Chapter 6

Data Analysis, Statistics, and Probability

Content Strand Description

Questions in this content strand assessed students' skills in collecting, organizing, reading, representing, and interpreting data. Also assessed were students' understanding of the basic elements of sampling, data analysis, and probability as well as their competence in calculating simple statistics and probabilities. Many questions required a constructed response and asked students to do a variety of tasks, such as completing or discussing charts and graphs or describing the best ways to collect or display data.

Students at grade 4 were expected to be familiar with a variety of types of graphs (typically pictorial), make predictions from data and explain their reasoning, and use the basic concept of chance. At grade 8, students were expected to analyze statistical claims and design experiments, demonstrate some understanding of sampling, and be able to make predictions based on complex data. Students at grade 12 were expected to use a wide variety of statistical techniques to model situations and solve problems. They also were expected to understand and apply concepts of probability to dependent and independent events and to have some knowledge of conditional probability.

Examples of Individual Questions and Student Performance

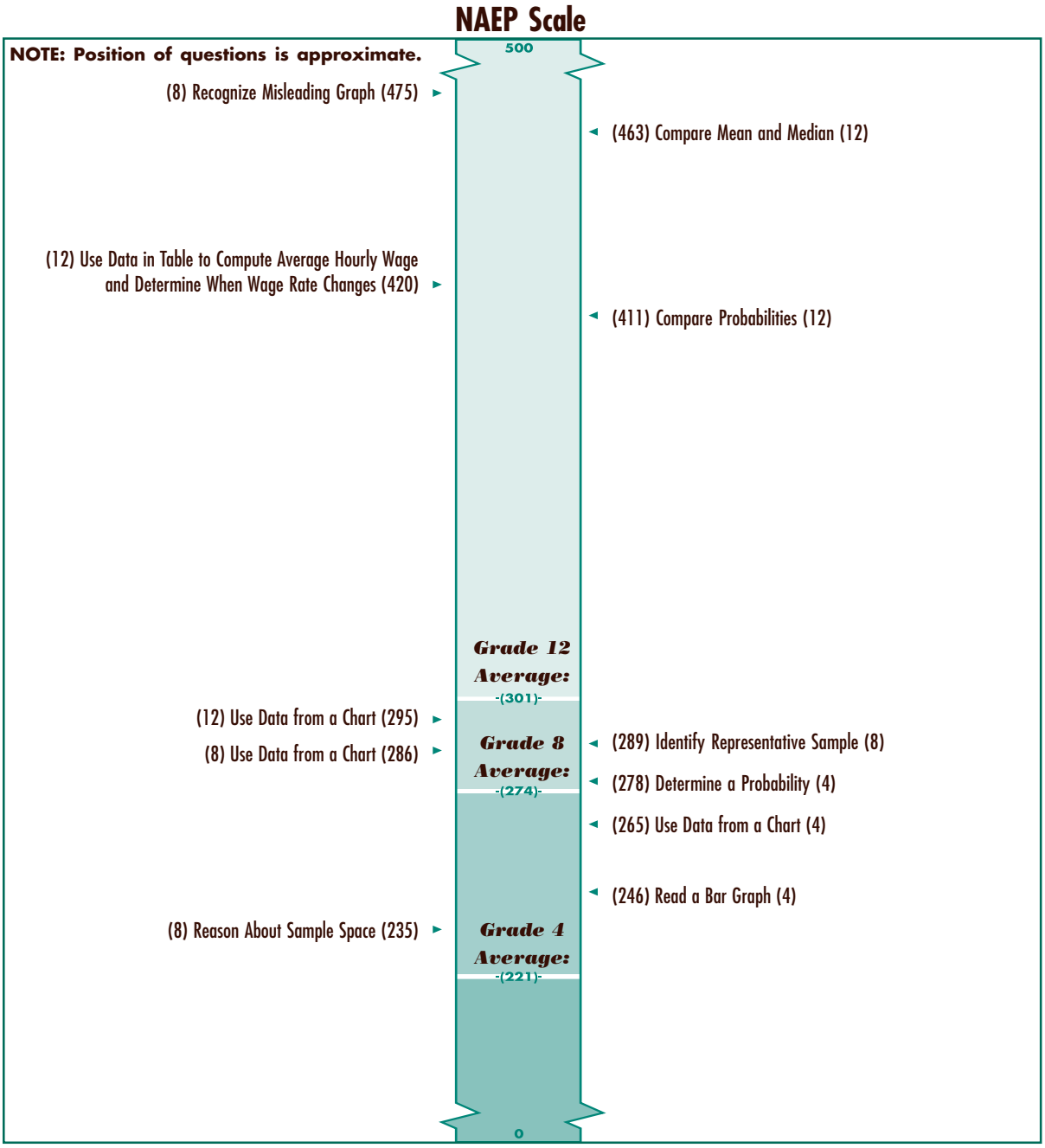
A number of the Data Analysis, Statistics, and Probability questions from the NAEP 1996 mathematics assessment are shown in this chapter. Presentation of the questions is organized around three areas of emphasis. *Tables, graphs, and charts* includes questions that assessed students' abilities to interpret and display data; *sampling and statistics* includes questions that assessed students' knowledge and skills in these areas; and *probability* includes questions that assessed students' understanding of and ability to calculate the probability of simple and related events.

