="list-style-type: none"> <li>Opening and closing</li> <li>Single focus</li> <li>Sense of unity and coherence</li> <li>Key ideas developed</li> <li>Logical progression of ideas</li> <li>Moderately fluent</li> <li>Attempts compositional risks</li> <li>Details appropriate and varied</li> <li>Few errors</li> <li>Variety in syntax appropriate and effective</li> <li>Few errors</li> <li>Few errors</li> </ul>	<ul style="list-style-type: none"> <li>Opening and closing</li> <li>Single, distinct focus</li> <li>Unified and coherent</li> <li>Well-developed</li> <li>Logical progression of ideas</li> <li>Fluent, cohesive</li> <li>Compositional risks successful</li> <li>Details effective, vivid, explicit, and/or pertinent</li> <li>Very few, if any, errors</li> <li>Precision and/or sophistication</li> <li>Very few, if any, errors</li> <li>Very few, if any, errors</li> </ul>
<b>Usage</b>						
<b>Sentence Construction</b>						
<b>Mechanics</b>						

NON-SCORABLE RESPONSES	Content/Organization	Usage	Sentence Construction	Mechanics
<p><b>NR = No Response</b> Student wrote too little to allow a reliable judgment, of his/her writing.</p> <p><b>OT = Off Topic/ Off Task</b> Student did not write on the assigned topic/task, or the student attempted to copy the prompt.</p> <p><b>NE = Not English</b> Student wrote in a language other than English.</p> <p><b>WF = Wrong Format</b> Student refused to write on the topic, or the writing task folder was blank.</p>	<ul style="list-style-type: none"> <li>Communicates intended message to intended audience</li> <li>Relates to topic</li> <li>Opening and closing</li> <li>Focused</li> <li>Logical progression of ideas</li> <li>Transitions</li> <li>Appropriate details and information</li> </ul>	<ul style="list-style-type: none"> <li>Tense formation</li> <li>Subject-verb agreement</li> <li>Pronouns</li> <li>usage/agreement</li> <li>Word choice/meaning</li> <li>Proper Modifiers</li> </ul>	<ul style="list-style-type: none"> <li>Variety of type, structure, and length</li> <li>Correct construction</li> </ul>	<ul style="list-style-type: none"> <li>Spelling</li> <li>Capitalization</li> <li>Punctuation</li> </ul>

Note: All unscorable responses, (NSRs), with the exception of NR, must be coded by the Scoring Director.

© New Jersey Department of Education

## OPEN-ENDED SCORING RUBRIC

Points	Criteria
<b>4</b>	A 4-point response clearly demonstrates understanding of the task, completes all requirements, and provides an insightful explanation/opinion that links to or extends aspects of the text.
<b>3</b>	A 3-point response demonstrates an understanding of the task, completes all requirements, and provides some explanation/opinion using situations or ideas from the text as support.
<b>2</b>	A 2-point response may address all of the requirements, but demonstrates a partial understanding of the task, and uses text incorrectly or with limited success resulting in an inconsistent or flawed explanation.
<b>1</b>	A 1-point response demonstrates minimal understanding of the task, does not complete the requirements, and provides only a vague reference to or no use of the text.
<b>0</b>	A 0-point response is irrelevant or off-topic.

© New Jersey Department of Education

## Holistic Scoring Guide for Mathematics Open-Ended (OE) Items (Generic Rubric)

### 3-Point Response

The response shows complete understanding of the problem's essential mathematical concepts. The student executes procedures completely and gives relevant responses to all parts of the task. The response contains few minor errors, if any. The response contains a clear, effective explanation detailing how the problem was solved so that the reader does not need to infer how and why decisions were made.

### 2-Point Response

The response shows nearly complete understanding of the problem's essential mathematical concepts. The student executes nearly all procedures and gives relevant responses to most parts of the task. The response may have minor errors. The explanation detailing how the problem was solved may not be clear, causing the reader to make some inferences.

### 1-Point Response

The response shows limited understanding of the problem's essential mathematical concepts. The response and procedures may be incomplete and/or may contain major errors. An incomplete explanation of how the problem was solved may contribute to questions as to how and why decisions were made.

### 0-Point Response

The response shows insufficient understanding of the problem's essential mathematical concepts. The procedures, if any, contain major errors. There may be no explanation of the solution or the reader may not be able to understand the explanation. The reader may not be able to understand how and why decisions were made.

The above generic rubric is used as a guide to develop specific scoring guides or rubrics for each of the Open-Ended (OE) items which appear on the New Jersey statewide assessments in Mathematics. These scoring rubrics provide the criteria for evaluating and scoring student performance and are developed by a committee of mathematicians and teachers. Rubrics ensure that there is consistency, fairness, and accuracy in scoring open-ended questions.

## HOLISTIC SCORING GUIDE (GENERIC RUBRIC) FOR SCIENCE OPEN-ENDED QUESTIONS

The zero-to-three-point generic scoring rubric below was created to help readers score open-ended responses consistently. In scoring, the reader should accept the use of appropriate diagrams, charts, formulas, and/or symbols which are part of a correct answer even when the question does not specifically request their use.

- |   |
|---|
| <p><b>3-Point Response:</b> Student response is reasonably complete, clear, and satisfactory.</p> <p><b>2-Point Response:</b> Student response has minor omissions and/or some incorrect information.</p> <p><b>1-Point Response:</b> Student response includes some correct information, but most information included in the response is either incorrect or not relevant.</p> <p><b>0-Point Response:</b> Student attempts the task but the response is incorrect, not relevant, or inappropriate.</p> |
|---|

The above generic rubric is used as a guide to develop item specific scoring guides or rubrics for each of the open-ended (OE) questions that appear on the New Jersey statewide assessments in Science. These scoring rubrics provide the criteria for evaluating and scoring student performance and are developed by a committee of scientists and teachers. Rubrics ensure that there is consistency, fairness, and accuracy in scoring open-ended questions.

Table 1

**Score Calculation Chart**Used for Means ( $\bar{x}$ )(Used for 3<sup>rd</sup> reading equal to or adjacent and all valid)

Absolute Difference ( $ 1^{\text{st}} - 2^{\text{nd}} $ )	Additional Conditions*	Additional Conditions*	Score Calculation*
0 <i>No 3<sup>rd</sup> Reading</i>	Both readings are valid	--	$(1^{\text{st}} + 2^{\text{nd}})/2$
1 <i>No 3<sup>rd</sup> Reading</i>	Both readings are valid	--	$(1^{\text{st}} + 2^{\text{nd}})/2$
2	$1^{\text{st}} < 3^{\text{rd}} < 2^{\text{nd}}$ or $2^{\text{nd}} < 3^{\text{rd}} < 1^{\text{st}}$	--	$(1^{\text{st}} + 2^{\text{nd}})/2$
2	$3^{\text{rd}} < ((1^{\text{st}} + 2^{\text{nd}})/2)$	$1^{\text{st}} < 2^{\text{nd}}$	$(1^{\text{st}} + 3^{\text{rd}})/2$
		$2^{\text{nd}} < 1^{\text{st}}$	$(2^{\text{nd}} + 3^{\text{rd}})/2$
	$3^{\text{rd}} > ((1^{\text{st}} + 2^{\text{nd}})/2)$	$1^{\text{st}} < 2^{\text{nd}}$	$(2^{\text{nd}} + 3^{\text{rd}})/2$
		$2^{\text{nd}} < 1^{\text{st}}$	$(1^{\text{st}} + 3^{\text{rd}})/2$
3	$3^{\text{rd}} = 1^{\text{st}}$ or $(3^{\text{rd}} \pm 1) = 1^{\text{st}}$	--	$(1^{\text{st}} + 3^{\text{rd}})/2$
	$3^{\text{rd}} = 2^{\text{nd}}$ or $(3^{\text{rd}} \pm 1) = 2^{\text{nd}}$	--	$(2^{\text{nd}} + 3^{\text{rd}})/2$
4 and 5	$3^{\text{rd}} = 1^{\text{st}}$ or $(3^{\text{rd}} \pm 1) = 1^{\text{st}}$	--	$(1^{\text{st}} + 3^{\text{rd}})/2$
	$3^{\text{rd}} = 2^{\text{nd}}$ or $(3^{\text{rd}} \pm 1) = 2^{\text{nd}}$	--	$(2^{\text{nd}} + 3^{\text{rd}})/2$

If both readings are invalid and equal, the score is 0.

Used for Means ( $\chi$ )  
 (Used for 3<sup>rd</sup> Reading NOT equal to or adjacent but all valid)

Condition	Score Calculation
1 <sup>st</sup> < 3 <sup>rd</sup> < 2 <sup>nd</sup> or 2 <sup>nd</sup> < 3 <sup>rd</sup> < 1 <sup>st</sup>	Use 3 <sup>rd</sup> reading
1 <sup>st</sup> < 2 <sup>nd</sup> < 3 <sup>rd</sup> or 3 <sup>rd</sup> < 2 <sup>nd</sup> < 1 <sup>st</sup>	$(2^{\text{nd}} + 3^{\text{rd}})/2$
2 <sup>nd</sup> < 1 <sup>st</sup> < 3 <sup>rd</sup> or 3 <sup>rd</sup> < 1 <sup>st</sup> < 2 <sup>nd</sup>	$(1^{\text{st}} + 3^{\text{rd}})/2$

**Table 2A**  
 Used for Means ( $\chi$ )  
 (Used for 1<sup>st</sup> or 2<sup>nd</sup> reading invalid and 3<sup>rd</sup> Reading valid)

Condition	Additional Condition	Score Calculation
1 <sup>st</sup> Reading Invalid 2 <sup>nd</sup> Reading Valid	Absolute difference between 2 <sup>nd</sup> Reading and 3 <sup>rd</sup> reading is 0 or 1	$(2^{\text{nd}} + 3^{\text{rd}}) / 2$
	Absolute difference between 2 <sup>nd</sup> Reading and 3 <sup>rd</sup> reading is greater than 1	Use 3 <sup>rd</sup> Reading
1 <sup>st</sup> Reading Valid 2 <sup>nd</sup> Reading Invalid	Absolute difference between 1 <sup>st</sup> Reading and 3 <sup>rd</sup> reading is 0 or 1	$(1^{\text{st}} + 3^{\text{rd}}) / 2$
	Absolute difference between 1 <sup>st</sup> Reading and 3 <sup>rd</sup> reading is greater than 1	Use 3 <sup>rd</sup> Reading
Both 1 <sup>st</sup> and 2 <sup>nd</sup> Readings are invalid		Use 3 <sup>rd</sup> Reading

If the 3<sup>rd</sup> Reading is invalid, use the 3<sup>rd</sup> reading score.

**Table 3**  
**Score Calculation Chart**

Used for Sum ( $\Sigma$ )  
(Used for 3<sup>rd</sup> Reading equal to or adjacent and all valid)

Absolute Difference ( 1 <sup>st</sup> - 2 <sup>nd</sup>  )	Additional Conditions*	Additional Conditions*	Score Calculation*
0 <i>No 3<sup>rd</sup> Reading</i>	--	--	(1 <sup>st</sup> + 2 <sup>nd</sup> )
1 <i>No 3<sup>rd</sup> Reading</i>	--	--	(1 <sup>st</sup> + 2 <sup>nd</sup> )
2 – 5	Equal to or Adjacent	--	((1 <sup>st</sup> + 2 <sup>nd</sup> + 3 <sup>rd</sup> ) *2) /3

**Table 4****Additional Score Calculations**Used for Sum ( $\Sigma$ )(Used for 3<sup>rd</sup> Reading NOT equal to or adjacent but all valid)

Condition	Score Calculation
NOT Equal to or Adjacent	$((1^{\text{st}} + 2^{\text{nd}} + 3^{\text{rd}}) * 2) / 3$

If the 3<sup>rd</sup> Reading is invalid, use the 3<sup>rd</sup> reading score.**Table 4A**Used for Sum ( $\Sigma$ )(Used for 1<sup>st</sup> or 2<sup>nd</sup> reading invalid and 3<sup>rd</sup> Reading valid)

Condition	Additional Condition	Score Calculation
1 <sup>st</sup> Reading Invalid 2 <sup>nd</sup> Reading Valid	Absolute difference between 2 <sup>nd</sup> Reading and 3 <sup>rd</sup> reading is 0 or 1	$(2^{\text{nd}} + 3^{\text{rd}})$
	Absolute difference between 2 <sup>nd</sup> Reading and 3 <sup>rd</sup> reading is greater than 1	$(3^{\text{rd}} * 2)$
1 <sup>st</sup> Reading Valid 2 <sup>nd</sup> Reading Invalid	Absolute difference between 1 <sup>st</sup> Reading and 3 <sup>rd</sup> reading is 0 or 1	$(1^{\text{st}} + 3^{\text{rd}})$
	Absolute difference between 1 <sup>st</sup> Reading and 3 <sup>rd</sup> reading is greater than 1	$(3^{\text{rd}} * 2)$
Both 1 <sup>st</sup> and 2 <sup>nd</sup> Readings are invalid		Use 3 <sup>rd</sup> Reading

If the 3<sup>rd</sup> Reading is invalid, use the 3<sup>rd</sup> reading score.

**Table 5**  
**When to Use the Mean vs. Sum Scoring Rules**

Subject	Valid scores	Grade 8
Reading OE	0-4 *	Mean
Writing – Picture	Grade 8 1-6 **	Mean
Writing – Persuasive	1-6 **	Sum
Math OE	0-3 *	Mean
Science OE	0-3 *	Mean

Designation Codes:

- \* = 7 = NR, for No Response  
(blank, fragmented, refusing or unable to write on topic, copy of item)  
8 = OT, for Off Topic  
9 = NE, for Not English
- \*\* = 0 = NR, for No Response  
7 = WF, for Wrong Format  
8 = OT, for Off Topic  
9 = NE, for Not English
- \*\*\* = 7 = NR, No Response

There are three types of situations that will require a third reading:

1. First and second reading are valid scores and not equal or adjacent.
2. One reading is a valid score and the other reading is not a valid score.
3. Both readings are not a valid score and are not equal.



# ***Appendix B***

## ***Cycle II Test Results***

## 2007 Grade Eight Proficiency Assessment Executive Summary

The 2007 New Jersey Grade Eight Proficiency Assessment (GEPA) consisted of three content areas: Language Arts Literacy, Mathematics, and Science. The GEPA is used as a primary indicator for identifying those students who may need instructional intervention in the three content areas. It is designed to give an indication of the progress students are making in mastering the skills they will need to pass the High School Proficiency Assessment (HSPA).

The GEPA Language Arts Literacy, Mathematics, and Science scores are reported as scale scores with a range of 100 to 300. Please note that 100 and 300 are a theoretical floor and ceiling, which may not actually be observed. The score ranges are as follows:

<i>Advanced Proficient</i>	<i>250-300</i>
<i>Proficient</i>	<i>200-249</i>
<i>Partially Proficient</i>	<i>100-199</i>

Students who scored Partially Proficient are considered to be below the state minimum level of proficiency. These students may need additional instructional support such as individual or programmatic intervention. It is important that districts consider multiple measures with all students before making final decisions about students' instructional placement.

The GEPA was administered between March 12 and March 15, 2007. Of the 108,474 students enrolled, 105,865 students received valid scale scores in Language Arts Literacy, 106,980 students received valid scale scores in Mathematics, and 106,913 students received valid scale scores in Science.

For the total group of students, 26.4% scored Partially Proficient, 62.4% Proficient, and 11.3% Advanced Proficient in Language Arts Literacy. In Mathematics, 31.6% scored Partially Proficient, 45.9% Proficient, and 22.5% Advanced Proficient. In Science, 21.1% scored Partially Proficient, 54.3% Proficient, and 24.6% Advanced Proficient. The mean scale score was 214.9 in Language Arts Literacy, 215.5 in Mathematics, and 223.3 in Science.

This executive summary includes four tables summarizing statewide performance by demographic groups. Tables 1–3 present the performance in Language Arts Literacy, Mathematics, and Science, respectively. Table 4 presents the performance for the state, Special Needs districts, and Non-Special Needs districts.

The performance data include only students with valid scale scores. Students whose answer folders were voided are excluded. Students may receive a scale score in one content area, but not in others.

In 2007, results for the General Education group are being reported in the state summary. The General Education group included students with answer folders not coded as special education or limited English proficient. In 2006, the General Education group was not reported in the state summary.

Previously, a major change for the 2006 State Summary was the reporting of the Limited English Proficient (LEP). LEP was reported as LEP (Current and Former) with two subcategories: LEP Current and LEP Former.

Student performance is summarized by total students, education program, and student demographic subgroups: Total, General Education, Special Education (SE), Limited English Proficient (LEP), Gender, Ethnicity, Economic status (disadvantaged vs. not disadvantaged), and Migrant status.

