



DELAWARE

STUDENT TESTING PROGRAM

Frequently Asked Questions



REVISED — October 2008

ISBN 999-8752-51-5



9 789998 752511

1 2 3 4 5 6 7 8 9 10 11 12 A B C D E

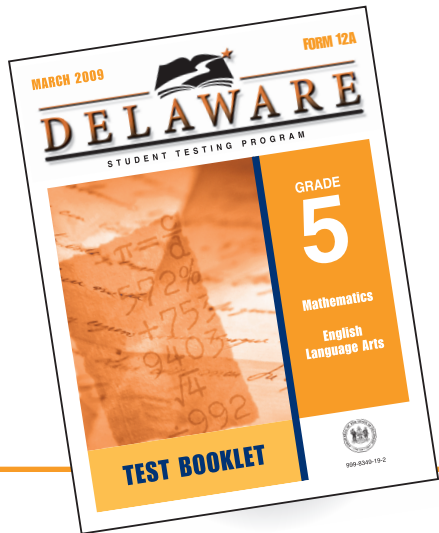
Copyright © 2009 by the State of Delaware, Department of Education.
Copyright © 2009 by NCS Pearson, Inc.

Users may reproduce this document for Delaware Student Testing Program purposes.

Printed in the United States of America.

Purpose of this brochure

This brochure is designed to provide general information regarding the Delaware Student Testing Program (DSTP). The questions included cover a broad spectrum and are representative of those often asked about the DSTP.





What is the purpose of the Delaware Student Testing Program (DSTP)?

- To serve as a measure of progress toward the Delaware Content Standards;
- To ensure students apply their academic skills to realistic, everyday problems;
- To promote better instruction and curriculum by providing timely reports of students' strengths and weaknesses;
- To serve as a primary indicator in the statewide accountability system.

How does the DSTP relate to the Delaware Content Standards?

- The DSTP is designed to measure how well students are learning the material identified by the Delaware Content Standards.
- Each question on the DSTP is designed to assess part of one of the Delaware Content Standards.
- The Delaware Content Standards identify the content expectations for students at the end of specific grade clusters. These content expectations explain what students should know and be able to do in each content area.

What types of questions are on the DSTP?

A combination of different types of questions is included on the test: multiple-choice, short-answer, and extended-response. Here are some samples of what DSTP items look like.

Sample Reading Multiple-Choice Question (1 point)

1

The word that best describes
Sally is —

- (A) carefree
- (B) hopeful
- (C) pleasant
- (D) creative

Sample Reading Short-Answer Question (2 points)

2

Write another good title for this story. Use information from the story to explain why your title is a good one.

Sample Mathematics Short-Answer Question and Scoring Rubric (2 points)

3

Retro Cab Company charges a base price of \$2.50 and 20¢ per mile. A competitor, **Rational Cab Company**, charges a base price of \$1.50 and 30¢ per mile. For what range of distances would Rational Cab Company be the cheaper choice? Show how you got your answer using tables, graphs, and/or equations.

Figure 1. Example of a Scoring Rubric for Item 3

Score Point	Description of Response
2	Response contains a correct answer with supporting evidence. Correctly identifies the range as < 10 miles and gives reasonable evidence.
1	Response contains a correct answer with weak supporting evidence. Correctly identifies the range but does not provide reasonable evidence, OR Provides evidence of understanding the procedure for finding the answer, but gives an incorrect response such as using the incorrect symbol (\leq) or contains an answer that is partially correct (≤ 9).
0	Response contains insufficient evidence of appropriate skills/knowledge to successfully accomplish the task.

How many points are possible on the test? *

Type of Question & Point Value	Reading	Mathematics	Science	Social Studies
Multiple-Choice (1 point each)	55 points (55 questions)	50	32	28
Short-Answer (2 points each)	16 points (8 questions)	16	36	40
Extended-Response (4 points each)	16 points (4 questions)	12**	0	0
Total Points	87 points	78	68	68

* Grade 2 is different. In reading, it has 30 multiple-choice questions, three short-answer questions, and two extended-response questions. In mathematics, it has 30 multiple-choice and 17 short-answer questions.

** There are no extended-response questions on grades 2 and 3 mathematics.

Writing

Type of Prompt	Writing (Point Value)
Stand-Alone Writing Prompt	10
Text-Based Writing Prompt	5
Total Points	15 points

How long is the test?