All sample questions from this content strand are mapped onto the NAEP mathematics scale as shown in Figure 6.1. Specific instructions on how to interpret this map are given at the end of Chapter 2. The map is included to provide an indication of the relative difficulty of each example question and, thus, to indicate the type of material mastered within this content strand by students with varying degrees of mathematics proficiency. As noted in previous chapters, however, the difficulty of any question is a function of the relationship between characteristics specific to the question (e.g., format, absence or presence of graphics, real-world application), the specific mathematics content associated with the question, and students' opportunities to learn this content. It should be remembered also that overall performance on the Data Analysis, Statistics, and Probability content strand is not determined solely by performance on the examples presented here. These examples illustrate only some of what students know and can do.

Figure 6.1

Map of Selected Data Analysis, Statistics, and Probability Questions on the NAEP Composite Mathematics Scale (Item Map)

THE NATION'S REPORT CARD



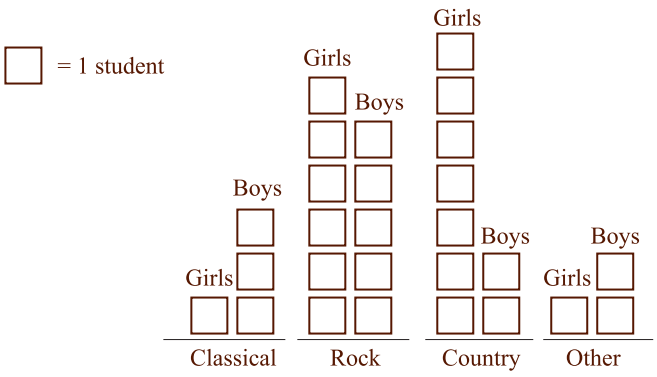
NOTE: Each mathematics question was mapped onto the NAEP 0 to 500 mathematics scale. The position of the question on the scale represents the scale score obtained by students who had a 65 percent probability of successfully answering the question. (The probability was 74 percent for a 4-option multiple-choice question and 72 percent for a 5-option multiple-choice question.) Only selected questions are presented. The number 4, 8, or 12 in parentheses is the grade level at which the question was asked.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 1996 Mathematics Assessment.

Tables, graphs, and charts

These questions assessed students' abilities to interpret and display data in tables, graphs, and charts. At all grade levels, students had to read and interpret data, make predictions, compute with data, and interpolate and extrapolate. They also had to translate data into tables and graphs. Questions for fourth-grade students often used pictographs, with symbols representing single or multiple units. Fourth-grade students also were evaluated on their ability to interpret simple pie charts. Questions for older students included stem-and-leaf and box-and-whisker plots. Graphs and charts often involved percents, and graphs often compared units on two dimensions. Students in eighth and twelfth grade were asked to make decisions about the best representation of data for certain situations or to compare data in two different tables, graphs, or charts.

Four examples of questions are presented here — one at each grade level and one that appeared at all three grade levels. The first example is a multiple-choice question that appeared on the assessment for fourth-grade students. The question presented students with a bar graph representing class votes on favorite types of music. Results for three types of music and a residual "other" category were displayed separately for boys and girls. A legend indicated that the square symbol used in the graph represented one student. Students were to determine the type of music preferred by most of the students in the class. In order to respond correctly, students had to add the number of votes for boys and girls together within categories and compare the totals.

4. Each boy and girl in the class voted for his or her favorite kind of music. Here are the results.



Which kind of music did most students in the class prefer?

- (A) Classical
- (B) Rock
- (C) Country
- (D) Other

Did you use a calculator on this question?

Yes No

The correct option is B.

This question was not very difficult for fourth-grade students. It mapped at a score of 246 on the NAEP composite mathematics scale. Student performance data are presented in Tables 6.1 and 6.2. Nearly 60 percent of the students responded correctly to the question. Another 36 percent of the students chose Option C (country music) as the appropriate response. These students may not have understood that they had to sum the data for girls and boys and may have simply chosen the category with the longest bar. Table 6.2 shows that approximately two-thirds of the students at the *Basic* achievement level and more than 80 percent of those at the *Proficient* level responded correctly to the question.

Table 6.1 ***Percentage Correct for "Read a Bar Graph"***

THE NATION'S REPORT CARD

	Percentage Correct
Grade 4	
Overall	59
Males	61
Females	57
White	67
Black	33
Hispanic	45
Asian/Pacific Islander	***
American Indian	***

*** Sample size is insufficient to permit a reliable estimate.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 1996 Mathematics Assessment.