For each demographic group, the number of students participating, the percent of students in each proficiency level, and the mean scale score are reported for each content area. The percentages of students for the three proficiency levels may not total to one hundred due to rounding. The percentage of students in Proficient or Advanced Proficient is calculated by subtracting the percentage of students in Partially Proficient from one hundred.

Demographic information originates from the data collected on the students' answer folders. School district personnel were given an opportunity to review the demographic information they provided on the answer folders and correct any errors prior to reporting.

This executive summary includes information from the state level Performance by Demographic Groups Report from Cycle II reporting. The complete state summary data file with District Factor Groups and longitudinal data is available at <http://www.state.nj.us/education/schools/achievement/>.

### **Reporting Rules for State Summary Data File**

The state summary data files contain the same type of information shown on the Statewide Performance by Demographic Groups Report for schools and districts included with the Cycle II reporting. In order to safeguard student confidentiality, certain information is suppressed in the files according to the following reporting rules:

- Data are not reported if the number of students with valid scale scores for a particular group is fewer than 11.
- Data are not reported where demographic groups are mutually exclusive (e.g., gender) and there are one or two students with a valid scale score in one of the groups (e.g., male).
- Data are not reported if it is possible to identify individual student performance.

## Highlights from the 2007 GEPA Performance Results

The percentages of students scoring at Proficient or Advanced Proficient by content areas are described below:

### Total

- Language Arts Literacy – 73.6% of the students scored Proficient or Advanced Proficient and 11.3% of the students scored Advanced Proficient.
- Mathematics – 68.4% of the students scored Proficient or Advanced Proficient and 22.5% of the students scored Advanced Proficient.
- Science – 78.9% of the students scored Proficient or Advanced Proficient and 24.6% of the students scored Advanced Proficient.

### General Education

- Language Arts Literacy – 83.1% of the students scored Proficient or Advanced Proficient and 13.5% of the students scored Advanced Proficient.
- Mathematics – 77.6% of the students scored Proficient or Advanced Proficient and 26.9% of the students scored Advanced Proficient.
- Science – 86.2% of the students scored Proficient or Advanced Proficient and 29.1% of the students scored Advanced Proficient.

### Special Education

- Language Arts Literacy – 32.9% of the students scored Proficient or Advanced Proficient and 1.2% of the students scored Advanced Proficient.
- Mathematics – 28.8% of the students scored Proficient or Advanced Proficient and 3.6% of the students scored Advanced Proficient.
- Science – 50.5% of the students scored Proficient or Advanced Proficient and 6.1% of the students scored Advanced Proficient.

### Limited English Proficient (LEP)

- Language Arts Literacy – 27.4% of the LEP Current and Former students scored Proficient or Advanced Proficient and 0.9% of the LEP Current and Former students scored Advanced Proficient. About 18.7% of the Current LEP students scored Proficient or Advanced Proficient and 0.4% of the Current LEP students scored Advanced Proficient and about 47.8% of the Former LEP students scored Proficient or Advanced Proficient and 2.0% of the Former LEP students scored Advanced Proficient.
- Mathematics – 31.0% of the LEP Current and Former students scored Proficient or Advanced Proficient and 5.7% of the group scored Advanced Proficient. Of the Current LEP students, 25.6% scored Proficient or Advanced Proficient and 5.0% scored Advanced Proficient. Of the Former LEP students, 46.6% scored Proficient or Advanced Proficient and 7.8% scored Advanced Proficient.
- Science – 33.9% of the LEP Current and Former students scored Proficient or Advanced Proficient and 2.0% of the group scored Advanced Proficient. Of the Current LEP students, 27.6% scored Proficient or Advanced Proficient and 1.4% scored Advanced Proficient. Of the Former LEP students, 52.3% scored Proficient or Advanced Proficient and 4.0% scored Advanced Proficient.

**TABLE 1**  
**2007 Grade Eight Proficiency Assessment**  
**Statewide Performance**  
**Language Arts Literacy**

	Number of Students Enrolled	Number of APA Students	Number Not Present	Number of Voids	Number of Students with Valid Scale Scores	% Partially Proficient	% Proficient	% Advanced Proficient	Scale Score Mean
<b>Total Students</b>	108,474	680	667	1,262	105,865	26.4	62.4	11.3	214.9
<b>General Education</b>	87,396	0	311	466	86,619	16.9	69.5	13.5	221.8
<b>Special Education</b>	18,197	680	229	317	16,971	67.1	31.7	1.2	185.1
<b>LEP (current and former)</b>	4,020	6	132	486	3,396	72.6	26.5	0.9	180.6
<b>Current LEP</b>	2,999	3	129	483	2,384	81.3	18.3	0.4	173.4
<b>Former LEP</b>	1,021	3	3	3	1,012	52.2	45.8	2.0	197.6
<b>Gender</b>									
Female	52,599	238	274	508	51,579	19.3	65.1	15.6	221.1
Male	55,803	436	385	747	54,235	33.0	59.8	7.2	209.0
<b>Ethnicity</b>									
White	61,596	360	188	309	60,739	15.9	69.8	14.3	222.5
Black	19,277	140	197	350	18,590	50.1	47.0	2.9	197.3
Asian	7,729	28	47	89	7,565	12.9	61.6	25.5	229.4
Pacific Islander	293	3	3	2	285	17.2	69.8	13.0	221.3
Hispanic	19,004	132	205	494	18,173	42.4	53.7	3.9	201.7
American Indian/Alaskan Native	114	2	1	2	109	30.3	64.2	5.5	210.3
Other	461	15	26	16	404	42.8	50.7	6.4	201.5
<b>Economic Status</b>									
Economically Disadvantaged	29,783	214	312	699	28,558	48.3	48.8	3.0	198.1
Non-Economically Disadvantaged	78,691	466	355	563	77,307	18.3	67.4	14.4	221.1
<b>Migrant Status</b>									
Migrant	41	0	0	3	38	63.2	36.8	0.0	188.8
Non-Migrant	108,433	680	667	1,259	105,827	26.3	62.4	11.3	214.9

**TABLE 2**  
**2007 Grade Eight Proficiency Assessment**  
**Statewide Performance**  
**Mathematics**

	Number of Students Enrolled	Number of APA Students	Number Not Present	Number of Voids	Number of Students with Valid Scale Scores	% Partially Proficient	% Proficient	% Advanced Proficient	Scale Score Mean
<b>Total Students</b>	108,474	681	661	152	106,980	31.6	45.9	22.5	215.5
<b>General Education</b>	87,396	0	385	71	86,940	22.4	50.8	26.9	222.5
<b>Special Education</b>	18,197	681	250	76	17,190	71.2	25.2	3.6	185.2
<b>LEP (current and former)</b>	4,020	5	32	8	3,975	69.0	25.3	5.7	187.5
<b>Current LEP</b>	2,999	3	27	7	2,962	74.4	20.7	5.0	183.8
<b>Former LEP</b>	1,021	2	5	1	1,013	53.4	38.8	7.8	198.4
<b>Gender</b>									
Female	52,599	245	266	39	52,049	31.7	48.0	20.2	214.6
Male	55,803	430	389	113	54,871	31.4	43.8	24.8	216.4
<b>Ethnicity</b>									
White	61,596	368	223	62	60,943	19.1	52.1	28.8	224.8
Black	19,277	137	260	57	18,823	61.7	33.3	4.9	191.8
Asian	7,729	26	12	3	7,688	12.3	39.4	48.3	238.1
Pacific Islander	293	3	2	0	288	19.4	48.6	31.9	226.5
Hispanic	19,004	129	146	29	18,700	49.5	41.1	9.4	200.1
American Indian/Alaskan Native	114	3	0	0	111	43.2	36.0	20.7	207.9
Other	461	15	18	1	427	49.9	34.2	15.9	201.8
<b>Economic Status</b>									
Economically Disadvantaged	29,783	211	311	79	29,182	55.4	37.1	7.5	196.1
Non-Economically Disadvantaged	78,691	470	350	73	77,798	22.6	49.2	28.2	222.8
<b>Migrant Status</b>									
Migrant	41	0	0	0	41	53.7	46.3	0.0	195.2
Non-Migrant	108,433	681	661	152	106,939	31.6	45.9	22.6	215.5

## Gender

- Language Arts Literacy – 80.7% of the female students and 67.0% of the male students scored Proficient or Advanced Proficient while 15.6% of the female students and 7.2% of the male students scored Advanced Proficient.
- Mathematics – 68.3% of the female students and 68.6% of the male students scored Proficient or Advanced Proficient while 20.2% of the female students and 24.8% of the male students scored Advanced Proficient.
- Science – 78.2% of the female students and 79.5% of the male students scored Proficient or Advanced Proficient while 21.2% of the female students and 27.9% of the male students scored Advanced Proficient.

## Ethnicity

- Language Arts Literacy – percentages of Proficient and Advanced Proficient ranged from 87.1% of Asian students to 49.9% of Black students while the percentages of Advanced Proficient ranged from 25.5% of Asian students to 2.9% of Black students and 3.9% of Hispanic students. (The percentages of the Proficient and Advanced Proficient scores in the other ethnic groups fell between the Asian and Black groups.)
- Mathematics – percentages of Proficient and Advanced Proficient ranged from 87.7% of Asian students to 38.3% of Black students while percentages of Advanced Proficient ranged from 48.3% of Asian students to 4.9% of Black students.
- Science – percentages of Proficient and Advanced Proficient ranged from 90.5% of Asian students and 90.4% of White students to 54.6% of Black students while percentages of Advanced Proficient ranged from 43.4% of Asian students to 5.7% of Black students.

**TABLE 3**  
**2007 Grade Eight Proficiency Assessment**  
**Statewide Performance**  
**Science**

	Number of Students Enrolled	Number of APA Students	Number Not Present	Number of Voids	Number of Students with Valid Scale Scores	% Partially Proficient	% Proficient	% Advanced Proficient	Scale Score Mean
<b>Total Students</b>	108,474	645	767	149	106,913	21.1	54.3	24.6	223.3
<b>General Education</b>	87,396	0	430	71	86,895	13.8	57.1	29.1	228.9
<b>Special Education</b>	18,197	645	306	73	17,173	49.5	44.4	6.1	200.7
<b>LEP (current and former)</b>	4,020	4	41	5	3,970	66.1	31.8	2.0	190.9
Current LEP	2,999	2	34	5	2,958	72.4	26.2	1.4	187.2
Former LEP	1,021	2	7	0	1,012	47.7	48.3	4.0	201.7
<b>Gender</b>									
Female	52,599	233	326	44	51,996	21.8	57.0	21.2	221.4
Male	55,803	406	435	105	54,857	20.5	51.7	27.9	225.1
<b>Ethnicity</b>									
White	61,596	348	250	67	60,931	9.6	57.3	33.2	232.7
Black	19,277	130	290	45	18,812	45.4	48.9	5.7	203.0
Asian	7,729	24	16	8	7,681	9.5	47.1	43.4	237.7
Pacific Islander	293	3	0	0	290	12.4	58.3	29.3	228.6
Hispanic	19,004	123	189	27	18,665	38.9	52.8	8.3	207.1
American Indian/Alaskan Native	114	2	0	0	112	21.4	58.0	20.5	220.1
Other	461	15	22	2	422	36.3	49.5	14.2	212.0
<b>Economic Status</b>									
Economically Disadvantaged	29,783	199	364	71	29,149	43.3	49.9	6.8	204.3
Non-Economically Disadvantaged	78,691	446	403	78	77,764	12.8	55.9	31.3	230.4
<b>Migrant Status</b>									
Migrant	41	0	0	0	41	51.2	41.5	7.3	203.1
Non-Migrant	108,433	645	767	149	106,872	21.1	54.3	24.6	223.3

### **Economic Status**

- Language Arts Literacy – 51.7% of Economically Disadvantaged students and 81.7% of Non-Economically Disadvantaged students scored Proficient or Advanced Proficient while 3.0% of Economically Disadvantaged students and 14.4% of Non-Economically Disadvantaged students scored Advanced Proficient.
- Mathematics – 44.6% of Economically Disadvantaged students and 77.4% of Non-Economically Disadvantaged students scored Proficient or Advanced Proficient while 7.5% of Economically Disadvantaged students and 28.2% of Non-Economically Disadvantaged students scored Advanced Proficient.
- Science – 56.7% of the Economically Disadvantaged students and 87.2% of Non-Economically Disadvantaged students scored Proficient or Advanced Proficient while 6.8% of Economically Disadvantaged students and 31.3% of Non-Economically Disadvantaged students scored Advanced Proficient.

### **Migrant**

Only 0.038% of the enrolled grade 8 students were migrant students. The percentage of Migrant students scoring at Proficient or Advanced Proficient was 36.8% for Language Arts Literacy, 46.3% for Mathematics, and 48.8% for Science. The percentage of Migrant students scoring at Advanced Proficient was 0% for Language Arts Literacy, 0% for Mathematics, and 7.3% for Science.

**TABLE 4**  
**2007 Grade Eight Proficiency Assessment - Statewide Performance**  
**Non-Special Needs and Special Needs Districts**

	<b>Number of Students with Valid Scores</b>	<b>% Partially Proficient</b>	<b>% Proficient</b>	<b>% Advanced Proficient</b>	<b>Scale Score Mean</b>
<b>LANGUAGE ARTS LITERACY</b>					
<b>STATEWIDE TOTAL</b>	<b>105,865</b>	<b>26.4</b>	<b>62.4</b>	<b>11.3</b>	<b>214.9</b>
Non-Special Needs	86,190	20.5	66.3	13.2	219.3
Special Needs	19,675	52.2	45.0	2.8	195.3
<b>GENERAL EDUCATION</b>	<b>86,619</b>	<b>16.9</b>	<b>69.5</b>	<b>13.5</b>	<b>221.8</b>
Non-Special Needs	72,141	12.3	72.2	15.5	225.2
Special Needs	14,478	40.0	56.2	3.8	204.9
<b>SPECIAL EDUCATION</b>	<b>16,971</b>	<b>67.1</b>	<b>31.7</b>	<b>1.2</b>	<b>185.1</b>
Non-Special Needs	12,983	61.3	37.2	1.5	190.3
Special Needs	3,988	86.2	13.6	0.2	168.3
<b>LEP CURRENT &amp; FORMER</b>	<b>3,396</b>	<b>72.6</b>	<b>26.5</b>	<b>0.9</b>	<b>180.6</b>
Non-Special Needs	1,548	67.7	31.3	1.0	185.1
Special Needs	1,848	76.8	22.5	0.8	176.9
<b>CURRENT LEP</b>	<b>2,384</b>	<b>81.3</b>	<b>18.3</b>	<b>0.4</b>	<b>173.4</b>
Non-Special Needs	1,113	76.6	22.7	0.6	178.3
Special Needs	1,271	85.4	14.4	0.2	169.2
<b>FORMER LEP</b>	<b>1,012</b>	<b>52.2</b>	<b>45.8</b>	<b>2.0</b>	<b>197.6</b>
Non-Special Needs	435	44.8	53.3	1.8	202.6
Special Needs	577	57.7	40.2	2.1	193.8
<b>MATHEMATICS</b>					
<b>STATEWIDE TOTAL</b>	<b>106,980</b>	<b>31.6</b>	<b>45.9</b>	<b>22.5</b>	<b>215.5</b>
Non-Special Needs	86,738	24.9	49.0	26.1	220.7
Special Needs	20,242	60.3	32.4	7.3	193.3
<b>GENERAL EDUCATION</b>	<b>86,940</b>	<b>22.4</b>	<b>50.8</b>	<b>26.9</b>	<b>222.5</b>
Non-Special Needs	72,284	16.8	52.9	30.4	226.9
Special Needs	14,656	50.1	40.3	9.6	200.8
<b>SPECIAL EDUCATION</b>	<b>17,190</b>	<b>71.2</b>	<b>25.2</b>	<b>3.6</b>	<b>185.2</b>
Non-Special Needs	13,082	65.5	30.0	4.5	189.4
Special Needs	4,108	89.2	10.2	0.6	171.7
<b>LEP CURRENT &amp; FORMER</b>	<b>3,975</b>	<b>69.0</b>	<b>25.3</b>	<b>5.7</b>	<b>187.5</b>
Non-Special Needs	1,860	61.6	30.1	8.3	193.3
Special Needs	2,115	75.6	21.0	3.4	182.5
<b>CURRENT LEP</b>	<b>2,962</b>	<b>74.4</b>	<b>20.7</b>	<b>5.0</b>	<b>183.8</b>
Non-Special Needs	1,423	66.3	26.1	7.6	190.0
Special Needs	1,539	81.9	15.6	2.5	178.1
<b>FORMER LEP</b>	<b>1,013</b>	<b>53.4</b>	<b>38.8</b>	<b>7.8</b>	<b>198.4</b>
Non-Special Needs	437	46.5	43.0	10.5	204.0
Special Needs	576	58.7	35.6	5.7	194.2
<b>SCIENCE</b>					
<b>STATEWIDE TOTAL</b>	<b>106,913</b>	<b>21.1</b>	<b>54.3</b>	<b>24.6</b>	<b>223.3</b>
Non-Special Needs	86,713	14.3	56.5	29.2	228.7
Special Needs	20,200	50.5	44.4	5.1	200.1
<b>GENERAL EDUCATION</b>	<b>86,895</b>	<b>13.8</b>	<b>57.1</b>	<b>29.1</b>	<b>228.9</b>
Non-Special Needs	72,262	8.5	58.0	33.6	233.5
Special Needs	14,633	40.4	52.9	6.7	206.1
<b>SPECIAL EDUCATION</b>	<b>17,173</b>	<b>49.5</b>	<b>44.4</b>	<b>6.1</b>	<b>200.7</b>
Non-Special Needs	13,080	41.1	51.2	7.7	205.6
Special Needs	4,093	76.4	22.8	0.8	184.9
<b>LEP CURRENT &amp; FORMER</b>	<b>3,970</b>	<b>66.1</b>	<b>31.8</b>	<b>2.0</b>	<b>190.9</b>
Non-Special Needs	1,859	59.5	37.3	3.2	195.2
Special Needs	2,111	72.0	27.0	1.0	187.2
<b>CURRENT LEP</b>	<b>2,958</b>	<b>72.4</b>	<b>26.2</b>	<b>1.4</b>	<b>187.2</b>
Non-Special Needs	1,423	65.3	32.4	2.3	191.5
Special Needs	1,535	79.1	20.4	0.5	183.3
<b>FORMER LEP</b>	<b>1,012</b>	<b>47.7</b>	<b>48.3</b>	<b>4.0</b>	<b>201.7</b>
Non-Special Needs	436	40.6	53.4	6.0	207.2
Special Needs	576	53.1	44.4	2.4	197.6