In March:

Test	Time	Number of Questions
Mathematics	two approximately 2½ hour sessions	Grade 2: 47 Grade 3: 70 Grades 4–10: 67
Reading	two approximately 2½ hour sessions	Grade 2: 35 Grades 3–10: 75
Writing	two sessions: Stand-Alone Prompt: 2½ hrs Text-Based Prompt: 1½ hrs	Grades 5, 8, and 10: 2

In May (grades 8 and 11) or in October (grades 4 and 6):

Test	Time	Number of Questions
Science	Grades 4 & 6: two approximately 1½ hour sessions Grades 8 & 11: approximately 2½ hours	60
Social Studies	Grades 4 & 6: two approximately 1½ hour sessions Grades 8 & 11: approximately 2½ hours	58

How is the DSTP developed?

■ Who writes the questions?

Most of the reading and mathematics questions are written by Delaware teachers, but some of them are written by the testing contractor. For grades 3, 5, 8, and 10, Delaware teachers write two-thirds of the test. For grades 2, 4, 6, 7, and 9, Delaware teachers review, re-write, and edit questions written by the testing contractor. The contractor writes 30 questions on the reading and mathematics tests, grades 2–10. All of the science and social studies questions are written or reviewed by Delaware teachers.

■ Are Delaware teachers involved in creating the test?

- Delaware teachers are involved in many parts of the development and scoring of the DSTP.
- They write, review, and edit test questions and find reading passages.
- They review test questions for possible bias before they are used.
- They review Delaware student responses collected in trial runs of the questions. They then choose **papers** that illustrate how Delaware wants the items to be scored.
- They review the Spanish translations of the Mathematics, Science, and Social Studies tests to make sure the tests are translated accurately and use language that is accessible to all Spanish-speaking students.
- They also participate in meetings to recommend where the performance levels should be set.

■ Who reviews the questions?

After a test question has been written, it is reviewed by several committees, content specialists, and editors. Each test question must be reviewed by the Bias committee, whose job it is to make sure nothing in the test is potentially offensive or unfair. The Content Advisory committee also reviews it for content accuracy. In addition, the testing contractor’s content specialists and editors review it for content validity, clarity, typographical errors, and grammar. If a test question survives all these reviews, it may be tried out with students—**“field tested.”**



■ Who decides what goes on the test?

In order to decide what goes on the test, a **blueprint** must be followed. The **blueprint** is the plan that tells how many multiple-choice, short-answer, and extended-response questions should be on the test. It also describes how many questions are needed for each Delaware Content Standard. Finally, it tells what kinds of reading passages should be on the test. The questions are chosen that will give the best possible estimates of students' progress towards learning the Delaware Content Standards.

■ Are the DSTP test questions tried out before they are used on the test?

A **live** test question is a question that counts towards a student's score. A **field test** question does not count towards a student's score. Before any question is used as a live question, it must first be tried out as a field test question. After the field test question is scored, the data for the question are reviewed. The data help determine how difficult the question is and whether it is able to distinguish between students who have learned the Content Standards and students who haven't. Finally, the data tell whether or not the question might be biased.

■ Are DSTP test items or sample/practice tests released to the public and to the school community?

The Department of Education knows how important it is to release test questions to the public and school community. There are several **Item Samplers** available on the DSTP website for each content area (English language arts, mathematics, science, and social studies). The Department of Education is currently working on releasing more questions each year along with sample student responses to show how the items are scored.

It is easy to see that writing quality test questions takes time and hard work. When test questions are released to the public, it means they cannot be used again, and others must be written to replace them.



How is the DSTP scored?

■ Who scores the test questions?

The multiple-choice questions are scored by machine. The written responses to open-ended questions are scored by human scorers, or readers, who are hired by the testing contractor. All readers who score the DSTP must have college degrees.

In order to be hired as a scorer, an applicant must provide a resumé; pass a reading, writing, and mathematics test; and be interviewed by the testing contractor. Acceptable candidates go on to **reader training**. In reader training, the applicants are trained to score the Delaware questions.



Not all applicants are hired to be readers. After reader training, each applicant must **qualify** before being hired as a reader. To qualify, a reader shows that he or she can score papers accurately and consistently, the way Delaware wants them to be scored. Readers are monitored during scoring to ensure that the scores given are consistent.

Readers learn how to use a **rubric** and they review **anchor papers**. A **rubric** is a set of rules for scoring a test question (see Figure 1 to see an example of a rubric). Each short-answer and extended-response test question has its own rubric. The **rubric** describes the characteristics of each possible score. The **anchor papers** are actual examples of Delaware students' work that were chosen as good examples for each score point.

■ How does the Department of Education know the tests are scored correctly and reliably?