Table 6.2 ***Percentage Correct Within Achievement-Level Intervals for "Read a Bar Graph"***

THE NATION'S REPORT CARD

	NAEP Grade 4 Composite Scale Range			
Overall	Below Basic	Basic	Proficient	Advanced
59	38	66	82	***

*** Sample size is insufficient to permit a reliable estimate.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 1996 Mathematics Assessment.

The second example for this area was a question that appeared on the assessment for all three grade levels. It is a short constructed-response question for which students had to interpret data from a table and then explain their interpretation. The data in this question again represented votes, this time regarding shapes that were being considered for a class symbol. (The question fell within a block for which students were supplied with cardboard shapes or manipulatives. The designations *N*, *P*, and *Q* that are used in the question refer to these shapes.) Based on the preference data from three classes, students were to determine which shape should be selected for the symbol and tell why. The correct response was shape *N* because it received more total votes than the other two shapes; students also could have stated that it was the first choice in one class and the second choice in the others.

5. This question refers to pieces *N*, *P*, and *Q*.

In Mr. Bell's classes, the students voted for their favorite shape for a symbol. Here are the results.

	Class 1	Class 2	Class 3
Shape *N*	9	14	11
Shape *P*	1	9	17
Shape *Q*	22	7	2

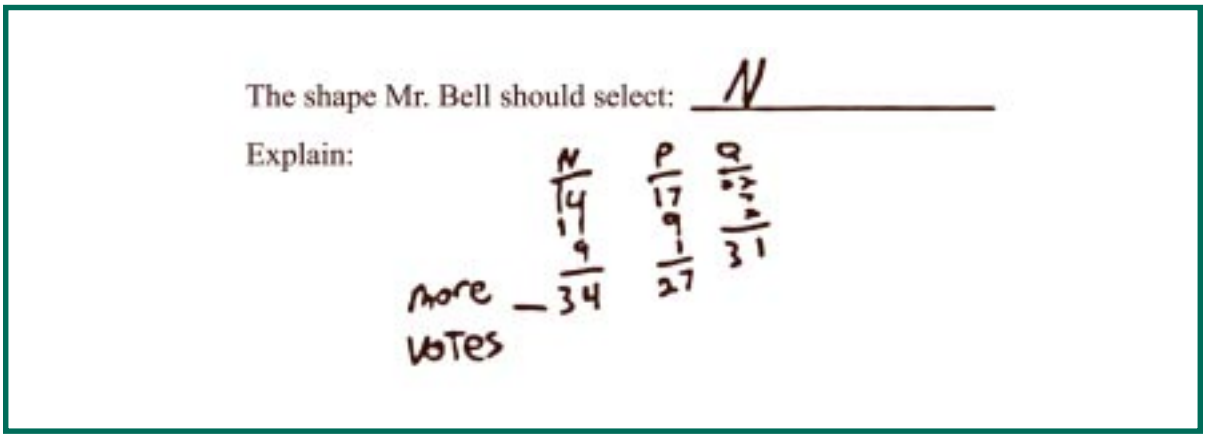
Using the information in the chart, Mr. Bell must select one of the shapes to be the symbol. Which one should he select and why?

The shape Mr. Bell should select: _____

Explain:

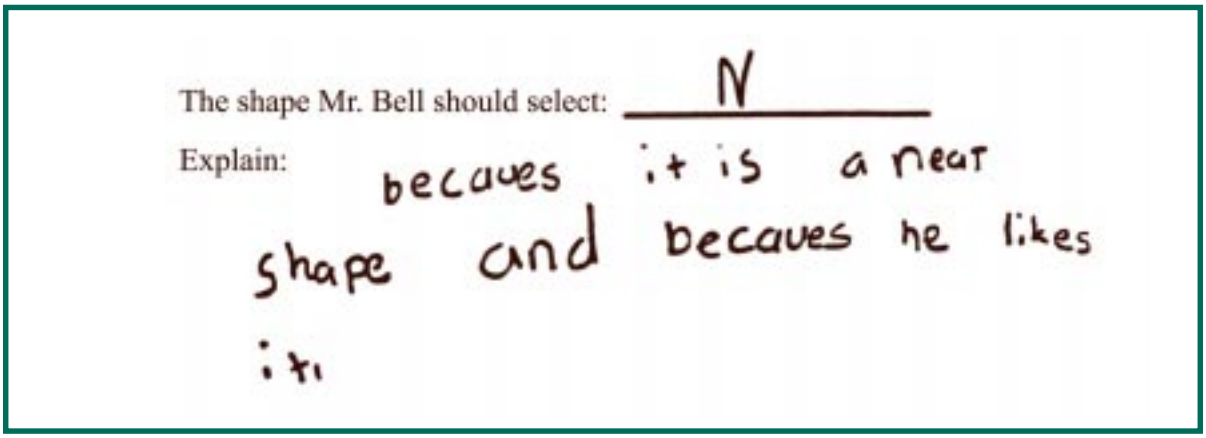
A sample "correct" response follows. In this response, the student chose shape *N*, supporting this choice by adding up the total number of votes for each shape and showing that shape *N* received the most votes overall.