## New Jersey Statewide Testing System Grade Eight Proficiency Assessment Performance by Demographic Groups State

Test Date: March 2007  
CYCLE II

	Language Arts Literacy										Mathematics										Science									
	Students Enrolled	APA Students	Not Present	Valid Scale Scores	% Partially Proficient	% Proficient	% Advanced Proficient	Scale Score Mean	APA Students	Not Present	Valid Scale Scores	% Partially Proficient	% Proficient	% Advanced Proficient	Scale Score Mean	APA Students	Not Present	Valid Scale Scores	% Partially Proficient	% Proficient	% Advanced Proficient	Scale Score Mean								
<b>Total Students<sup>2</sup></b>	108,474	680	667	1,262	105,865	26.4	62.4	11.3	214.9	681	152	106,980	31.6	45.9	22.5	215.5	645	767	149	106,913	21.1	54.3	24.6	223.3						
General Education <sup>3</sup>	87,396	0	311	466	86,619	16.9	69.5	13.5	221.8	0	385	71	86,940	22.4	50.8	26.9	222.5	0	430	71	86,895	13.8	57.1	29.1	228.9					
Special Education	18,197	680	229	317	16,971	67.1	31.7	1.2	185.1	681	250	76	17,190	71.2	25.2	3.6	185.2	645	306	73	17,173	49.5	44.4	6.1	200.7					
Limited English Proficient <sup>4</sup>	4,020	6	132	486	3,396	72.6	26.5	0.9	180.6	5	32	8	3,975	69.0	25.3	5.7	187.5	4	41	5	3,970	66.1	31.8	2.0	190.9					
Current LEP	2,999	3	129	483	2,384	81.3	18.3	0.4	173.4	3	27	7	2,962	74.4	20.7	5.0	183.8	2	34	5	2,958	72.4	26.2	1.4	187.2					
Former LEP	1,021	3	3	1,012	52.2	45.8	2.0	197.6	2	5	1	1,013	53.4	38.8	7.8	198.4	2	7	0	1,012	47.7	48.3	4.0	201.7						
<b>Gender<sup>5</sup></b>																														
Female	52,599	238	274	508	51,579	19.3	65.1	15.6	221.1	245	266	39	52,049	31.7	48.0	20.2	214.6	233	326	44	51,996	21.8	57.0	21.2	221.4					
Male	55,803	436	385	747	54,255	33.0	59.8	7.2	209.0	430	389	113	54,871	31.4	43.8	24.8	216.4	406	435	105	54,857	20.5	51.7	27.9	225.1					
<b>Migrant Status</b>																														
Migrant	41	0	0	3	38	63.2	36.8	0.0	188.8	0	0	0	41	53.7	46.3	0.0	195.2	0	0	0	41	51.2	41.5	7.3	203.1					
Non-Migrant	108,433	680	667	1,259	105,827	26.3	62.4	11.3	214.9	681	152	106,939	31.6	45.9	22.6	215.5	645	767	149	106,872	21.1	54.3	24.6	223.3						
<b>Ethnicity</b>																														
White	61,596	360	188	309	60,739	15.9	69.8	14.3	222.5	368	223	62	60,943	19.1	52.1	28.8	224.8	348	250	67	60,931	9.6	57.3	33.2	232.7					
Black	19,277	140	197	350	18,590	50.1	47.0	2.9	197.3	137	260	57	18,823	61.7	33.3	4.9	191.8	130	290	45	18,812	45.4	48.9	5.7	203.0					
Asian	7,729	28	47	89	7,565	12.9	61.6	25.5	229.4	26	12	3	7,688	12.3	39.4	48.3	238.1	24	16	8	7,681	9.5	47.1	43.4	237.7					
Pacific Islander	293	3	3	2	285	17.2	69.8	13.0	221.3	3	2	0	288	19.4	48.6	31.9	226.5	3	0	0	290	12.4	58.3	29.3	228.6					
Hispanic	19,004	132	205	494	18,173	42.4	53.7	3.9	201.7	129	146	29	18,700	49.5	41.1	9.4	200.1	123	189	27	18,665	38.9	52.8	8.3	207.1					
Amer Indian/AK Native	114	2	1	2	109	30.3	64.2	5.5	210.3	3	0	0	111	43.2	36.0	20.7	207.9	2	0	0	112	21.4	58.0	20.5	220.1					
Other <sup>6</sup>	461	15	26	16	404	42.8	50.7	6.4	201.5	15	18	1	427	49.9	34.2	15.9	201.8	15	22	2	422	36.3	49.5	14.2	212.0					
<b>Economic Status</b>																														
Econ Disadvantaged	29,783	214	312	699	28,558	48.3	48.8	3.0	198.1	211	311	79	29,182	55.4	37.1	7.5	196.1	199	364	71	29,149	43.3	49.9	6.8	204.3					
Non-Econ Disadvantaged	78,691	466	355	563	77,507	18.3	67.4	14.4	221.1	470	350	73	77,798	22.6	49.2	28.2	222.8	446	403	78	77,764	12.8	55.9	31.3	230.4					

1 Includes students coded as LEP exempt (LAL only).  
 2 Students appear in each applicable category, but are included in Total Students only once.  
 3 Includes students coded as former LEP.  
 4 Includes students coded LEP or former LEP.  
 5 Excludes students who did not have gender coded.  
 6 Includes students who did not have an ethnicity coded and students who had more than one ethnicity coded.



**New Jersey Statewide Testing System  
Grade Eight Proficiency Assessment  
Performance by Demographic Groups  
DFG A**

Test Date: March 2007  
CYCLE II

	Language Arts Literacy										Mathematics										Science									
	Students Enrolled	APA Students	Not Present	Voids <sup>1</sup>	Valid Scale Scores	% Partially Proficient	% Proficient	Advanced Proficient	Scale Score Mean	APA Students	Not Present	Voids	Valid Scale Scores	% Partially Proficient	% Proficient	Advanced Proficient	Scale Score Mean	APA Students	Not Present	Voids	Valid Scale Scores	% Partially Proficient	% Proficient	Advanced Proficient	Scale Score Mean					
<b>Total Students<sup>2</sup></b>	17,401	102	264	509	16,526	53.1	44.2	2.7	194.6	99	245	45	17,012	61.7	31.8	6.5	192.2	96	294	58	16,973	51.8	43.8	4.5	199.2					
General Education <sup>3</sup>	12,497	0	132	193	12,172	41.2	55.2	3.6	204.0	0	138	22	12,337	52.1	39.4	8.6	199.4	0	160	21	12,316	42.0	52.1	5.9	204.8					
Special Education	3,631	102	81	123	3,325	86.4	13.5	0.2	168.2	99	92	20	3,420	89.2	10.3	0.5	171.6	96	116	15	3,404	76.9	22.4	0.7	184.5					
Limited English Proficient <sup>4</sup>	1,901	2	55	195	1,649	76.3	22.9	0.8	177.2	2	21	5	1,673	75.8	21.3	2.9	182.2	2	26	2	1,671	71.9	27.1	1.0	187.3					
Current LEP	1,326	1	52	193	1,080	86.0	13.8	0.2	168.4	1	16	4	1,305	82.6	15.5	1.9	177.4	1	20	2	1,303	79.6	19.8	0.6	183.0					
Former LEP	575	1	3	2	569	57.8	40.1	2.1	193.8	1	5	1	568	60.2	34.7	5.1	193.2	1	6	0	568	54.2	43.8	1.9	197.1					
<b>Gender<sup>5</sup></b>																														
Female	8,493	34	109	197	8,153	44.0	51.9	4.0	201.3	31	103	14	8,545	60.8	33.2	5.9	193.1	31	128	16	8,318	52.5	44.1	3.4	198.6					
Male	8,889	67	153	309	8,360	61.9	36.8	1.3	188.1	67	141	31	8,650	62.6	30.5	7.0	191.4	64	165	22	8,638	51.0	43.5	5.5	199.8					
<b>Migrant Status</b>																														
Migrant	14	0	0	0	14	92.9	7.1	0.0	166.7	0	0	0	14	85.7	14.3	0.0	180.1	0	0	0	14	71.4	28.6	0.0	184.7					
Non-Migrant	17,387	102	264	509	16,512	53.1	44.3	2.7	194.6	99	245	45	16,998	61.7	31.8	6.5	192.2	96	294	58	16,959	51.7	43.8	4.5	199.2					
<b>Ethnicity</b>																														
White	1,751	5	10	25	1,711	34.1	60.3	5.6	207.5	6	13	3	1,729	36.3	48.6	15.1	209.6	6	17	3	1,725	26.3	59.9	13.8	215.5					
Black	7,352	48	101	205	6,998	60.9	37.8	1.3	189.7	46	134	28	7,144	72.8	24.7	2.5	184.5	43	154	23	7,132	60.3	37.8	1.9	194.1					
Asian	267	2	5	7	253	32.8	58.1	9.1	211.3	2	1	1	263	34.2	43.7	22.1	214.3	2	1	1	263	33.5	49.0	17.5	213.8					
Pacific Islander	23	0	1	1	21	76.2	23.8	0.0	176.6	0	2	0	21	76.2	19.0	4.8	182.3	0	0	0	23	82.6	17.4	0.0	177.9					
Hispanic	7,685	44	133	266	7,442	50.6	46.4	3.0	195.9	42	88	13	7,742	57.9	34.4	7.6	194.8	42	116	10	7,717	49.9	45.7	4.3	199.9					
Amer Indian/AK Native	22	0	0	0	22	40.9	50.0	9.1	207.0	0	0	0	22	54.5	36.4	9.1	195.2	0	0	0	22	36.4	59.1	4.5	204.2					
Other <sup>6</sup>	101	3	14	5	79	72.2	25.3	2.5	181.2	3	7	0	91	83.5	13.2	3.3	177.4	3	6	1	91	72.5	25.3	2.2	188.3					
<b>Economic Status</b>																														
Econ Disadvantaged	12,691	65	170	360	12,096	54.9	42.6	2.5	193.4	63	158	35	12,435	63.3	30.8	5.8	191.1	62	197	27	12,405	53.7	42.6	3.7	197.9					
Non-Econ Disadvantaged	4,710	37	94	149	4,430	48.1	48.7	3.3	197.8	36	87	10	4,577	57.4	34.4	8.2	195.1	34	97	11	4,568	46.5	47.0	6.5	202.6					

<sup>1</sup> Includes students coded as LEP exempt (LAL only).  
<sup>2</sup> Students appear in each applicable category, but are included in Total Students only once.  
<sup>3</sup> Includes students coded as former LEP.  
<sup>4</sup> Includes students coded LEP or former LEP.  
<sup>5</sup> Excludes students who did not have gender coded.  
<sup>6</sup> Includes students who did not have an ethnicity coded and students who had more than one ethnicity coded.



**New Jersey Statewide Testing System  
Grade Eight Proficiency Assessment  
Performance by Demographic Groups  
DFG B**

Test Date: March 2007  
CYCLE II

	Language Arts Literacy										Mathematics										Science									
	Students Enrolled	APA Students	Not Present	Voide <sup>1</sup>	Valid Scale Scores	% Proficient	Advanced Proficient	Scale Score Mean	APA Students	Not Present	Voide	Valid Scale Scores	% Partially Proficient	% Proficient	Advanced Proficient	Scale Score Mean	APA Students	Not Present	Voide	Valid Scale Scores	% Partially Proficient	% Proficient	Advanced Proficient	Scale Score Mean						
<b>Total Students<sup>2</sup></b>	11,498	108	77	190	11,123	38.6	56.8	4.6 204.7	108	70	20	11,300	42.2	45.1	12.7 205.2	103	79	23	11,293	31.3	56.4	12.4 212.5								
General Education <sup>3</sup>	8,613	0	29	50	8,534	26.3	67.8	6.0 213.5	0	35	3	8,575	30.7	53.3	16.0 213.2	0	39	6	8,568	21.2	63.0	15.8 219.1								
Special Education	2,337	108	34	54	2,141	78.5	21.3	0.2 176.5	108	34	17	2,178	79.0	19.2	1.8 179.2	103	36	17	2,181	60.4	37.6	2.0 193.4								
Limited English Proficient <sup>4</sup>	681	0	14	88	579	78.4	21.4	0.2 176.9	0	1	0	680	72.2	22.6	5.1 185.4	0	5	0	676	68.3	30.5	1.2 189.3								
Current LEP	568	0	14	88	466	83.9	15.9	0.2 171.9	0	1	0	567	76.4	19.4	4.2 182.4	0	4	0	564	73.0	26.6	0.4 186.3								
Former LEP	113	0	0	0	113	55.8	44.2	0.0 197.5	0	0	0	113	51.3	38.9	9.7 200.8	0	1	0	112	44.6	50.0	5.4 204.6								
<b>Gender<sup>5</sup></b>																														
Female	5,531	35	26	74	5,396	30.1	63.2	6.7 210.9	36	21	3	5,471	41.9	46.7	11.4 204.9	33	28	8	5,462	32.0	58.0	10.0 211.0								
Male	5,961	71	49	115	5,726	46.6	50.7	2.6 198.8	70	47	17	5,827	42.4	43.6	14.1 205.5	68	50	15	5,828	30.5	54.8	14.6 214.0								
<b>Migrant Status</b>																														
Migrant	7	0	0	1	6	16.7	83.3	0.0 202.3	0	0	0	7	14.3	85.7	0.0 205.1	0	0	0	7	42.9	42.9	14.3 211.9								
Non-Migrant	11,491	108	77	189	11,117	38.6	56.7	4.6 204.7	108	70	20	11,293	42.2	45.1	12.8 205.2	103	79	23	11,286	31.3	56.4	12.4 212.5								
<b>Ethnicity</b>																														
White	4,652	50	20	50	4,532	26.8	66.3	6.9 212.5	49	22	5	4,576	29.3	52.3	18.4 214.6	48	24	7	4,573	18.0	61.9	20.1 222.3								
Black	2,785	30	25	42	2,688	56.5	42.1	1.4 193.2	31	33	11	2,710	63.7	32.6	3.7 189.9	28	31	8	2,718	47.9	48.3	3.7 200.7								
Asian	577	1	4	12	560	28.6	60.4	11.1 214.9	1	0	0	576	23.3	48.8	28.0 222.6	1	0	2	574	22.0	55.9	22.1 222.2								
Pacific Islander	36	0	1	0	37	10.8	73.0	16.2 225.9	0	0	0	38	15.8	63.2	21.1 223.4	0	0	0	38	5.3	73.7	21.1 227.1								
Hispanic	3,360	23	24	83	3,250	42.3	54.7	2.9 201.5	23	13	4	3,340	46.1	44.4	9.5 201.5	22	22	6	3,330	37.8	55.2	7.0 206.9								
Amer Indian/AK Native	24	0	1	1	22	22.7	68.2	9.1 215.5	0	0	0	24	29.2	37.5	33.3 214.3	0	0	0	24	25.0	50.0	25.0 221.0								
Other <sup>6</sup>	42	4	2	2	34	50.0	47.1	2.9 197.4	4	2	0	36	38.9	55.6	5.6 198.2	4	2	0	36	33.3	58.3	8.3 209.6								
<b>Economic Status</b>																														
Econ Disadvantaged	5,639	53	43	131	5,412	47.8	49.6	2.6 198.4	53	40	14	5,532	52.2	39.6	8.2 197.8	49	46	15	5,529	40.7	52.7	6.6 205.3								
Non-Econ Disadvantaged	5,859	55	34	59	5,711	29.9	63.6	6.5 210.7	55	30	6	5,768	32.6	50.3	17.1 212.3	54	33	8	5,764	22.2	59.9	17.9 219.4								

1 Includes students coded as LEP exempt (LAL only).  
 2 Students appear in each applicable category, but are included in Total Students only once.  
 3 Includes students coded as LEP or former LEP.  
 4 Excludes students who did not have gender coded.  
 5 Excludes students who did not have an ethnicity coded and students who had more than one ethnicity coded.  
 6 Includes students who did not have an ethnicity coded and students who had more than one ethnicity coded.