The student responses are stored on a computer system. The scoring program shows the reader an image of a student response, which the reader must then score. The computer records the score that the reader assigns. To make sure that scoring is done reliably, several quality checks are in place. For example:

1. During scoring, papers that have already been scored by expert scorers are put into the system. If a reader scores many of these papers incorrectly, he or she will be removed from the scoring staff.

2. A percentage of the student responses are scored independently by a second reader who ensures agreement with the scores.
3. For writing, any score that is within one point of the total score that “meets the standard” will automatically be re-scored.
4. Any student response that a reader scores as “off-topic” will be double-checked to be sure that it really is “off-topic.”
5. If the two readers who score the stand-alone writing prompt are not within one point of each other, then the response will be scored by a third reader to determine what the total score should be.

■ What is the difference between a raw score and a scale score?

The **raw score** is the sum of the score points earned by a student. A multiple-choice question is worth 1 score point, a short-answer question is worth 2 score points, and an extended-response question is worth 4 score points. The raw scores range from 0 to about 80, depending on the test.

Scale scores make it possible to compare test scores from year to year. For example, because of scale scores, meaningful comparisons can be made of the achievement of this year’s fifth graders to that of last year’s fifth graders. Reading and mathematics scale scores range from about 150 to 800. The science and social studies scale scores range from about 100 to about 500.

■ What is the relationship of percentile rank to total score on DSTP?

The **percentile rank** gives information about how a student performed on the nationally normed portion of the test (30 questions on the DSTP Reading or the DSTP Mathematics test). For some grades, some of the questions on this portion of the test are excluded from the total DSTP score because they do not directly measure Delaware Content Standards.

If a student receives a percentile rank of 79, it means that he or she scored higher than 79% of students in a national sample of students in the same grade. Because the percentile rank is based only on 30 questions and the total DSTP score is based on all the questions, sometimes two students who have the same percentile rank will have different scale scores or raw scores.



■ How are the DSTP performance levels determined?

Several years ago, the Department of Education held a “Standard Setting” meeting. Committees of educators, parents, and community members gathered to review the tests and set the performance standards. (These are not the same as the content standards.) **Standard setting** is the process of choosing the “cut scores,” or minimum scores needed for each performance level for each test at each grade. The committees made recommendations to the State Board of Education, who approved the cut scores for each of the five performance levels (Well Below the Standard, Below the Standard, Meets the Standard, Exceeds the Standard, and Distinguished). In the summer of 2005, the cut scores for the performance levels for grades 2–10 in reading, writing, and mathematics were reviewed and updated—effective for the spring of 2006.

■ How long does it take to get scores back?

Test scores are available several weeks after the test is given. The reason it takes time to get the scores back is because the DSTP has short-answer and extended-response questions. These questions take longer to score because they are scored by people, not machines. It is important to include these types of questions on the test because students need to know how to explain concepts, in addition to being able to pick the correct answer out of a list.

■ How are the test results used by the schools?

Test results are used to help evaluate a school’s curriculum. Districts want to know which content areas of the curriculum need improvement and whether or not a curriculum is working. The DSTP is one of many pieces of evidence that districts can use to investigate these questions.

■ Sometimes a student's grades may be good but his or her DSTP scores are below the standard. How can this be explained?

Sometimes a student has good grades but does poorly on the DSTP. There are many possible reasons for this type of difference. Some of these are: (1) district/teacher grading policies may not be the same as DSTP performance standards; (2) it is possible that a school's curriculum may not wholly reflect the Delaware Content Standards; (3) it is possible that the assignments or tests that a student took during the school year may not have been as challenging as the questions on the DSTP; (4) it is possible that there may have been some non-school factors that affected a student's performance during testing. There are other possible reasons for these discrepancies, and they vary from student to student.



■ How does the Department of Education know how Delaware students are doing compared to students in other states?

There are at least two ways to find out how Delaware students compare to students in other states: (1) the percentile rank scores from the nationally normed part of the DSTP; and (2) the report of scores from a test called the National Assessment of Educational Progress (NAEP). This test is created by the U.S. Department of Education and has been given regularly since 1969. You can see Delaware's NAEP scores at <http://nces.ed.gov/nationsreportcard/>.

■ Who should I contact if I want to find out more about the DSTP?

You can find lots of information about the DSTP and Delaware's accountability system on the DSTP website at <http://www.doe.k12.de.us/aab>. If you have a specific question for which you cannot find an answer, please call the DSTP hotline at 1-877-838-DSTP.



Frequently Asked Questions