Sample "correct" response

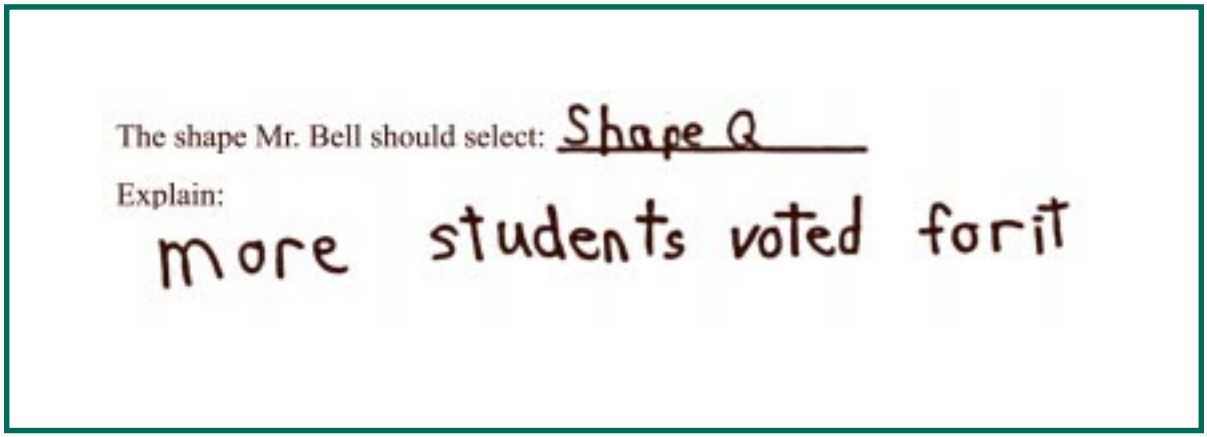


These next two samples are "incorrect" responses. In the first, the student correctly chose shape *N* but provided an incorrect explanation. It is followed by a sample response from a student who chose shape *Q*.

Sample "incorrect" response 1



Sample "incorrect" response 2



This question was somewhat difficult for students in grade 4 but easier for students in grades 8 and 12. The performance results are displayed in Tables 6.3 and 6.4. Table 6.3 shows the percentage of students at each grade who 1) chose shape *N* and had a correct explanation, 2) who chose shape *N* but had no or an incorrect explanation, 3) who chose shape *Q*, and 4) who made some other incorrect response.[1] Only the responses of students who chose shape *N* and had a correct explanation were rated "correct."

Approximately one-third of the fourth-grade students, one-half of the eighth-grade students, and two-thirds of the twelfth-grade students chose shape *N* for the symbol and had correct explanations. At each of the three grades, the percentage of students who chose shape *N* but had no or incorrect explanations was between 12 and 17 percent. Perhaps the most interesting difference was in the percent of students who chose shape *Q*. Thirty-two percent of the fourth-grade students (equivalent to the percentage who answered correctly) chose shape *Q*. At grade 8, this percentage dropped by half, and at grade 12, only nine percent of the students chose shape *Q*. At the earlier grades, students may be more inclined than at later grades simply to base their response on the largest single number in the table rather than to sum the data across classes. Another possible explanation is that, at the fourth-grade level, students simply answered their favorite shape.

[1] Student responses for this and all other constructed-response questions also could have been scored as "off task," which means that the student provided a response, but it was deemed not related in content to the question asked. There are many examples of these types of responses, but a simple one would be "I don't like this test." Responses of this sort could not be rated. In contrast, responses scored as "incorrect" were valid attempts to answer the question that were simply wrong.

Table 6.3 *Score Percentages for "Use Data from a Chart"*

THE NATION'S REPORT CARD

	Correct	Incorrect			Omit
Grade 4	Shape N–Correct Explanation	Shape N–No, or Incorrect, Explanation	Shape Q	Other	
Overall	32	12	32	21	3
Males	31	12	30	23	3
Females	33	12	33	18	3
White	38	12	32	14	3
Black	13	11	34	39	3
Hispanic	16	13	25	39	6
Asian/Pacific Islander	33	18	31	16	2!
American Indian	***	***	***	***	***
Grade 8					
Overall	58	15	16	10	0
Males	57	16	16	11	0
Females	60	14	16	10	0!
White	64	15	13	7	0!
Black	38	20	21	20	2
Hispanic	52	9	20	19	0!
Asian/Pacific Islander	– –	– –	– –	– –	– –
American Indian	***	***	***	***	***
Mathematics Course Taking:					
Eighth-Grade Mathematics	56	16	19	10	0
Pre-Algebra	57	14	17	11	0
Algebra	67	15	9	10	0
Grade 12					
Overall	67	17	9	6	1
Males	66	20	7	6	2
Females	67	16	10	6	1
White	70	18	7	4	1
Black	58	15	16	8	2
Hispanic	55	21	9	12	4
Asian/Pacific Islander	67	14	10	7	2
American Indian	***	***	***	***	***
Geometry Taken	68	18	8	5	1
Highest Algebra-Calculus Course Taken:					
Pre-Algebra	65	10	11	9	3
First-Year Algebra	61	20	10	6	2
Second-Year Algebra	69	18	8	5	0
Third-Year Algebra/Pre-Calculus	71	17	6	5	1
Calculus	75	10	10	4	0

NOTE: Row percentages may not total 100 due to rounding. Responses that could not be rated were excluded.