**New Jersey Statewide Testing System  
Grade Eight Proficiency Assessment  
Performance by Demographic Groups  
DFG CD**

Test Date: March 2007  
CYCLE II

	Language Arts Literacy					Mathematics					Science															
	Students Enrolled	APA Students	Not Present	Valid Scale Scores	% Partially Proficient	% Proficient	Advanced Proficient	Scale Score Mean	APA Students	Not Present	Valid Scale Scores	% Partially Proficient	% Proficient	Advanced Proficient	Scale Score Mean											
<b>Total Students<sup>2</sup></b>	10,391	90	50	105	10,146	30.2	65.7	6.1	210.5	88	64	17	10,222	35.0	48.5	16.5	211.0	82	71	11	10,227	22.4	59.7	17.9	219.7	
General Education <sup>3</sup>	8,542	0	29	26	8,287	19.8	72.8	7.4	217.2	0	43	6	8,293	24.9	55.2	19.9	218.0	0	43	5	8,294	14.5	64.2	21.3	225.3	
Special Education	1,821	90	12	32	1,687	75.8	23.9	0.4	181.6	88	18	9	1,706	78.9	19.6	1.5	180.5	82	25	3	1,711	53.8	43.0	3.2	197.2	
Limited English Proficient <sup>4</sup>	346	1	9	49	287	71.1	28.6	0.3	183.8	1	3	2	340	67.1	28.2	4.7	186.5	1	3	3	339	65.5	32.4	2.1	191.9	
Current LEP	234	1	9	48	176	85.8	14.2	0.0	173.2	1	3	2	228	75.0	21.1	3.9	182.5	1	3	3	227	75.8	22.5	1.8	186.3	
Former LEP	112	0	0	1	111	47.7	51.4	0.9	200.7	0	0	0	112	50.9	42.9	6.3	200.6	0	0	0	112	44.6	52.7	2.7	203.3	
<b>Gender<sup>5</sup></b>																										
Female	4,972	34	19	38	4,881	20.8	70.1	9.1	217.1	34	26	3	4,909	34.5	50.9	14.6	210.8	32	30	2	4,908	22.9	61.9	15.3	218.3	
Male	5,414	55	31	66	5,262	39.0	57.7	3.3	204.4	53	58	14	5,309	35.4	46.3	18.3	211.2	49	41	9	5,315	21.9	57.8	20.3	221.1	
<b>Migrant Status</b>																										
Migrant	2	0	0	0	2	100.0	0.0	0.0	197.0	0	0	0	2	100.0	0.0	0.0	173.0	0	0	0	2	100.0	0.0	0.0	195.0	
Non-Migrant	10,389	90	50	105	10,144	30.2	65.7	6.1	210.5	88	64	17	10,220	35.0	48.6	16.5	211.0	82	71	11	10,225	22.4	59.7	17.9	219.7	
<b>Ethnicity</b>																										
White	5,542	38	17	29	5,458	23.7	68.6	7.7	215.1	37	26	10	5,469	25.7	53.5	20.8	217.8	34	29	6	5,473	14.5	61.3	24.2	226.4	
Black	2,087	23	18	26	2,020	43.1	53.8	3.2	201.6	23	26	3	2,035	53.7	40.3	6.0	196.4	22	28	1	2,036	35.8	57.5	6.7	207.9	
Asian	581	4	4	8	565	21.4	65.8	12.7	218.8	4	1	0	576	18.1	46.0	35.9	228.4	4	2	0	575	14.6	57.6	27.8	227.9	
Pacific Islander	27	0	0	0	27	11.1	81.5	7.4	221.8	0	0	0	27	3.7	70.4	25.9	230.4	0	0	0	27	3.7	70.4	25.9	233.6	
Hispanic	2,024	25	9	41	1,949	37.8	58.9	3.3	206.9	24	10	3	1,987	46.2	44.1	9.7	202.1	22	11	3	1,988	32.7	57.8	9.4	211.3	
Amer Indian/AK Native	12	0	0	0	12	25.0	75.0	0.0	211.0	0	0	0	12	50.0	50.0	0.0	202.0	0	0	0	12	16.7	75.0	8.3	224.0	
Other <sup>6</sup>	118	0	2	1	115	33.9	65.2	0.9	205.5	0	1	1	116	40.5	44.0	15.5	207.1	0	1	1	116	25.9	63.8	10.3	216.9	
<b>Economic Status</b>																										
Econ Disadvantaged	3,349	29	29	58	3,233	42.3	54.8	3.0	202.0	28	36	10	3,275	47.6	42.9	9.5	201.1	24	37	8	3,280	34.5	56.2	9.4	210.0	
Non-Econ Disadvantaged	7,042	61	21	47	6,913	24.6	67.8	7.6	214.5	60	28	7	6,947	29.0	51.2	19.8	215.6	58	34	3	6,947	16.7	61.4	21.9	224.3	

1 Includes students coded as LEP exempt (LAL only).  
 2 Students appear in each applicable category, but are included in Total Students only once.  
 3 Includes students coded as former LEP.  
 4 Includes students coded LEP or former LEP.  
 5 Excludes students who did not have gender coded.  
 6 Includes students who did not have an ethnicity coded and students who had more than one ethnicity coded.



## New Jersey Statewide Testing System Grade Eight Proficiency Assessment Performance by Demographic Groups DFG DE

Test Date: March 2007  
CYCLE II

	Language Arts Literacy										Mathematics										Science									
	Students Enrolled	APA Students	Not Present	Voids <sup>1</sup>	Valid Scale Scores	% Proficient	% Partially Proficient	% Advanced Proficient	Scale Score Mean	APA Students	Not Present	Voids	Valid Scale Scores	% Proficient	% Partially Proficient	% Advanced Proficient	Scale Score Mean	APA Students	Not Present	Voids	Valid Scale Scores	% Proficient	% Partially Proficient	% Advanced Proficient	Scale Score Mean					
<b>Total Students<sup>2</sup></b>	14,042	129	83	107	13,723	24.1	67.1	8.8	215.4	129	84	15	13,814	30.6	50.3	19.1	214.7	131	89	12	13,810	16.6	59.5	23.9	225.0					
General Education <sup>3</sup>	11,509	0	37	55	11,417	15.3	74.3	10.4	221.1	0	52	9	11,448	22.3	55.2	22.4	220.6	0	55	7	11,447	10.6	61.8	27.7	229.7					
Special Education	2,559	129	35	25	2,170	67.4	31.7	1.0	187.3	129	29	6	2,195	70.2	27.1	2.7	186.5	131	33	5	2,190	44.2	49.6	6.1	203.7					
Limited English Proficient <sup>4</sup>	230	1	11	27	191	69.1	29.8	1.0	184.0	0	3	0	227	70.5	26.9	2.6	186.7	0	1	0	229	63.3	34.9	1.7	192.6					
Current LEP	180	1	11	27	141	76.6	23.4	0.0	177.5	0	3	0	177	77.4	19.8	2.8	182.9	0	1	0	179	69.3	30.2	0.6	188.7					
Former LEP	50	0	0	0	50	48.0	48.0	4.0	202.3	0	0	0	50	46.0	52.0	2.0	200.2	0	0	0	50	42.0	52.0	6.0	206.4					
<b>Gender<sup>5</sup></b>																														
Female	6,690	42	35	47	6,566	16.2	71.3	12.5	221.5	44	30	7	6,609	31.5	51.9	16.6	213.5	46	37	5	6,602	17.5	62.1	20.4	222.9					
Male	7,346	87	46	59	7,154	31.3	63.2	5.5	209.7	85	52	8	7,201	29.8	48.9	21.3	215.9	85	49	7	7,205	15.8	57.1	27.1	227.0					
<b>Migrant Status</b>																														
Migrant	11	0	0	0	11	45.5	54.5	0.0	204.1	0	0	0	11	27.3	72.7	0.0	213.7	0	0	0	11	27.3	54.5	18.2	221.5					
Non-Migrant	14,031	129	83	107	13,712	24.1	67.1	8.8	215.4	129	84	15	13,803	30.6	50.3	19.1	214.7	131	89	12	13,799	16.6	59.5	23.9	225.0					
<b>Ethnicity</b>																														
White	9,475	97	36	44	9,296	19.2	70.7	10.1	218.9	99	41	10	9,325	23.7	54.0	22.3	219.7	100	40	8	9,327	11.1	60.3	28.6	229.9					
Black	2,055	13	22	27	1,993	41.2	55.3	3.5	203.5	12	27	4	2,012	56.6	37.5	5.9	195.7	13	32	2	2,008	34.0	57.3	8.7	209.9					
Asian	794	3	3	8	780	15.9	66.5	17.6	223.1	3	1	0	790	17.2	48.7	34.1	228.4	3	1	1	789	12.4	55.6	31.9	231.1					
Pacific Islander	9	0	1	0	8	25.0	62.5	12.5	216.5	0	0	0	9	66.7	22.2	11.1	197.8	0	0	0	9	22.2	55.6	22.2	215.1					
Hippanic	1,665	15	16	25	1,609	34.9	61.2	3.9	206.3	14	12	1	1,638	43.7	46.8	9.5	203.6	14	10	1	1,640	28.5	59.4	12.1	213.5					
Amer Indian/AK Native	8	0	0	0	8	50.0	50.0	0.0	200.4	0	0	0	8	50.0	25.0	25.0	215.1	0	0	0	8	12.5	50.0	37.5	226.9					
Other <sup>6</sup>	36	1	3	3	29	44.8	55.2	0.0	201.4	1	3	0	32	56.3	28.1	15.6	199.3	1	6	0	29	27.6	58.6	13.8	211.6					
<b>Economic Status</b>																														
Econ Disadvantaged	2,725	34	31	39	2,621	39.3	56.8	3.9	203.9	31	31	4	2,659	48.5	42.5	9.0	200.9	34	28	3	2,660	30.8	57.1	12.1	212.3					
Non-Econ Disadvantaged	11,317	95	52	68	11,102	20.5	69.5	10.0	218.1	98	53	11	11,155	26.3	52.2	21.4	218.0	97	61	9	11,150	13.3	60.0	26.7	228.1					

<sup>1</sup> Includes students coded as LEP exempt (LAL only).  
<sup>2</sup> Students appear in each applicable category, but are included in Total Students only once.  
<sup>3</sup> Includes students coded as former LEP.  
<sup>4</sup> Includes students who did not have gender coded.  
<sup>5</sup> Excludes students who did not have gender coded.  
<sup>6</sup> Includes students who did not have an ethnicity coded and students who had more than one ethnicity coded.

Test Date: March 2007  
CYCLE II

**New Jersey Statewide Testing System  
Grade Eight Proficiency Assessment  
Performance by Demographic Groups  
DFG FG**



	Language Arts Literacy										Mathematics										Science									
	Students Enrolled	APA Students	Not Present	Voids <sup>1</sup>	Valid Scale Scores	% Proficient	% Partially Proficient	% Advanced Proficient	Scale Mean	APA Students	Not Present	Voids	Valid Scale Scores	% Proficient	% Partially Proficient	% Advanced Proficient	Scale Mean	APA Students	Not Present	Voids	Valid Scale Scores	% Proficient	% Partially Proficient	% Advanced Proficient	Scale Mean					
<b>Total Students<sup>2</sup></b>	14,392	70	42	94	14,186	19.8	69.9	10.3	218.4	75	43	17	14,257	23.6	53.5	22.9	219.7	69	56	14	14,253	12.3	61.3	26.4	228.5					
General Education <sup>3</sup>	12,030	0	20	40	11,970	11.6	76.3	12.1	223.8	0	23	10	11,997	15.5	58.1	26.4	225.5	0	35	8	11,987	6.9	62.8	30.3	232.9					
Special Education	2,188	70	14	31	2,073	63.3	35.6	1.1	189.7	75	20	7	2,086	67.1	29.2	3.7	188.9	69	20	6	2,093	38.9	54.6	6.5	205.9					
Limited English Proficient <sup>4</sup>	218	1	8	23	186	70.4	29.6	0.0	184.3	1	0	0	217	60.8	32.7	6.5	192.6	1	1	0	216	62.5	35.2	2.3	193.9					
Current LEP	178	0	8	23	147	74.1	25.9	0.0	181.6	0	0	0	178	62.9	29.8	7.3	191.8	0	1	0	177	64.4	33.3	2.3	193.1					
Former LEP	40	1	0	0	39	56.4	43.6	0.0	194.6	1	0	0	39	51.3	46.2	2.6	196.1	1	0	0	39	53.8	43.6	2.6	198.0					
<b>Gender<sup>5</sup></b>																														
Female	7,065	30	24	43	6,968	12.7	72.3	15.0	224.9	33	21	4	7,007	24.0	55.7	20.2	218.5	30	27	6	7,002	13.2	65.8	22.9	226.5					
Male	7,310	40	17	51	7,202	26.6	67.5	5.9	212.2	42	21	13	7,234	23.2	51.3	25.4	221.0	39	28	8	7,235	11.4	58.8	29.9	230.4					
<b>Migrant Status</b>																														
Migrant	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0					
Non-Migrant	14,392	70	42	94	14,186	19.8	69.9	10.3	218.4	75	43	17	14,257	23.6	53.5	22.9	219.7	69	56	14	14,253	12.3	61.3	26.4	228.5					
<b>Ethnicity</b>																														
White	10,847	48	26	55	10,718	17.0	72.1	10.9	220.2	53	28	13	10,753	19.9	55.6	24.5	222.2	47	40	13	10,747	9.2	61.8	29.0	231.2					
Black	1,157	6	7	14	1,130	36.1	59.1	4.8	206.7	6	11	3	1,137	48.4	44.1	7.6	200.5	6	8	1	1,142	29.0	61.8	9.2	212.4					
Asian	924	3	4	8	909	12.9	68.4	18.7	226.6	2	0	0	922	13.0	46.4	40.6	233.3	3	0	0	921	9.3	54.6	36.0	234.1					
Pacific Islander	94	1	0	1	92	16.3	77.2	6.5	220.8	1	0	0	93	14.0	55.9	30.1	227.3	1	0	0	93	6.5	61.3	32.3	235.1					
Hispanic	1,304	6	3	16	1,279	33.2	62.0	4.8	208.2	6	2	1	1,295	40.2	50.2	9.6	205.8	6	6	0	1,292	25.2	61.5	13.4	215.8					
Amer Indian/AK Native	14	0	0	0	14	21.4	71.4	7.1	214.1	1	0	0	13	38.5	30.8	30.8	220.5	0	0	0	14	14.3	57.1	28.6	225.6					
Other <sup>6</sup>	52	6	2	0	44	38.6	52.3	9.1	204.4	6	2	0	44	43.2	38.6	18.2	209.6	6	2	0	44	20.5	59.1	20.5	222.5					
<b>Economic Status</b>																														
Econ Disadvantaged	1,875	13	10	34	1,818	37.6	58.6	3.8	205.7	14	10	3	1,848	44.3	47.8	7.9	202.6	13	14	4	1,844	27.7	62.1	10.2	213.4					
Non-Econ Disadvantaged	12,517	57	32	60	12,368	17.2	71.5	11.3	220.3	61	33	14	12,409	20.5	54.4	25.1	222.3	56	42	10	12,409	10.0	61.1	28.8	230.7					

<sup>1</sup> Includes students coded as LEP exempt (LAL only).  
<sup>2</sup> Students appear in each applicable category, but are included in Total Students only once.  
<sup>3</sup> Includes students coded as former LEP.  
<sup>4</sup> Includes students coded LEP or former LEP.  
<sup>5</sup> Excludes students who did not have gender coded.  
<sup>6</sup> Includes students who did not have an ethnicity coded and students who had more than one ethnicity coded.



**New Jersey Statewide Testing System  
Grade Eight Proficiency Assessment  
Performance by Demographic Groups  
DFG GH**

Test Date: March 2007  
CYCLE II

	Language Arts Literacy										Mathematics										Science									
	Students Enrolled	APA Students	Not Present	Voide <sup>1</sup>	Valid Scale Scores	% Partially Proficient	% Proficient	% Advanced Proficient	Scale Score Mean	APA Students	Not Present	Voide	Valid Scale Scores	% Partially Proficient	% Proficient	% Advanced Proficient	Scale Score Mean	APA Students	Not Present	Voide	Valid Scale Scores	% Partially Proficient	% Proficient	% Advanced Proficient	Scale Score Mean					
<b>Total Students<sup>2</sup></b>	14,249	79	60	118	13,992	16.0	67.1	16.8	223.6	83	60	19	14,087	20.5	49.0	30.5	224.9	77	63	17	14,092	11.4	55.5	35.1	231.9					
General Education <sup>3</sup>	11,895	0	25	36	11,832	8.6	71.8	19.6	229.2	0	32	10	11,853	12.5	52.4	35.1	231.1	0	31	6	11,858	5.7	56.6	37.8	236.9					
Special Education	2,142	79	22	22	2,019	55.9	42.3	1.7	193.4	83	25	9	2,025	62.9	31.4	5.7	191.4	77	29	11	2,025	39.5	51.7	8.8	206.9					
Limited English Proficient <sup>4</sup>	304	1	13	59	231	58.0	41.1	0.9	190.9	0	3	0	301	54.5	33.9	11.6	198.6	0	3	0	301	55.1	40.9	4.0	197.9					
Current LEP	225	0	13	59	153	73.2	26.8	0.0	181.3	0	3	0	222	62.2	26.6	11.3	193.7	0	3	0	222	64.9	32.4	2.7	192.3					
Former LEP	79	1	0	0	78	28.2	69.2	2.6	209.6	0	0	0	79	32.9	54.4	12.7	212.3	0	0	0	79	27.8	64.6	7.6	213.5					
<b>Gender<sup>5</sup></b>																														
Female	6,991	32	31	53	6,875	10.1	67.6	22.3	229.5	38	27	4	6,922	20.6	51.6	27.8	223.6	32	33	3	6,923	11.9	59.3	28.9	229.7					
Male	7,250	45	28	65	7,112	21.7	66.7	11.6	217.8	43	33	15	7,159	20.4	46.5	35.1	226.1	43	30	14	7,163	10.8	51.9	37.2	234.1					
<b>Migrant Status</b>																														
Migrant	2	0	0	0	2	50.0	50.0	0.0	207.0	0	0	0	2	50.0	50.0	0.0	210.0	0	0	0	2	0.0	100.0	0.0	225.0					
Non-Migrant	14,247	79	60	118	13,990	16.0	67.1	16.8	223.6	83	60	19	14,085	20.5	49.0	30.5	224.9	77	63	17	14,090	11.4	55.5	35.1	231.9					
<b>Ethnicity</b>																														
White	9,732	50	33	47	9,602	12.8	69.7	17.5	225.8	53	36	13	9,630	16.3	51.1	32.6	227.8	48	36	12	9,636	8.1	55.8	36.2	234.9					
Black	1,483	12	11	14	1,446	31.6	61.7	6.7	209.6	13	12	3	1,455	43.9	47.5	8.6	202.7	12	15	3	1,455	25.8	61.2	13.0	215.9					
Asian	1,742	3	7	29	1,703	10.4	61.0	28.7	232.2	3	5	0	1,734	10.4	40.0	49.6	240.0	3	4	0	1,735	7.3	47.1	45.5	239.8					
Pacific Islander	33	0	0	0	33	3.0	72.7	24.2	233.2	0	0	0	33	6.1	39.4	54.5	245.9	0	0	0	33	0.0	65.6	36.4	240.6					
Hispanic	1,211	12	7	27	1,165	31.2	62.1	6.6	210.4	12	6	3	1,190	39.9	47.8	12.3	206.5	12	7	2	1,190	25.9	59.2	15.0	216.5					
Amer Indian/AK Native	16	1	0	0	15	33.3	66.7	0.0	206.9	1	0	0	15	46.7	40.0	13.3	198.9	1	0	0	15	13.3	73.3	13.3	217.1					
Other <sup>6</sup>	32	1	2	1	28	46.4	32.1	21.4	204.9	1	1	0	30	50.0	30.0	20.0	207.1	1	1	0	30	40.0	33.3	26.7	214.4					
<b>Economic Status</b>																														
Econ Disadvantaged	1,486	11	10	32	1,433	36.2	58.9	4.9	206.2	13	11	4	1,458	43.1	46.2	10.7	204.3	11	11	1	1,463	30.6	58.8	10.6	213.3					
Non-Econ Disadvantaged	12,763	68	50	86	12,559	13.7	68.1	18.2	225.5	70	49	15	12,629	17.9	49.3	32.8	227.3	66	52	16	12,629	9.1	55.2	35.7	234.1					

<sup>1</sup> Includes students coded as LEP exempt (LAL only).  
<sup>2</sup> Students appear in each applicable category, but are included in Total Students only once.  
<sup>3</sup> Includes students coded LEP or former LEP.  
<sup>4</sup> Includes students coded LEP or former LEP.  
<sup>5</sup> Excludes students who did not have gender coded.  
<sup>6</sup> Includes students who did not have an ethnicity coded and students who had more than one ethnicity coded.



**New Jersey Statewide Testing System  
Grade Eight Proficiency Assessment  
Performance by Demographic Groups  
DFG I**

Test Date: March 2007  
CYCLE II

	Language Arts Literacy										Mathematics										Science									
	Students Enrolled	APA Students	Net Present	Voide <sup>1</sup>	Valid Scale Scores	% Partially Proficient	% Proficient	Advanced Proficient	Scale Score Mean	APA Students	Net Present	Voide	Valid Scale Scores	% Partially Proficient	% Proficient	Advanced Proficient	Scale Score Mean	APA Students	Net Present	Voide	Valid Scale Scores	% Partially Proficient	% Proficient	Advanced Proficient	Scale Score Mean					
<b>Total Students<sup>2</sup></b>	20,354	79	64	68	20,143	10.8	68.7	20.5	228.2	80	61	8	20,205	14.7	48.1	37.1	231.0	63	73	20	20,198	6.8	52.1	41.1	237.7					
General Education <sup>3</sup>	17,336	0	24	34	17,278	5.1	71.6	23.4	232.8	0	37	5	17,294	8.6	49.4	42.0	236.4	0	37	11	17,288	3.1	51.0	45.9	241.8					
Special Education	2,834	79	24	10	2,721	45.1	51.7	3.1	200.7	80	24	3	2,727	52.0	40.6	7.4	198.3	63	36	9	2,726	28.0	59.2	12.7	214.3					
Limited English Proficient <sup>4</sup>	236	0	17	24	195	51.3	45.6	3.1	198.5	1	0	1	234	38.0	38.9	23.1	212.9	0	1	0	235	37.9	51.5	10.6	207.7					
Current LEP	194	0	17	24	153	57.5	40.5	2.0	194.4	1	0	1	192	40.1	40.6	19.3	210.2	0	1	0	193	41.5	50.8	7.8	203.9					
Former LEP	42	0	0	0	42	28.6	64.3	7.1	213.4	0	0	0	42	28.6	31.0	40.5	225.6	0	0	0	42	21.4	54.8	23.8	225.3					
<b>Gender<sup>5</sup></b>																														
Female	9,798	24	26	27	9,721	6.4	65.7	27.9	230.1	23	29	2	9,744	15.6	51.2	33.3	228.8	21	29	2	9,746	7.2	56.8	36.0	235.1					
Male	10,545	55	38	40	10,412	15.0	71.4	13.6	222.7	57	32	6	10,450	13.9	45.3	40.8	233.1	42	44	18	10,441	6.5	47.7	45.8	240.1					
<b>Migrant Status</b>																														
Migrant	2	0	0	0	2	50.0	50.0	0.0	203.5	0	0	0	2	50.0	50.0	0.0	195.5	0	0	0	2	100.0	0.0	0.0	184.5					
Non-Migrant	20,352	79	64	68	20,141	10.8	68.7	20.5	228.2	80	61	8	20,203	14.7	48.1	37.2	231.0	63	73	20	20,196	6.8	52.1	41.1	237.7					
<b>Ethnicity</b>																														
White	16,059	56	38	38	15,927	9.2	70.5	20.3	229.0	57	47	5	15,950	12.7	50.4	36.9	231.8	47	55	16	15,941	5.2	52.5	42.3	239.0					
Black	1,154	4	4	1	1,145	30.9	60.5	8.6	212.4	3	6	0	1,145	43.7	44.1	12.2	205.0	3	4	0	1,147	24.7	62.2	13.2	216.9					
Asian	2,102	11	15	12	2,064	7.3	59.9	32.8	236.0	11	4	2	2,085	6.6	32.6	60.9	247.3	7	8	4	2,083	4.5	39.4	56.2	245.7					
Pacific Islander	61	2	0	0	59	11.9	66.1	22.0	228.2	2	0	0	59	18.6	42.4	39.0	232.1	2	0	0	59	8.5	52.5	39.0	235.1					
Hispanic	905	5	7	14	879	22.4	67.0	10.6	216.2	6	4	1	894	32.1	49.8	18.1	212.4	3	6	0	896	17.4	62.5	20.1	222.3					
Amer Indian/AK Native	17	1	0	1	15	26.7	66.7	6.7	212.7	1	0	0	16	37.5	31.3	31.3	215.6	1	0	0	16	18.8	43.8	37.5	231.8					
Other <sup>6</sup>	56	0	0	2	54	11.1	66.7	22.2	225.9	0	0	0	56	16.1	39.3	44.6	231.8	0	0	0	56	10.7	50.0	39.3	238.4					
<b>Economic Status</b>																														
Econ Disadvantaged	746	7	4	9	726	34.7	58.4	6.9	208.5	7	5	0	734	42.2	46.0	11.7	204.9	4	7	0	735	27.8	58.2	14.0	214.6					
Non-Econ Disadvantaged	19,608	72	60	59	19,417	10.0	69.1	21.0	228.9	73	56	8	19,471	13.7	48.2	38.1	232.0	59	66	20	19,463	6.0	51.9	42.1	238.6					

<sup>1</sup> Includes students coded as LEP exempt (LAL only).  
<sup>2</sup> Students appear in each applicable category, but are included in Total Students only once.  
<sup>3</sup> Includes students coded LEP or former LEP.  
<sup>4</sup> Includes students who did not have gender coded.  
<sup>5</sup> Excludes students who did not have ethnicity coded and students who had more than one ethnicity coded.  
<sup>6</sup> Includes students who did not have an ethnicity coded and students who had more than one ethnicity coded.



**New Jersey Statewide Testing System  
Grade Eight Proficiency Assessment  
Performance by Demographic Groups  
DFG J**

Test Date: March 2007  
CYCLE II

	Language Arts Literacy						Mathematics						Science											
	Students Enrolled	APA Students	Net Present	Voids <sup>1</sup>	Valid Scale Scores	% Partially Proficient	% Advanced Proficient	Scale Score Mean	APA Students	Net Present	Voids	Valid Scale Scores	% Partially Proficient	% Proficient	% Advanced Proficient	Scale Score Mean	APA Students	Net Present	Voids	Valid Scale Scores	% Partially Proficient	% Proficient	% Advanced Proficient	Scale Score Mean
<b>Total Students<sup>2</sup></b>	4,209	15	9	20	4,165	5.7	66.5	27.8 239.5	11	8	2	4,188	8.3	42.5	49.2 240.4	16	9	0	4,184	3.1	45.3	51.6	244.2	
General Education <sup>3</sup>	3,640	0	1	8	3,631	2.3	66.6	31.1 238.1	0	6	2	3,632	4.0	41.5	54.4 244.8	0	6	0	3,634	1.2	42.5	56.3	247.3	
Special Education	533	15	3	5	510	28.0	67.1	4.9 210.0	11	2	0	520	37.1	48.3	14.6 210.4	16	3	0	514	14.4	64.0	21.6	224.9	
Limited English Proficient <sup>4</sup>	48	0	5	7	36	44.4	47.2	8.3 198.4	0	0	0	48	25.0	50.0	25.0 222.9	0	0	0	48	31.3	66.7	2.1	209.4	
Current LEP	39	0	5	7	27	51.9	37.0	11.1 193.6	0	0	0	39	25.6	51.3	23.1 221.6	0	0	0	39	35.9	61.5	2.6	207.8	
Former LEP	9	0	0	0	9	22.2	77.8	0.0 212.8	0	0	0	9	22.2	44.4	33.3 228.6	0	0	0	9	11.1	88.9	0.0	216.6	
<b>Gender<sup>5</sup></b>																								
Female	2,096	3	3	8	2,082	3.4	60.7	35.9 239.4	1	3	1	2,091	9.0	45.8	45.2 237.9	3	3	0	2,090	3.3	51.3	45.4	241.3	
Male	2,113	12	6	12	2,083	8.1	72.2	19.7 229.5	10	5	1	2,097	7.5	39.2	53.3 242.8	13	6	0	2,094	2.9	39.4	57.8	247.1	
<b>Migrant Status</b>																								
Migrant	0	0	0	0	0				0	0	0	0				0	0	0	0					
Non-Migrant	4,209	15	9	20	4,165	5.7	66.5	27.8 239.5	11	8	2	4,188	8.3	42.5	49.2 240.4	16	9	0	4,184	3.1	45.3	51.6	244.2	
<b>Ethnicity</b>																								
White	3,285	14	4	15	3,252	5.3	69.0	25.8 233.9	11	7	2	3,265	7.8	46.2	46.0 238.7	15	7	0	3,263	2.5	47.0	50.5	243.8	
Black	87	0	0	0	87	16.1	71.3	12.6 222.6	0	0	0	87	36.8	44.8	18.4 213.9	0	0	0	87	11.5	57.5	31.0	228.5	
Asian	707	1	5	5	696	5.0	52.7	42.2 240.6	0	0	0	707	5.5	23.6	70.9 253.6	1	0	0	706	3.7	33.3	63.0	249.9	
Pacific Islander	3	0	0	0	3	0.0	100.0	0.0 224.0	0	0	0	3	0.0	33.3	66.7 241.3	0	0	0	3	0.0	33.3	66.7	239.0	
Hispanic	122	0	0	0	122	13.9	73.8	12.3 224.8	0	0	0	122	14.8	52.5	32.8 228.4	0	0	0	122	7.4	62.3	30.3	233.8	
Amer Indian/AK Native	1	0	0	0	1	0.0	100.0	0.0 200.0	0	0	0	1	100.0	0.0	0.0 188.0	0	0	0	1	0.0	100.0	0.0	223.0	
Other <sup>6</sup>	4	0	0	0	4	50.0	50.0	0.0 186.8	0	1	0	3	66.7	0.0	33.3 206.3	0	2	0	2	50.0	50.0	0.0	191.5	
<b>Economic Status</b>																								
Econ Disadvantaged	54	0	3	0	51	21.6	72.5	5.9 214.8	0	0	0	54	35.2	53.7	11.1 210.2	0	1	0	53	20.8	66.0	13.2	218.6	
Non-Econ Disadvantaged	4,155	15	6	20	4,114	5.5	66.4	28.1 234.7	11	8	2	4,134	7.9	42.3	49.7 240.7	16	8	0	4,131	2.8	45.1	52.1	244.5	

<sup>1</sup> Includes students coded as LEP exempt (LAL only).  
<sup>2</sup> Students appear in each applicable category, but are included in Total Students only once.  
<sup>3</sup> Includes students coded LEP or former LEP.  
<sup>4</sup> Includes students who did not have gender coded.  
<sup>5</sup> Excludes students who did not have gender coded.  
<sup>6</sup> Includes students who did not have an ethnicity coded and students who had more than one ethnicity coded.