*** Sample size is insufficient to permit a reliable estimate.

– – Data for grade 8 Asian/Pacific Islanders are not reported due to concerns about the accuracy and precision of the national estimates. See Appendix A for further detail.

! Statistical tests involving this value should be interpreted with caution. Standard error estimates may not be accurately determined and/or the sampling distribution of the statistics does not match statistical test assumptions (see Appendix A).

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 1996 Mathematics Assessment.

Table 6.4 shows, for each grade, the percentage of students within each of the NAEP achievement levels who responded correctly to the question. At grade 4, only students who were at least at the *Proficient* level had a greater than 50 percent chance of answering correctly. At grade 8, more than two-thirds of students at the *Basic* level provided "correct" responses, and at grade 12, even students below the *Basic* level had a 50 percent chance of correct response. The question mapped at a score of 265 on the NAEP composite mathematics scale for grade 4 students. At grade 8, the question mapped at 286, and at grade 12, the question mapped at 295.

Table 6.4 ***Percentage Correct Within Achievement-Level Intervals for "Use Data from a Chart"***

THE NATION'S REPORT CARD

		NAEP Grades 4, 8, and 12 Composite Scale Ranges			
	Overall	Below Basic	Basic	Proficient	Advanced
Grade 4	32	8	32	61	***
Grade 8	58	38	68	77	***
Grade 12	67	51	73	77	***

*** Sample size is insufficient to permit a reliable estimate.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 1996 Mathematics Assessment.

The third sample for this area is an extended constructed-response question for eighth-grade students. Students were presented with two graphs displaying the number of riders for the Metro Rail Company over a 6-month period. The difference between the graphs was that one displayed the scale for the number of riders in increments of 2,000 while the other displayed the scale in increments of 100. The question itself had two components, although these were scored together to provide a single rating for the question. First, students were instructed to choose, and justify their choice for, the graph that would best convince others that the Metro Rail Company made a lot more money from ticket sales in March than in October. Second, students were asked to explain why some people might consider the graph they chose to be misleading.

This question requires you to show your work and explain your reasoning. You may use drawings, words, and numbers in your explanation. Your answer should be clear enough so that another person could read it and understand your thinking. It is important that you show all of your work.

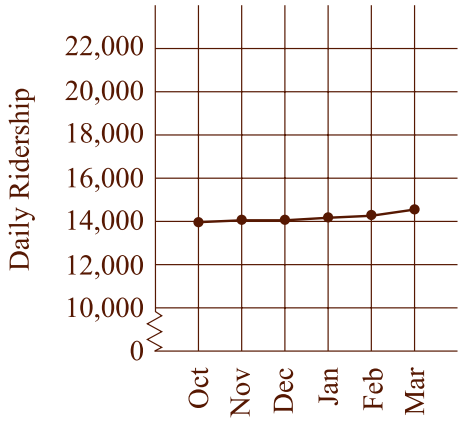
9.

METRO RAIL COMPANY

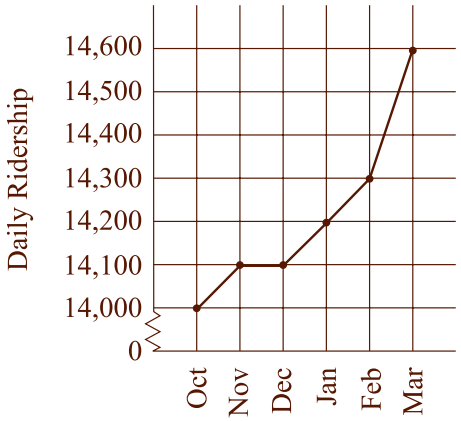
Month	Daily Ridership
October	14,000
November	14,100
December	14,100
January	14,200
February	14,300
March	14,600

The data in the table above has been correctly represented by both graphs shown below.

Graph A



Graph B



Which graph would be best to help convince others that the Metro Rail Company made a lot more money from ticket sales in March than in October?

Explain your reason for making this selection.

Why might people who thought that there was little difference between October and March ticket sales consider the graph you chose to be misleading?

Did you use a calculator on this question?

◯ Yes ◯ No

The correct answer for the first part of the question was Graph B because it appeared to show a large increase from October to March. Acceptable variations on this explanation included:

- The line in Graph B goes up more, has a more dramatic rise, or climbs higher.
- Graph B climbs faster.
- Graph B is steeper.
- Graph B shows a larger visual increase.

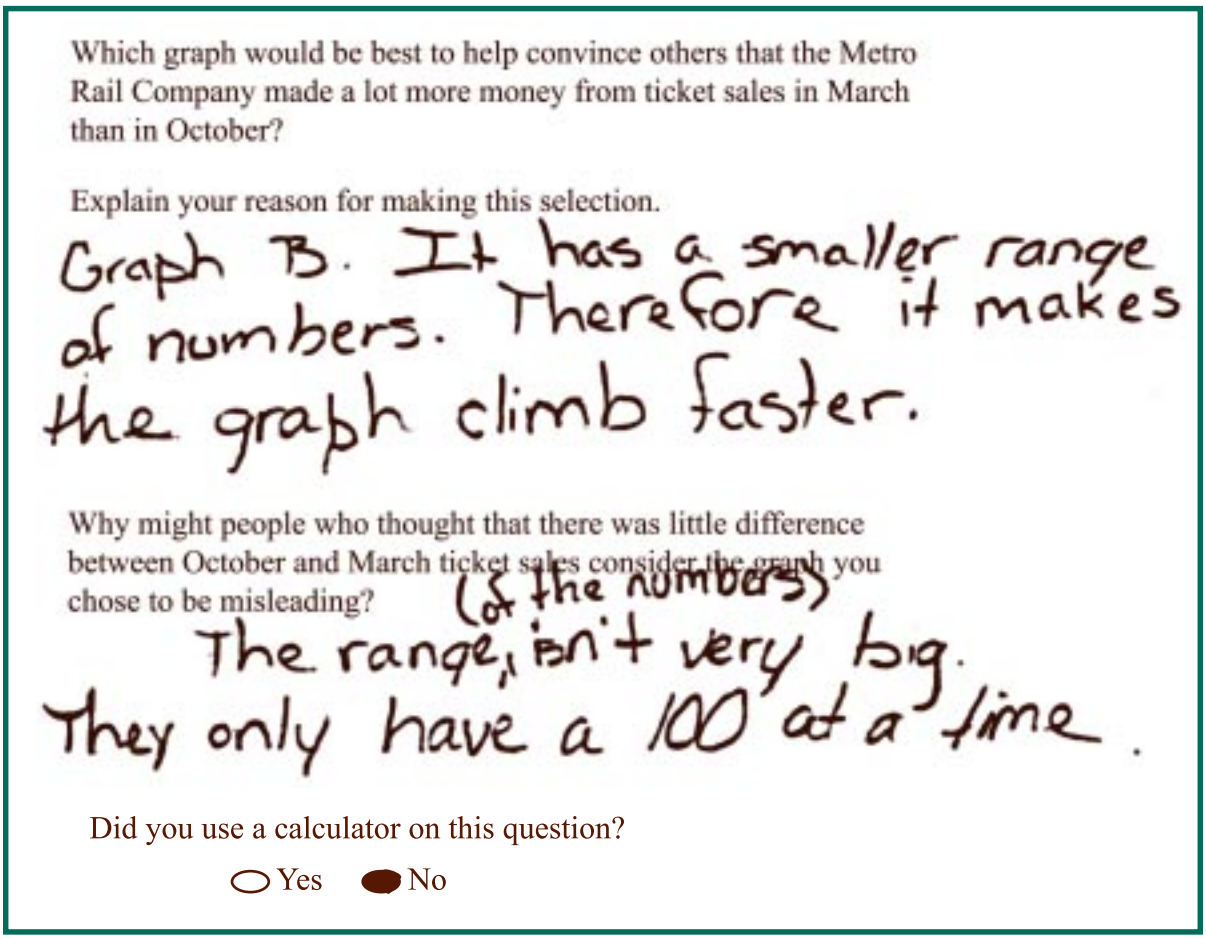
For the second part of the question, students were expected to recognize that Graph B might be considered misleading because it exaggerated a relatively small increase in ridership (misuse of scale). Acceptable variations of this reason included:

- Graph B has a smaller scale.
- Graph A has a larger scale.
- The numbers on B are smaller than those on A (they increase by 100s not 1,000s).

The use of the term "range" instead of "scale" was not considered acceptable.

Student responses were considered "correct" if they identified B as the best graph and had a complete explanation for both parts of the question. An example of a "correct" response follows. In this example, the student used the term "range" in both answers, which would not, in itself, be considered "correct"; however, the student also said that Graph B should be chosen because it climbs faster and is misleading because it only has 100 at a time. Both of these statements were considered "correct" responses.

Sample "correct" response



Which graph would be best to help convince others that the Metro Rail Company made a lot more money from ticket sales in March than in October?

Explain your reason for making this selection.

Graph B. It has a smaller range of numbers. Therefore it makes the graph climb faster.

Why might people who thought that there was little difference between October and March ticket sales consider the graph you chose to be misleading?

The range, (of the numbers) isn't very big. They only have a 100 at a time.

Did you use a calculator on this question?