**New Jersey Statewide Testing System  
Grade Eight Proficiency Assessment  
Performance by Demographic Groups  
DFG R**

Test Date: March 2007  
CYCLE II

	Language Arts Literacy										Mathematics										Science									
	Students Enrolled	APA Students	Not Present	Voids <sup>1</sup>	Valid Scale Scores	% Partially Proficient	% Proficient	% Advanced Proficient	Scale Score Mean	APA Students	Not Present	Voids	Valid Scale Scores	% Partially Proficient	% Proficient	% Advanced Proficient	Scale Score Mean	APA Students	Not Present	Voids	Valid Scale Scores	% Partially Proficient	% Proficient	% Advanced Proficient	Scale Score Mean					
<b>Total Students<sup>2</sup></b>	1,384	2	3	16	1,363	47.2	49.2	3.5	201.7	2	4	3	1,375	55.9	35.1	9.1	198.1	2	6	6	1,370	36.9	52.7	10.4	209.6					
General Education <sup>3</sup>	1,193	0	2	5	1,186	41.8	54.3	3.9	205.1	0	4	0	1,189	51.4	38.2	10.4	201.4	0	6	2	1,185	33.2	55.4	11.4	211.8					
Special Education	181	2	1	7	171	83.6	15.2	1.2	178.5	2	0	3	176	84.7	14.8	0.6	176.8	2	0	4	175	60.0	35.4	4.6	195.3					
Limited English Proficient <sup>4</sup>	12	0	0	4	8	87.5	12.5	0.0	188.1	0	0	0	12	83.3	16.7	0.0	174.5	0	0	0	12	58.3	41.7	0.0	194.0					
Current LEP	11	0	0	4	7	85.7	14.3	0.0	188.1	0	0	0	11	81.8	18.2	0.0	175.0	0	0	0	11	54.5	45.5	0.0	195.3					
Former LEP	1	0	0	0	1	100.0	0.0	0.0	188.0	0	0	0	1	100.0	0.0	0.0	169.0	0	0	0	1	100.0	0.0	0.0	180.0					
<b>Gender<sup>5</sup></b>																														
Female	744	1	0	7	736	40.2	54.8	5.0	206.2	1	1	0	742	55.7	36.1	8.2	197.5	1	3	1	739	38.3	53.7	8.0	207.8					
Male	640	1	3	9	627	55.5	42.7	1.8	196.4	1	3	3	633	56.1	35.8	10.1	198.8	1	3	5	631	35.2	51.5	13.3	211.6					
<b>Migrant Status</b>																														
Migrant	3	0	0	2	1	100.0	0.0	0.0	194.0	0	0	0	3	66.7	33.3	0.0	180.0	0	0	0	3	33.3	66.7	0.0	203.7					
Non-Migrant	1,381	2	3	14	1,362	47.2	49.3	3.5	201.7	2	4	3	1,372	55.8	35.1	9.1	198.1	2	6	6	1,367	36.9	52.7	10.5	209.6					
<b>Ethnicity</b>																														
White	169	0	1	3	165	27.3	61.8	10.9	217.8	0	1	0	168	29.2	43.5	27.4	220.3	0	1	0	168	10.7	56.0	33.3	233.1					
Black	878	0	2	9	867	53.1	45.2	1.7	197.4	0	2	2	874	64.8	31.5	3.8	191.3	0	3	4	871	44.7	50.6	4.7	203.1					
Asian	29	0	0	0	29	10.3	65.5	24.1	227.9	0	0	0	29	6.9	51.7	41.4	238.0	0	0	0	29	3.4	48.3	48.3	242.0					
Pacific Islander	5	0	0	0	5	20.0	60.0	20.0	231.2	0	0	0	5	20.0	0.0	80.0	250.4	0	0	0	5	20.0	60.0	20.0	224.2					
Hispanic	291	2	0	4	285	46.0	51.6	2.5	202.4	2	1	1	287	50.2	39.4	10.5	200.8	2	2	2	285	32.3	56.8	10.9	212.0					
Amer Indian/AK Native	0	0	0	0	0					0	0	0	0					0	0	0	0									
Other <sup>6</sup>	12	0	0	0	12	33.3	66.7	0.0	201.7	0	0	0	12	50.0	50.0	0.0	194.5	0	0	0	12	33.3	66.7	0.0	206.3					
<b>Economic Status</b>																														
Econ Disadvantaged	857	2	1	10	844	54.0	43.8	2.1	196.8	2	3	3	849	62.4	31.8	5.8	192.9	2	4	6	845	44.0	50.3	5.7	204.1					
Non-Econ Disadvantaged	527	0	2	6	519	36.2	59.0	5.8	209.7	0	1	0	526	45.2	40.3	14.4	206.4	0	2	0	525	25.3	56.6	18.1	218.4					

1 Includes students coded as LEP exempt (LAL only).  
 2 Students appear in each applicable category, but are included in Total Students only once.  
 3 Includes students coded as former LEP.  
 4 Includes students coded LEP or former LEP.  
 5 Excludes students who did not have gender coded.  
 6 Includes students who did not have an ethnicity coded and students who had more than one ethnicity coded.

New Jersey Statewide Testing System  
Grade Eight Proficiency Assessment  
Performance by Demographic Groups  
Special Needs

Test Date: March 2007  
CYCLE II



	Language Arts Literacy						Mathematics						Science									
	Students Enrolled	APA Students	Net Present	Valid Scale Scores	% Partially Proficient	% Advanced Proficient	Scale Score Mean	APA Students	Net Present	Valid Scale Scores	% Partially Proficient	% Advanced Proficient	Scale Score Mean	APA Students	Net Present	Valid Scale Scores	% Partially Proficient	% Advanced Proficient	Scale Score Mean			
<b>Total Students<sup>2</sup></b>	20,700	135	296	19,675	52.2	45.0	2.8 193.3	133	276	49	20,242	60.3	32.4	7.3 193.3	128	327	45	20,200	50.5	44.4	5.1 200.1	
General Education <sup>3</sup>	14,630	0	142	14,478	40.0	56.2	3.8 204.9	0	154	20	14,656	50.1	40.3	9.6 200.8	0	176	21	14,653	40.4	52.9	6.7 206.1	
Special Education	4,374	135	97	3,988	86.2	13.6	0.2 163.3	133	107	26	4,108	89.2	10.2	0.6 171.7	128	130	23	4,093	76.4	22.8	0.8 184.9	
Limited English Proficient <sup>4</sup>	2,143	2	61	1,848	76.8	22.5	0.8 176.9	2	21	5	2,115	75.6	21.0	3.4 182.5	2	29	1	2,111	72.0	27.0	1.0 187.2	
Current LEP	1,560	1	58	1,271	85.4	14.4	0.2 169.2	1	16	4	1,539	81.9	15.6	2.5 178.1	1	23	1	1,535	79.1	20.4	0.5 183.3	
Former LEP	583	1	3	577	57.7	40.2	2.1 193.8	1	5	1	576	58.7	35.6	5.7 194.2	1	6	0	576	53.1	44.4	2.4 197.6	
<b>Gender<sup>5</sup></b>																						
Female	10,050	42	114	9,664	43.3	52.6	4.2 202.0	39	113	11	9,887	59.2	34.1	6.6 194.1	39	136	15	9,860	51.3	44.7	4.0 199.5	
Male	10,650	90	179	10,000	60.7	37.8	1.5 189.0	91	161	38	10,340	61.3	30.8	7.9 192.5	86	190	30	10,324	49.7	44.2	6.1 200.8	
<b>Migrant Status</b>																						
Migrant	13	0	0	13	92.3	7.7	0.0 166.8	0	0	0	13	84.6	15.4	0.0 179.8	0	0	0	13	69.2	30.8	0.0 188.1	
Non-Migrant	20,687	135	296	19,662	52.1	45.1	2.8 195.4	133	276	49	20,229	60.3	32.4	7.3 193.3	128	327	45	20,187	50.5	44.4	5.1 200.1	
<b>Ethnicity</b>																						
White	2,355	11	16	2,282	34.7	59.2	6.1 207.5	12	20	3	2,320	37.1	46.5	16.4 209.7	12	21	4	2,318	27.4	58.1	14.5 215.4	
Black	8,698	66	112	8,286	60.2	38.5	1.3 190.2	65	153	32	8,448	71.7	25.5	2.8 185.2	60	170	30	8,438	59.1	38.7	2.2 194.9	
Asian	524	4	8	497	32.0	57.9	10.1 212.5	4	1	0	519	30.3	41.6	28.1 218.7	4	1	0	519	30.4	47.8	21.8 217.7	
Pacific Islander	37	0	1	35	51.4	40.0	8.6 195.6	0	2	0	35	51.4	34.3	14.3 200.2	0	0	0	37	54.1	40.5	5.4 195.6	
Hispanic	8,945	51	144	8,458	50.1	47.0	2.9 196.2	49	93	14	8,789	57.1	35.0	7.9 195.3	49	128	10	8,758	49.4	46.3	4.3 200.1	
Amer Indian/AK Native	30	0	1	28	35.7	53.6	10.7 211.8	0	0	0	30	50.0	30.0	20.0 201.1	0	0	0	30	36.7	43.3	20.0 212.9	
Other <sup>6</sup>	111	3	14	89	69.7	25.8	4.5 184.6	3	7	0	101	78.2	17.8	4.0 180.9	3	7	1	100	66.0	30.0	4.0 192.0	
<b>Economic Status</b>																						
Econ Disadvantaged	14,896	80	194	14,187	54.1	43.4	2.5 194.1	79	180	37	14,600	62.2	31.4	6.4 192.0	76	220	31	14,569	52.7	43.2	4.0 198.6	
Non-Econ Disadvantaged	5,804	55	102	5,488	47.2	49.2	3.6 198.7	54	96	12	5,642	55.4	35.1	9.5 196.7	52	107	14	5,631	44.8	47.5	7.7 204.2	

1 Includes students coded as LEP exempt (LAL only).  
 2 Includes students coded as LEP exempt (LAL only) and students coded as former LEP.  
 3 Includes students coded as former LEP.  
 4 Includes students coded LEP or former LEP.  
 5 Excludes students who did not have gender coded.  
 6 Includes students who did not have an ethnicity coded and students who had more than one ethnicity coded.

**New Jersey Statewide Testing System  
Grade Eight Proficiency Assessment  
Performance by Demographic Groups  
Non Special Needs**

Test Date: March 2007  
CYCLE II



	Students Enrolled	Language Arts Literacy						Mathematics						Science												
		APA students	Not Present	Valid Scale Scores	% Partially Proficient	% Advanced Proficient	Scale Score Mean	APA students	Not Present	Valid Scale Scores	% Partially Proficient	% Advanced Proficient	Scale Score Mean	APA students	Not Present	Valid Scale Scores	% Partially Proficient	% Advanced Proficient	Scale Score Mean							
				Voids <sup>1</sup>																						
<b>Total Students<sup>2</sup></b>	87,774	545	371	668	86,190	20.5	66.3	13.2	219.3	548	385	103	86,738	24.9	49.0	26.1	220.7	517	440	104	86,713	14.3	56.5	29.2	228.7	
General Education <sup>3</sup>	72,566	0	169	256	72,141	12.3	72.2	15.5	225.2	0	231	51	72,284	16.8	52.9	30.4	226.9	0	254	50	72,262	8.5	58.0	33.6	233.5	
Special Education	13,823	545	132	163	12,983	61.3	37.2	1.5	190.3	548	143	50	13,082	65.5	30.0	4.5	189.4	517	176	50	13,080	41.1	51.2	7.7	205.6	
United English Proficient <sup>4</sup>	1,877	4	71	254	1,548	67.7	31.3	1.0	185.1	3	11	3	1,860	61.6	30.1	8.3	193.3	2	12	4	1,859	59.5	37.3	3.2	195.2	
Current LEP	1,439	2	71	253	1,113	76.6	22.7	0.6	178.3	2	11	3	1,423	66.3	26.1	7.6	190.0	1	11	4	1,423	65.3	32.4	2.3	191.5	
Former LEP	438	2	0	1	435	44.8	53.3	1.8	202.6	1	0	0	437	46.5	43.0	18.5	204.0	1	1	0	436	40.6	53.4	6.0	207.2	
<b>Gender<sup>5</sup></b>																										
Female	42,549	196	160	278	41,915	13.8	68.0	18.2	225.5	206	153	28	42,162	25.3	51.3	23.4	219.4	194	190	29	42,136	14.8	59.9	25.3	226.5	
Male	45,173	346	206	386	44,235	26.7	64.8	8.5	213.6	339	228	75	44,831	24.5	46.8	28.7	221.9	320	245	75	44,533	13.7	53.4	32.9	230.7	
<b>Migrant Status</b>																										
Migrant	28	0	0	3	25	48.0	52.0	0.0	200.2	0	0	0	28	39.3	60.7	0.0	202.4	0	0	0	28	42.9	46.4	10.7	210.0	
Non-Migrant	87,746	545	371	665	86,165	20.5	66.3	13.2	219.4	548	385	103	86,710	24.9	49.0	26.1	220.7	517	440	104	86,685	14.3	56.5	29.2	228.7	
<b>Ethnicity</b>																										
White	59,241	349	172	263	58,457	15.2	70.2	14.7	223.1	356	203	59	58,623	18.4	52.3	29.3	225.3	336	229	63	58,613	8.9	57.2	33.9	233.4	
Black	10,579	74	85	116	10,304	42.1	53.7	4.2	203.0	72	107	25	10,375	53.6	39.7	6.7	197.2	70	120	15	10,374	34.3	57.3	8.5	209.5	
Asian	7,205	24	39	74	7,068	11.5	61.9	26.6	230.6	22	11	3	7,169	11.0	39.3	49.7	239.6	20	15	8	7,162	8.0	47.0	45.0	239.1	
Pacific Islander	256	3	2	1	250	12.4	74.0	13.6	224.9	3	0	0	253	15.0	50.6	34.4	230.2	3	0	0	253	6.3	60.9	32.8	233.4	
Hispanic	10,059	81	61	202	9,715	35.7	59.6	4.7	206.4	80	53	15	9,911	42.7	46.5	10.8	204.3	74	61	17	9,907	29.5	58.5	12.0	213.3	
Am. Indian/AK Native	84	2	0	1	81	28.4	67.9	3.7	209.8	3	0	0	81	40.7	38.3	21.0	210.4	2	0	0	82	15.9	63.4	20.7	222.2	
Other <sup>6</sup>	350	12	12	11	315	35.2	57.8	7.0	206.2	12	11	1	326	41.1	39.3	19.6	208.3	12	15	1	322	27.0	55.6	17.4	218.2	
<b>Economic Status</b>																										
Econ Disadvantaged	14,887	134	118	264	14,371	42.5	54.1	3.4	202.0	132	131	42	14,562	48.7	42.8	8.5	200.3	123	144	40	14,580	33.9	56.6	9.5	210.1	
Non-Econ Disadvantaged	72,887	411	253	404	71,819	16.1	68.8	15.2	222.8	416	254	61	72,156	20.1	50.3	29.7	224.8	394	296	64	72,133	10.3	56.5	33.2	232.4	

<sup>1</sup> Includes students coded as LEP exempt (LAL only).  
<sup>2</sup> Students appear in each applicable category, but are included in Total Students only once.  
<sup>3</sup> Includes students coded as former LEP.  
<sup>4</sup> Includes students coded LEP or former LEP.  
<sup>5</sup> Excludes students who did not have gender coded.  
<sup>6</sup> Includes students who did not have an ethnicity coded and students who had more than one ethnicity coded.

## **Appendix C**

### ***Modifications of Test Administration Procedures for Special Education Students and Students Eligible Under Section 504 of the Rehabilitation Act of 1973***

In accordance with the Individuals with Disabilities Education Act (IDEA), students who are receiving special education services must participate in each subject area of the age-appropriate statewide assessment with the following exception:

Students with disabilities shall participate in the Alternate Proficiency Assessment in each content area where the nature of the student’s disability is so severe that the student is not receiving instruction in any of the knowledge and skills measured by the general statewide assessment and the student cannot complete any of the types of questions on the assessment content area(s) even with accommodation and modifications. (New Jersey Administrative Code Chapter 6A:14-4.11[a]2)

Districts may use modifications of test administration procedures when administering the GEPA to special education students or to students eligible under Section 504 of the Rehabilitation Act of 1973. Decisions about participation and accommodations/modifications are made by the *Individualized Education Program* (IEP) or 504 team. Information about test content and item types from the test specifications booklets can be used to make this determination. Modifications in the areas listed below may be used separately or in combination.

Any accommodations or modifications of test administration procedures for students eligible for special education under the IDEA or eligible under Section 504 of the Rehabilitation Act of 1973 must be specified in the student’s IEP or 504 accommodation plan. Accommodations or modifications must be consistent with the instruction and assessment procedures used in the student’s classroom. Students eligible for modifications under Section 504 may not be classified but do have a permanent or temporary impairment in a major life function (for example: performing manual tasks, walking, seeing, hearing, speaking, etc.).

Advanced planning is integral to implementing accommodations/modifications effectively and ensuring that the security of test materials is maintained. If a student requires an accommodation or modification that is not listed below, contact the Office of State Assessments, GEPA Coordinator.

Accommodations must be recorded on the student’s answer folder by the codes (A, B, C, or D) listed in this appendix. Verify that the coding on the Pre-ID labels is correct.

## ACCEPTABLE ACCOMMODATIONS OR MODIFICATIONS

### Code

#### A. Setting Accommodations

1. Administering the assessment:
  - a. individually in a separate room
  - b. in a small group in a separate room
  - c. in the resource room
  - d. in a special education classroom
  - e. using carrels
  - f. at home or in a hospital (this will depend on the nature of the assessment task)
2. Seating the student in the front of the room near the examiner or proctor
3. Seating the student facing the examiner or proctor
4. Providing special lighting

5. Providing special furniture (e.g., desks, trays, carrels)

#### **B. Scheduling Accommodations**

1. Adding time as needed
2. Providing frequent breaks
3. Terminating a section of the test when a student has indicated that he/she has completed all the items he/she can. The test examiner must ensure that the student has attempted all items in a section since items are not ordered by difficulty. When this accommodation is used, the test must be administered in a small group or individually to avoid distraction.