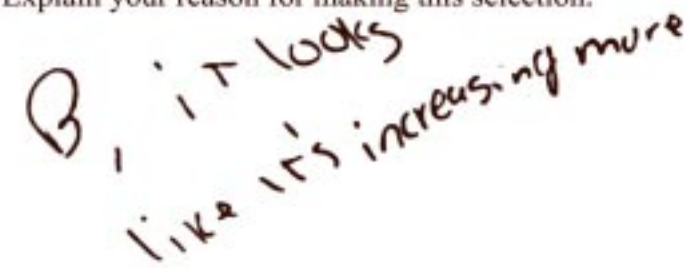
○ Yes ● No

Student responses were rated as "partial" if they chose Graph B and had an incomplete, but partially correct, explanation for one or both parts of the question. For example, in the following sample response the student gave a complete explanation of why B should be chosen, but an incomplete explanation of why Graph B could be considered misleading. When the question was anchored to the NAEP scale, the "correct" and "partial" rating categories were collapsed.

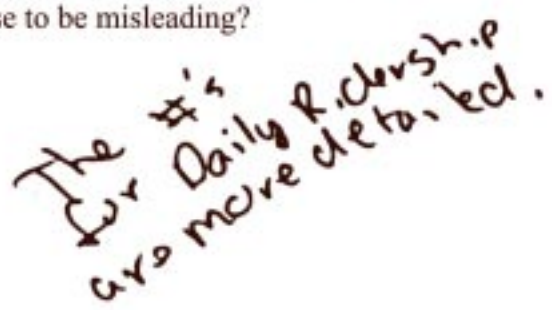
Sample "partial" response

Which graph would be best to help convince others that the Metro Rail Company made a lot more money from ticket sales in March than in October?

Explain your reason for making this selection.



Why might people who thought that there was little difference between October and March ticket sales consider the graph you chose to be misleading?



Did you use a calculator on this question?

○ Yes ● No

Students also could have had responses that identified Graph B as the best graph but offered no explanations or only incorrect explanations. Answers of this type, which are illustrated in the following example, were rated as "minimal."

Sample "minimal" response

Which graph would be best to help convince others that the Metro Rail Company made a lot more money from ticket sales in March than in October?

Explain your reason for making this selection.

Graph B because it is more clear to read.

Why might people who thought that there was little difference between October and March ticket sales consider the graph you chose to be misleading?

Because they might not be able to read the graph I have choosen

Did you use a calculator on this question?

○ Yes ● No

Student responses, such as the following, that did not identify B as the best graph were rated "incorrect."

Sample "incorrect" response

Which graph would be best to help convince others that the Metro Rail Company made a lot more money from ticket sales in March than in October?

Explain your reason for making this selection.

graph A because the dot is very close to 15,000 and in graph B it is in 14,600.

Why might people who thought that there was little difference between October and March ticket sales consider the graph you chose to be misleading?

Because the one I chose doesn't look like much difference because it has a longer range of amounts

Did you use a calculator on this question?

○ Yes ● No

Table 6.5 shows that only 2 percent of the students chose Graph B and gave complete explanations, while 19 percent chose Graph B and gave incomplete, but partially correct, explanations for at least one part of the question. However, 35 percent of the eighth-grade students who were taking algebra were able to provide at least partially correct explanations for choosing Graph B. This was a higher percentage than was obtained for students enrolled in pre-algebra or eighth-grade mathematics. Overall, 30 percent of the students did not even attempt the question.

Table 6.5 ***Score Percentages for "Recognize Misleading Graph"***

THE NATION'S REPORT CARD

	Correct Graph B–Complete Explanation	Partial Graph B–Incomplete but Partially Correct Explanation	Minimal Graph B–No or Incorrect Explanation	Incorrect	Omit
Grade 8					
Overall	2	19	34	14	30
Males	2	19	32	15	31
Females	2	18	36	14	30
White	2	23	34	11	28
Black	2!	6	38	21	33
Hispanic	0!	7	24	25	43
Asian/Pacific Islander	– –	– –	– –	– –	– –
American Indian	***	***	***	***	***
Mathematics Course Taking:					
Eighth-Grade Mathematics	1	16	32	19	32
Pre-Algebra	3	13	38	15	30
Algebra	3	32	31	6	28

NOTE: Row percentages may not total to 100 due to rounding. Responses that could not be rated were excluded.
*** Sample size is insufficient to permit a reliable estimate.
–– Data for grade 8 Asian/Pacific Islanders are not reported due to concerns about the accuracy and precision of the national estimates. See Appendix A for further detail.
! Statistical tests involving this value should be interpreted with caution. Standard error estimates may not be accurately determined and/or the sampling distribution of the statistics does not match statistical test assumptions (see Appendix A).
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 1996 Mathematics Assessment.

Table 6.6 shows that 35 percent of the eighth-grade students who were classified as *Proficient* on the NAEP composite mathematics scale, 22 percent of the students classified as *Basic*, and only 7 percent of the students classified as performing below *Basic* chose Graph B and responded with at least partially correct explanations to the two parts of the question. The question mapped at a score of 475.