#### **C. Test Materials Modifications**

1. Administering the large-print version of test materials
2. Administering the Braille version of test materials

#### **D. Test Procedures Modifications**

1. Administration modifications
  - a. reading directions aloud
  - b. reading test items aloud (**YOU MAY NOT READ ALOUD OR SIGN THE READING PASSAGES IN LANGUAGE ARTS LITERACY—YOU MAY READ ONLY THE READING ITEMS ASSOCIATED WITH THE PASSAGE**); **ONLY the teacher who must read test items aloud is permitted to have a test book assigned to them for this task.**
  - c. providing and ensuring that amplification (hearing aid and/or FM system) is in working order
  - d. using a sign language or cued speech interpreter for administration of directions or items **but not reading passages**
  - e. masking a portion of the test booklet and/or answer folder to eliminate visual distractors or providing reading windows
  - f. repeating, clarifying, or rewording directions
  - g. providing written directions on a separate sheet or transparency
  - h. using an examiner who is familiar with the student
  - i. using an examiner who can communicate fluently in sign language (American Sign Language or a form of Manually Coded English)
  - j. providing manipulatives for math items
  - k. using graph paper for math section
  - l. using a Braille ruler and talking calculator
  - m. using tactile or visual cues for deaf or hard of hearing students to indicate time to begin, time remaining, and time to end a particular part of the test

2. Response modifications
  - a. having an examiner record the student's identifying information on the answer folder, or grid corrections to the pre-ID label
  - b. dictating oral responses to a scribe (person who writes from dictation) – student must indicate all punctuation and must spell all key words (see FAQ – *Test Manual*, Appendix J)
  - c. using a Braille writer to record responses
  - d. signing responses to a sign language interpreter (student must indicate all punctuation and must spell all key words)
  - e. recording responses on a word processor
  - f. using large-face calculators
  - g. using talking calculators
  - h. providing an Augmentative Communication device
  - i. using a larger diameter or modified special grip #2 pencil
  - j. masking portions of the answer folder to eliminate visual distractors
  - k. marking answers in the test booklet (an examiner would transfer the answers to an answer folder)
  - l. Allowing separate additional continuation pages for writing tasks. These pages **MUST** be properly marked to link them to the correct student for credit.

## OTHER CONSIDERATIONS

Ensure that:

- a. any medication has been appropriately adjusted so it will not interfere with the student's functioning.
- b. eyeglasses are used, if needed.
- c. hearing aids, FM systems, Augmentative Communication devices, word processors, or other equipment are functioning properly.
- d. source and strength of light are appropriate.
- e. all students can clearly see and hear the examiner.
- f. all deaf or hard of hearing students who communicate aurally/orally are watching the examiner when instructions are given.
- g. responses to open-ended items and writing tasks which are written or typed on separate sheets of paper by students eligible for this accommodation are labeled with student data paper-clipped to the front of the answer folder, and placed in the fluorescent orange envelope provided. **Follow packaging instructions in this manual or the student's responses cannot be linked to their responses on the other sections of the test and they will receive incomplete scores.** Copies of these pages should be made and retained on file by the school district until scores are received.

- h. students using the large-print test booklets
  - 1. mark their answers in the large-print answer folder. All responses must be transcribed into the regular answer folder provided in the large print kit.
  - 2. may be instructed to skip items identified in the LP instructions. The spaces for these items must be left blank on the student's answer folder (included in the large-print kit).
  - 3. who dictate responses on open-ended items and writing tasks indicate all punctuation and spell all key words.
- i. students using the Braille test booklets
  - 1. are instructed to bring a Braille ruler and a talking calculator to the test session.
  - 2. are instructed to skip dropped items identified in the Braille instructions. The spaces for these items must be left blank on the student transcription answer folder (included in the Braille kit).
  - 3. have answer folders transcribed from the Braille version by the Examiner.
  - 4. dictate their answers to the examiner or use a device that produces Braille. For dictations and responses recorded in Braille:
    - Students must indicate all punctuation and must spell all key words.
    - Examiners must transcribe the Braille responses into the regular answer folder included in the Braille kit.
- j. students who communicate in sign language
  - 1. have an interpreter to translate oral directions and test items (**but not the Reading passages in the Language Arts Literacy section of the test**). The interpreter should be able to communicate in the mode used by the student, American Sign Language or a form of Manually Coded English. The interpreter should be instructed to interpret so as not to give the answer to the student through the use of a particular sign or finger spelling.
  - 2. using American Sign Language for open-ended and writing task responses will sign the responses to the interpreter who will interpret them into spoken English and a scribe will record the responses in the answer folder.
  - 3. using Signed English or cued speech will sign/cue to the interpreter who will transliterate (word for word) into spoken English and a scribe will record the responses.

For any unresolved questions, contact the Office of Special Education Programs at (609) 292-2912.

# ***Appendix D***

## ***Raw to Scale Scores Conversions***

## 2007 GEPA LAL Raw Score to Scale Score

RS	Ability	SS
0.0	-6.1116	103
0.5	-5.5145	103
1.0	-4.9174	105
1.5	-4.5735	106
2.0	-4.2296	108
2.5	-4.0258	109
3.0	-3.8220	110
3.5	-3.6744	112
4.0	-3.5268	113
4.5	-3.4091	114
5.0	-3.2914	116
5.5	-3.1917	117
6.0	-3.0920	118
6.5	-3.0037	120
7.0	-2.9155	121
7.5	-2.8347	122
8.0	-2.7538	124
8.5	-2.6777	126
9.0	-2.6016	127
9.5	-2.5284	129
10.0	-2.4552	131
10.5	-2.3837	132
11.0	-2.3122	134
11.5	-2.2416	136
12.0	-2.1710	138
12.5	-2.1008	140
13.0	-2.0307	142
13.5	-1.9608	144
14.0	-1.8909	145
14.5	-1.8212	147
15.0	-1.7515	149
15.5	-1.6822	151
16.0	-1.6128	153
16.5	-1.5438	154
17.0	-1.4747	156
17.5	-1.4061	158
18.0	-1.3375	159
18.5	-1.2693	161
19.0	-1.2011	163
19.5	-1.1332	164
20.0	-1.0654	166
20.5	-0.9978	167
21.0	-0.9302	169

RS	Ability	SS
21.5	-0.8627	171
22.0	-0.7952	172
22.5	-0.7275	174
23.0	-0.6598	176
23.5	-0.5916	178
24.0	-0.5235	179
24.5	-0.4545	181
25.0	-0.3856	183
25.5	-0.3155	185
26.0	-0.2454	187
26.5	-0.1738	188
27.0	-0.1022	190
27.5	-0.0286	192
28.0	0.0450	194
28.5	0.1210	196
29.0	0.1969	198
29.5	0.2758	200*
30.0	0.3546	202
30.5	0.4369	204
31.0	0.5192	206
31.5	0.6055	209
32.0	0.6918	211
32.5	0.7828	213
33.0	0.8739	215
33.5	0.9707	217
34.0	1.0674	220
34.5	1.1710	222
35.0	1.2746	224
35.5	1.3866	226
36.0	1.4986	229
36.5	1.6210	231
37.0	1.7435	233
37.5	1.8790	236
38.0	2.0145	238
38.5	2.1659	241
39.0	2.3174	243
39.5	2.4868	246
40.0	2.6563	250*
40.5	2.8431	251
41.0	3.0299	254
41.5	3.2283	257
42.0	3.4268	260
42.5	3.6284	264

RS	Ability	SS
43.0	3.8299	267
43.5	4.0284	270
44.0	4.2270	273
44.5	4.4205	276
45.0	4.6141	279
45.5	4.8018	281
46.0	4.9895	284
46.5	5.1722	286
47.0	5.3549	288
47.5	5.5420	290
48.0	5.7290	291
48.5	5.9441	293
49.0	6.1592	295
49.5	6.4444	297
50.0	6.7296	299
50.5	7.1039	300
51.0	7.4782	300
51.5	7.8967	300
52.0	8.3153	300
52.5	8.8311	300
53.0	9.3469	300
53.5	10.0449	300
54.0	10.7429	300

\* Rounding was applied for this cut score.

### 2007 GEPA Mathematics Raw Score to Scale Score

RS	Ability	SS
0.0	-5.4832	137
1.0	-4.2537	139
2.0	-3.5240	141
3.0	-3.0817	143
4.0	-2.7571	146
5.0	-2.4968	148
6.0	-2.2772	150
7.0	-2.0855	152
8.0	-1.9142	155
9.0	-1.7583	157
10.0	-1.6146	160
11.0	-1.4806	162
12.0	-1.3545	164
13.0	-1.2349	167
14.0	-1.1207	169
15.0	-1.0111	172
16.0	-0.9052	175
17.0	-0.8026	177
18.0	-0.7025	180
19.0	-0.6047	183
20.0	-0.5085	186
21.0	-0.4137	188
22.0	-0.3199	191
23.0	-0.2266	195
24.0	-0.1337	200*

RS	Ability	SS
25.0	-0.0407	201
26.0	0.0527	204
27.0	0.1468	208
28.0	0.2422	211
29.0	0.3390	215
30.0	0.4380	219
31.0	0.5394	223
32.0	0.6441	227
33.0	0.7526	231
34.0	0.8657	235
35.0	0.9846	239
36.0	1.1103	243
37.0	1.2444	250*
38.0	1.3887	252
39.0	1.5457	256
40.0	1.7185	260
41.0	1.9114	264
42.0	2.1304	268
43.0	2.3845	271
44.0	2.6877	274
45.0	3.0650	277
46.0	3.5705	280
47.0	4.3737	283
48.0	5.6598	300*

\* Rounding was applied for this cut score.

### 2007 GEPA Science Raw Score to Scale Score

RS	Ability	SS
0.0	-5.166	126
1.0	-3.943	129
2.0	-3.222	132
3.0	-2.788	136
4.0	-2.472	140
5.0	-2.220	143
6.0	-2.009	146
7.0	-1.825	150
8.0	-1.662	153
9.0	-1.514	157
10.0	-1.378	160
11.0	-1.252	163
12.0	-1.133	167
13.0	-1.021	170
14.0	-0.914	173
15.0	-0.812	177
16.0	-0.714	180
17.0	-0.619	183
18.0	-0.527	186
19.0	-0.437	189
20.0	-0.349	192
21.0	-0.264	195
22.0	-0.179	200*
23.0	-0.096	202
24.0	-0.014	205
25.0	0.068	208
26.0	0.149	211
27.0	0.230	214

RS	Ability	SS
28.0	0.310	217
29.0	0.391	220
30.0	0.472	223
31.0	0.554	226
32.0	0.637	228
33.0	0.720	231
34.0	0.805	234
35.0	0.891	237
36.0	0.980	240
37.0	1.071	242
38.0	1.165	245
39.0	1.262	250*
40.0	1.363	251
41.0	1.470	253
42.0	1.582	256
43.0	1.701	259
44.0	1.829	262
45.0	1.967	264
46.0	2.118	267
47.0	2.285	269
48.0	2.474	272
49.0	2.693	275
50.0	2.954	278
51.0	3.282	280
52.0	3.730	282
53.0	4.470	286
54.0	5.709	300*

\* Rounding was applied for this cut score.

## 2007 GEPA LAL Braille Raw Score to Scale Score

RS	Ability	SS
0.0	-5.8957	103
0.5	-5.2869	103
1.0	-4.6781	106
1.5	-4.3269	107
2.0	-3.9757	109
2.5	-3.7710	111
3.0	-3.5665	113
3.5	-3.4204	114
4.0	-3.2744	116
4.5	-3.1582	117
5.0	-3.0421	119
5.5	-2.9428	121
6.0	-2.8435	122
6.5	-2.7541	124
7.0	-2.6648	126
7.5	-2.5812	128
8.0	-2.4977	130
8.5	-2.4175	131
9.0	-2.3373	133
9.5	-2.2591	136
10.0	-2.1808	138
10.5	-2.1037	140
11.0	-2.0266	142
11.5	-1.9502	144
12.0	-1.8738	146
12.5	-1.7980	148
13.0	-1.7223	150
13.5	-1.6472	152
14.0	-1.5720	154
14.5	-1.4976	155
15.0	-1.4232	157
15.5	-1.3495	159
16.0	-1.2758	161
16.5	-1.2027	162
17.0	-1.1296	164
17.5	-1.0571	166
18.0	-0.9845	168
18.5	-0.9122	170

RS	Ability	SS
19.0	-0.8399	171
19.5	-0.7675	173
20.0	-0.6952	175
20.5	-0.6225	177
21.0	-0.5497	179
21.5	-0.4762	181
22.0	-0.4026	182
22.5	-0.3278	184
23.0	-0.2530	186
23.5	-0.1764	188
24.0	-0.0998	190
24.5	-0.0209	192
25.0	0.0580	194
25.5	0.1399	197
26.0	0.2217	200*
26.5	0.3071	201
27.0	0.3925	203
27.5	0.4822	205
28.0	0.5718	208
28.5	0.6666	210
29.0	0.7614	212
29.5	0.8623	215
30.0	0.9633	217
30.5	1.0717	220
31.0	1.1801	222
31.5	1.2979	225
32.0	1.4156	227
32.5	1.5452	230
33.0	1.6748	232
33.5	1.8197	235
34.0	1.9646	237
34.5	2.1292	240
35.0	2.2938	243
35.5	2.4824	246
36.0	2.6710	250*
36.5	2.8848	252
37.0	3.0986	255
37.5	3.3311	259

RS	Ability	SS
38.0	3.5637	263
38.5	3.8031	266
39.0	4.0425	270
39.5	4.2801	274
40.0	4.5178	277
40.5	4.7485	280
41.0	4.9793	283
41.5	5.2009	286
42.0	5.4225	288
42.5	5.6541	291
43.0	5.8856	293
43.5	6.1830	295
44.0	6.4804	297
44.5	6.9014	300
45.0	7.3225	300
45.5	7.7827	300
46.0	8.2429	300
46.5	8.7765	300
47.0	9.3102	300
47.5	10.0143	300
48.0	10.7184	300

# ***Appendix E***

## ***Scale Scores with Frequencies***

## 2007 GEPA Language Arts Literacy Scale Scores with Frequencies

Scale Score	Number of Students	Percent of Students	Cumulative Number of Students	Cumulative Percent of Students	Scale Score	Number of Students	Percent of Students	Cumulative Number of Students	Cumulative Percent of Students
105	4	0	4	0	178	614	0.6	12,040	11.4
108	8	0	12	0	179	1,285	1.2	13,325	12.6
110	23	0	35	0	181	671	0.6	13,996	13.2
113	39	0	74	0.1	183	1,470	1.4	15,466	14.6
114	2	0	76	0.1	185	789	0.7	16,255	15.4
116	68	0.1	144	0.1	187	1,628	1.5	17,883	16.9
117	7	0	151	0.1	188	947	0.9	18,830	17.8
118	79	0.1	230	0.2	190	1,869	1.8	20,699	19.6
120	9	0	239	0.2	192	1,123	1.1	21,822	20.6
121	91	0.1	330	0.3	194	2,220	2.1	24,042	22.7
122	24	0	354	0.3	196	1,302	1.2	25,344	23.9
124	135	0.1	489	0.5	198	2,557	2.4	27,901	26.4
126	29	0	518	0.5	200	1,483	1.4	29,384	27.8
127	147	0.1	665	0.6	202	2,927	2.8	32,311	30.5
129	48	0	713	0.7	203	1	0	32,312	30.5
131	209	0.2	922	0.9	204	1,790	1.7	34,102	32.2
132	57	0.1	979	0.9	206	3,274	3.1	37,376	35.3
134	237	0.2	1,216	1.1	207	1	0	37,377	35.3
136	83	0.1	1,299	1.2	209	2,118	2.0	39,495	37.3
138	246	0.2	1,545	1.5	210	2	0	39,497	37.3
140	96	0.1	1,641	1.6	211	3,666	3.5	43,163	40.8
142	302	0.3	1,943	1.8	212	2	0	43,165	40.8
144	94	0.1	2,037	1.9	213	2,362	2.2	45,527	43.0
145	355	0.3	2,392	2.3	215	4,063	3.8	49,590	46.8
147	140	0.1	2,532	2.4	216	1	0	49,591	46.8
149	433	0.4	2,965	2.8	217	2,697	2.5	52,288	49.4
151	182	0.2	3,147	3.0	218	2	0	52,290	49.4
153	453	0.4	3,600	3.4	220	4,589	4.3	56,879	53.7
154	222	0.2	3,822	3.6	221	1	0	56,880	53.7
156	563	0.5	4,385	4.1	222	2,683	2.5	59,563	56.3
158	221	0.2	4,606	4.4	223	2	0	59,565	56.3
159	613	0.6	5,219	4.9	224	4,707	4.4	64,272	60.7
161	257	0.2	5,476	5.2	226	2,623	2.5	66,895	63.2
163	667	0.6	6,143	5.8	227	3	0	66,898	63.2
164	315	0.3	6,458	6.1	229	4,878	4.6	71,776	67.8
166	752	0.7	7,210	6.8	231	2,657	2.5	74,433	70.3
167	382	0.4	7,592	7.2	232	1	0	74,434	70.3
169	832	0.8	8,424	8.0	233	4,925	4.7	79,359	75.0
171	468	0.4	8,892	8.4	235	2	0	79,361	75.0
172	934	0.9	9,826	9.3	236	2,325	2.2	81,686	77.2
174	526	0.5	10,352	9.8	237	1	0	81,687	77.2
176	1,074	1.0	11,426	10.8	238	4,580	4.3	86,267	81.5

### 2007 GEPA Language Arts Literacy Scale Scores with Frequencies (continued)

Scale Score	Number of Students	Percent of Students	Cumulative Number of Students	Cumulative Percent of Students	Scale Score	Number of Students	Percent of Students	Cumulative Number of Students	Cumulative Percent of Students
239	3	0	86,270	81.5	273	377	0.4	105,087	99.3
241	2,024	1.9	88,294	83.4	276	238	0.2	105,325	99.5
243	4,022	3.8	92,316	87.2	279	203	0.2	105,528	99.7
246	1,604	1.5	93,920	88.7	281	107	0.1	105,635	99.8
247	1	0	93,921	88.7	284	101	0.1	105,736	99.9
250	3,115	2.9	97,036	91.7	286	38	0	105,774	99.9
251	1,334	1.3	98,370	92.9	288	42	0	105,816	100.0
254	2,168	2.0	100,538	95.0	290	7	0	105,823	100.0
255	1	0	100,539	95.0	291	18	0	105,841	100.0
257	953	0.9	101,492	95.9	293	10	0	105,851	100.0
260	1,414	1.3	102,906	97.2	295	6	0	105,857	100.0
264	626	0.6	103,532	97.8	297	1	0	105,858	100.0
267	769	0.7	104,301	98.5	299	6	0	105,864	100.0
270	409	0.4	104,710	98.9	300	1	0	105,865	100.0

**N-COUNT = 105,865 MEAN = 214.8857 STANDARD DEVIATION = 28.9677 SEM = 13.167**

## 2007 GEPA Mathematics Scale Scores with Frequencies

Scale Score	Number of Students	Percent of Students	Cumulative Number of Students	Cumulative Percent of Students	Scale Score	Number of Students	Percent of Students	Cumulative Number of Students	Cumulative Percent of Students
139	3	0	3	0	206	1	0	44,723	41.8
141	17	0	20	0	208	3,685	3.4	48,408	45.2
143	38	0	58	0.1	210	3	0	48,411	45.3
146	108	0.1	166	0.2	211	3,691	3.5	52,102	48.7
148	238	0.2	404	0.4	215	3,942	3.7	56,044	52.4
150	394	0.4	798	0.7	216	3	0	56,047	52.4
152	626	0.6	1,424	1.3	219	4,116	3.8	60,163	56.2
155	852	0.8	2,276	2.1	220	3	0	60,166	56.2
157	1,035	1.0	3,311	3.1	223	3,922	3.7	64,088	59.9
158	1	0	3,312	3.1	227	3,950	3.7	68,038	63.6
160	1,226	1.1	4,538	4.2	231	3,902	3.6	71,940	67.2
162	1,411	1.3	5,949	5.6	235	3,891	3.6	75,831	70.9
164	1,571	1.5	7,520	7.0	239	3,414	3.2	79,245	74.1
165	2	0	7,522	7.0	242	1	0	79,246	74.1
167	1,719	1.6	9,241	8.6	243	3,612	3.4	82,858	77.5
169	1,904	1.8	11,145	10.4	246	1	0	82,859	77.5
170	2	0	11,147	10.4	250	3,997	3.7	86,856	81.2
172	2,022	1.9	13,169	12.3	252	3,476	3.2	90,332	84.4
175	2,116	2.0	15,285	14.3	256	3,248	3.0	93,580	87.5
177	2,197	2.1	17,482	16.3	258	1	0	93,581	87.5
178	3	0	17,485	16.3	260	2,862	2.7	96,443	90.2
180	2,365	2.2	19,850	18.6	262	1	0	96,444	90.2
183	2,535	2.4	22,385	20.9	263	1	0	96,445	90.2
186	2,692	2.5	25,077	23.4	264	2,639	2.5	99,084	92.6
188	2,836	2.7	27,913	26.1	266	3	0	99,087	92.6
189	2	0	27,915	26.1	268	2,282	2.1	101,369	94.8
191	2,705	2.5	30,620	28.6	270	2	0	101,371	94.8
194	1	0	30,621	28.6	271	1,894	1.8	103,265	96.5
195	3,167	3.0	33,788	31.6	274	1,441	1.3	104,706	97.9
197	2	0	33,790	31.6	277	1,073	1.0	105,779	98.9
200	3,633	3.4	37,423	35.0	280	720	0.7	106,499	99.6
201	3,622	3.4	41,045	38.4	283	373	0.3	106,872	99.9
203	3	0	41,048	38.4	300	108	0.1	106,980	100.0
204	3,674	3.4	44,722	41.8					

**N-COUNT = 106,980 MEAN = 215.5038 STANDARD DEVIATION = 33.4964 SEM = 13.234**

## 2007 GEPA Science Scale Scores with Frequencies

Scale Score	Number of Students	Percent of Students	Cumulative Number of Students	Cumulative Percent of Students	Scale Score	Number of Students	Percent of Students	Cumulative Number of Students	Cumulative Percent of Students
129	1	0	1	0	211	3,399	3.2	38,604	36.1
132	4	0	5	0	212	1	0	38,605	36.1
136	2	0	7	0	214	3,381	3.2	41,986	39.3
140	7	0	14	0	217	3,528	3.3	45,514	42.6
143	30	0	44	0	220	3,510	3.3	49,024	45.9
144	1	0	45	0	223	3,476	3.3	52,500	49.1
146	60	0.1	105	0.1	226	3,519	3.3	56,019	52.4
150	117	0.1	222	0.2	228	3,646	3.4	59,665	55.8
153	200	0.2	422	0.4	229	1	0	59,666	55.8
154	1	0	423	0.4	231	3,549	3.3	63,215	59.1
157	332	0.3	755	0.7	232	1	0	63,216	59.1
158	1	0	756	0.7	234	3,628	3.4	66,844	62.5
160	532	0.5	1,288	1.2	235	3	0	66,847	62.5
163	741	0.7	2,029	1.9	237	3,594	3.4	70,441	65.9
166	1	0	2,030	1.9	240	3,566	3.3	74,007	69.2
167	974	0.9	3,004	2.8	242	3,149	2.9	77,156	72.2
169	5	0	3,009	2.8	244	2	0	77,158	72.2
170	1,259	1.2	4,268	4.0	245	3,419	3.2	80,577	75.4
172	2	0	4,270	4.0	246	1	0	80,578	75.4
173	1,480	1.4	5,750	5.4	248	1	0	80,579	75.4
176	2	0	5,752	5.4	250	3,643	3.4	84,222	78.8
177	1,759	1.6	7,511	7.0	251	3,301	3.1	87,523	81.9
180	1,970	1.8	9,481	8.9	252	2	0	87,525	81.9
183	2,318	2.2	11,799	11.0	253	2,960	2.8	90,485	84.6
186	2,478	2.3	14,277	13.4	254	3	0	90,488	84.6
187	3	0	14,280	13.4	256	2,772	2.6	93,260	87.2
189	2,662	2.5	16,942	15.8	259	2,624	2.5	95,884	89.7
190	6	0	16,948	15.9	262	2,321	2.2	98,205	91.9
192	2,715	2.5	19,663	18.4	264	2,044	1.9	100,249	93.8
194	5	0	19,668	18.4	267	1,822	1.7	102,071	95.5
195	2,907	2.7	22,575	21.1	269	1,545	1.4	103,616	96.9
197	1	0	22,576	21.1	271	2	0	103,618	96.9
200	3,197	3.0	25,773	24.1	272	1,181	1.1	104,799	98.0
201	2	0	25,775	24.1	275	898	0.8	105,697	98.9
202	3,125	2.9	28,900	27.0	276	1	0	105,698	98.9
204	3	0	28,903	27.0	278	588	0.5	106,286	99.4
205	3,108	2.9	32,011	29.9	280	358	0.3	106,644	99.7
207	5	0	32,016	29.9	282	170	0.2	106,814	99.9
208	3,183	3.0	35,199	32.9	286	85	0.1	106,899	100.0
210	6	0	35,205	32.9	300	14	0	106,913	100.0

**N-COUNT = 106,913 MEAN = 223.2722 STANDARD DEVIATION = 28.9512 SEM = 10.692**

# ***Appendix F***

## ***Reporting Limited English Proficiency (LEP) and Special Education (SE)***

**LIMITED ENGLISH PROFICIENT (LEP)**

A limited English proficient student is a student whose native language is one other than English. This student has sufficient difficulty speaking, reading, writing, or understanding the English language, as measured by an English language proficiency test, so as to be denied the opportunity to learn successfully in the classroom where the language of instruction is English.

School staff were instructed to mark a circle to designate the number of academic years each limited English proficient student participated in a language assistance program (Bilingual, English as a Second Language, or English Language Services) in **ANY** school in their **DISTRICT**. The codes for LEP are:

- < = LEP student **entered** a language assistance program **AFTER July 1, 2006**, and is **currently enrolled in the program**. *These students do not have to take the LAL portion of the test but MUST take Math and Science.*
- 1 = LEP student **entered** a language assistance program **BETWEEN July 1, 2005, and June 30, 2006**, and is **currently enrolled in the program**.
- 2 = LEP student **entered** a language assistance program **BETWEEN July 1, 2004, and June 30, 2005**, and is **currently enrolled in the program**.
- 3 = LEP student **entered** a language assistance program **BEFORE July 1, 2004**, and is **currently enrolled in the program**.
- F1 = **Former** LEP student **exited** a language assistance program **BETWEEN July 1, 2005, and the current administration date**, and is **NO longer enrolled in the program**.
- F2 = **Former** LEP student **exited** a language assistance program **BETWEEN July 1, 2004, and June 30, 2005** and is **NO longer enrolled in the program**.

**LIMITED-ENGLISH PROFICIENT (LEP) EXEMPT – LAL**

- E = LEP student **entered the United States** as well as a language assistance program **AFTER July 1, 2006 [currently enrolled in the program]**. *These students do not have to take the LAL portion of the test but MUST take Math and Science.*

**SPECIAL EDUCATION (SE)**

There are 13 codes for Special Education categories. The categories are:

- A. Auditorily Impaired
- B. Other Health Impaired
- C. Communication Impaired
- D. Emotionally Disturbed
- E. Cognitively Impaired
- F. Multiply Disabled
- G. Traumatic Brain Injury
- H. Orthopedically Impaired
- I. Specific Learning Disability
- J. Social Maladjustment
- K. Visually Impaired
- L. Speech-Language Services Only
- M. Autistic

For reporting, category N is used to indicate multiple grids. This is also a default code when a school failed to provide the specific information listed above for an APA student.

# ***References***

**REFERENCES**

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education (1974). *Standards for Educational and Psychological Tests*. Washington, DC: American Psychological Association.

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education (1985). *Standards for Educational and Psychological Testing*. Washington, DC: American Psychological Association.

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

Camilli, G. & Shepard, L.A. (1994). *Methods for Identifying Biased Test Items*. Thousand Oaks, CA: Sage.

Cizek, G.J. (1995). *Standard setting as Psychometric Due Process*. Paper presented at the Annual Meeting of the National Council on Measurement in Education, San Francisco.

Crocker, L. & Algina, J. (1986). *Introduction to Classical & Modern Test Theory*. Belmont, CA: Wadsworth Group.

*Cycle II Criterion-Based Holistic Scoring: A Mathematics and Science Handbook*. Developed by the New Jersey Department of Education, January 2005.

*Cycle II Criterion-Based Holistic Scoring: A Reading Handbook*. Developed by the New Jersey Department of Education, January 2004.

*Cycle II Criterion-Based Holistic Scoring: A Writing Handbook*. Developed by the New Jersey Department of Education, September 2007.

DeMauro, G.E. (1995). *Construct Validation of Minimum Competence*. Paper presented at the Annual Meeting of the National Council on Measurement in Education, San Francisco.

*Directory of Test Specifications and Sample Items for the Elementary School Proficiency Assessment (ESPA), Grade Eight Proficiency Assessment (GEPA), and High School Proficiency Assessment (HSPA) in Language Arts Literacy*. Developed by the New Jersey Department of Education, February 1998.

*Directory of Test Specifications and Sample Items for the Grade Eight Proficiency Assessment (GEPA) and the High School Proficiency Assessment (HSPA) in Mathematics*. Developed by the New Jersey Department of Education, February 1998.

*Directory of Test Specifications and Sample Items for the Grade Eight Proficiency Assessment (GEPA) and High School Proficiency Assessment (HSPA) in Science*. Developed by the New Jersey Department of Education, February 1998.

*Grade Eight Proficiency Assessment (GEPA) Standard Setting Study Report of Activities, 1999 Language Arts Literacy and Mathematics, and 2000 Science*. Developed for the New Jersey Department of Education by National Computer Systems, November 2000.

Holland, P.W. & Thayer, D.T. (1986). *Differential Item Functioning and the Mantel-Haenszel Procedure* (Technical Rep. No. 86-69). Princeton, NJ: Educational Testing Service.

- Jaeger, R.M. & Mills, C.N. (1998). *An integrated judgment procedure for setting standards on complex large-scale assessments*. Paper presented at that annual meeting of the American Educational Research Association, San Diego, CA.
- Kane, M. T. (2006). Validation. In R.L. Brennan (Ed.), *Educational Measurement* (4th ed., pp. 17-64). Westport, CT: American Council on Education and Praeger.
- Kolen, M.J. & Brennan, R.L. (2004). *Test Equating, Scaling and Linking*. New York: Springer Science and Business Media, Inc.
- Linacre, J. M. (2006) *A User's Guide to WINSTEPS MINISTEP Rasch-Model Computer Programs*. Chicago: Winsteps.com
- Linacre, J. M. (2006) *WINSTEPS Rasch measurement computer program*. Chicago: Winsteps.com
- Livingston, S.A. & Zieky, M.J. (1982) *Passing Scores*. Princeton, NJ: Educational Testing Service.
- Masters, G. N & Wright, B. (1997). The partial credit model. In W. J. van der Linden & R.K. Hambleton, (Eds.), *Handbook of Modern Item Response Theory*. New York. Springer-Verlag.
- Mehrens, W.A. & Lehmann, I.J. (1991) *Measurement and Evaluation in Education and Psychology* (4th ed.). New York: Holt, Rinehart and Winston.
- Messick, S. (1980). Test validity and the ethics of assessment. *American Psychologist*, 35, 1012-1027.
- Messick, S. (1995) Standards of validity and the validity of standards in performance assessment. *Educational Measurement: Issues and Practice* 14(4), 5-8.
- Millman, J. & Greene, J. (1989) The specification and development of tests of achievement and ability. In R. L. Linn (Ed.), *Educational Measurement* (3rd ed., pp. 335-366) New York: American Council on Education.
- New Jersey Core Curriculum Content Standards*. Developed by the New Jersey Department of Education, October 2004. Retrieved January 4, 2007, from <http://www.state.nj.us/njded/cccs/cccs.pdf>
- Score Interpretation Manual*. Developed by the New Jersey Department of Education, June 2007.
- Van Der Linden, W.J. & Hambleton, R.K. (Eds.)(1997). *Handbook of Modern Item Response Theory*. New York: Springer-Verlag.
- Webb, M.W. & Miller, E.R. (1995). *Standard Setting on Constructed-Response Items*. Paper presented at the Annual Meeting of the National Council on Measurement in Education, San Francisco.
- Zhao, H., Robinson, B., & Guo, F. (2006) *How well the New Jersey's grade eight proficiency assessment serves its designed purpose*. Paper presented at the Annual Meeting of the National Council on Measurement in Education, San Francisco.
- Zieky, M. (1993) Practical questions in the use of DIF statistics in test development. In P.W. Holland and H. Wainer (Eds.), *Differential Item Functioning*. Hillsdale, NJ: Lawrence Erlbaum.





PTM#1506.61

Copyright © 2007 New Jersey  
Department of Education

All rights reserved.