Table 6.6 ***Percentage at Least Partial Within Achievement-Level Intervals for "Recognize Misleading Graph"***

THE NATION'S REPORT CARD

	NAEP Grade 8 Composite Scale Range			
Overall	Below Basic	Basic	Proficient	Advanced
20	7	22	35	***

*** Sample size is insufficient to permit a reliable estimate.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 1996 Mathematics Assessment.

The final example for this area is a short constructed-response question that was used at grade 12. The question presented a table summarizing time of day and number of hours worked, average hourly wage, and daily earnings for an individual on each of 5 days. For the fourth day, the cells for average hourly wage and total earnings were left blank. As in the previous example, this question had two parts that were considered together in determining the student's score. In the first part, students were given the total earnings for all 5 days and asked to use this information, in conjunction with the table, to determine the average hourly wage for day 4. To respond correctly, students had to add the daily earnings for the 4 days presented ($119.00), subtract this from the total earnings of $153.50, and divide by the total number of hours worked on the fourth day. This yielded the correct answer of $5.75. For the second part of the question, students were to use the information on time of day and number of hours worked, along with the average hourly rate, to determine the time of day at which the hourly rate changed. The correct answer was 5:00 p.m.

TIME CARD Name: J. Jasmine	Number of Hours	Average Hourly Wage	Total Daily Earnings
Mon. 10:00 a.m. — 3:00 p.m.	5	5.50	27.50
Tue. 9:00 a.m. — 4:00 p.m.	7	5.50	38.50
Wed. 3:00 p.m. — 7:00 p.m.	4	5.75	23.00
Thur. 2:00 p.m. — 8:00 p.m.	6		
Fri. 5:00 p.m. — 10:00 p.m.	5	6.00	30.00

2. According to the information above, what is the average hourly wage for Thursday's earnings if the total earnings for five days was $153.50?

Answer:_____

The hourly wage rate changes at some hour during the day. At what time does the hourly wage rate change?

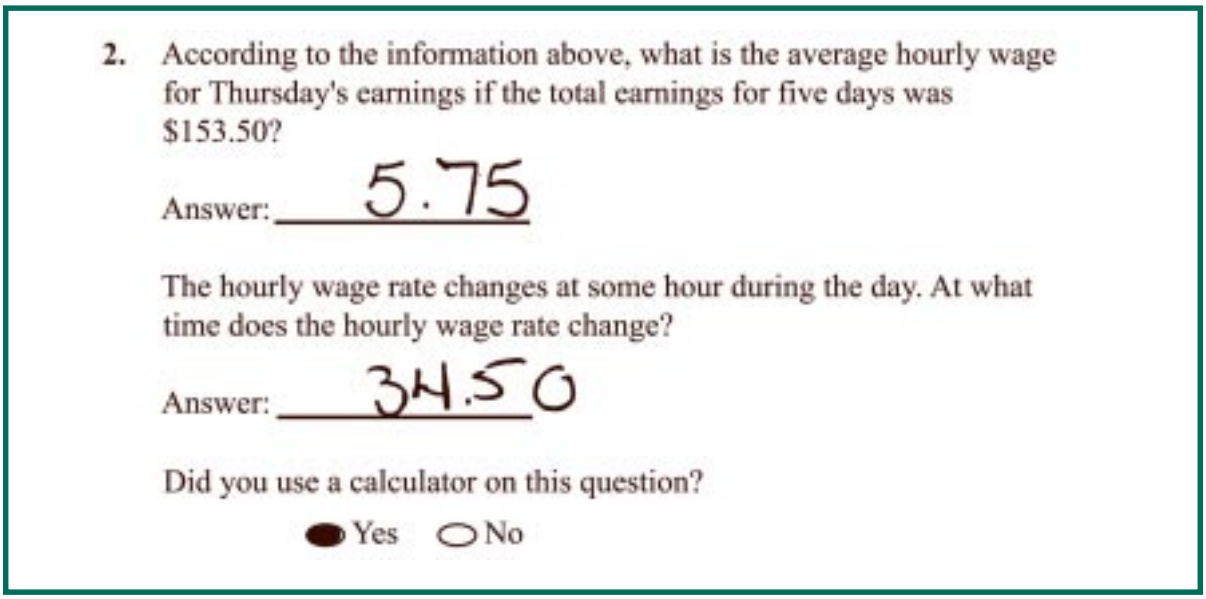
Answer:_____

Did you use a calculator on this question?

◯ Yes ◯ No

Student responses were rated "correct," "partial," or "incorrect." "Correct" responses identified both the correct hourly wage and the correct time of the rate change. "Partial" responses identified either the correct average hourly wage or the correct time of the rate change, and "incorrect" responses did not correctly identify either. Following are three sample responses. The first two responses were rated "partial"; each student correctly computed the average hourly wage for the fourth day, but in the second part of the question the first respondent entered the total daily earnings for day 4 instead of the time of the rate change, and the second respondent entered 2:00 (the time the individual started work on day 4).

Sample "partial" response 1



2. According to the information above, what is the average hourly wage for Thursday's earnings if the total earnings for five days was $153.50?

Answer: 5.75

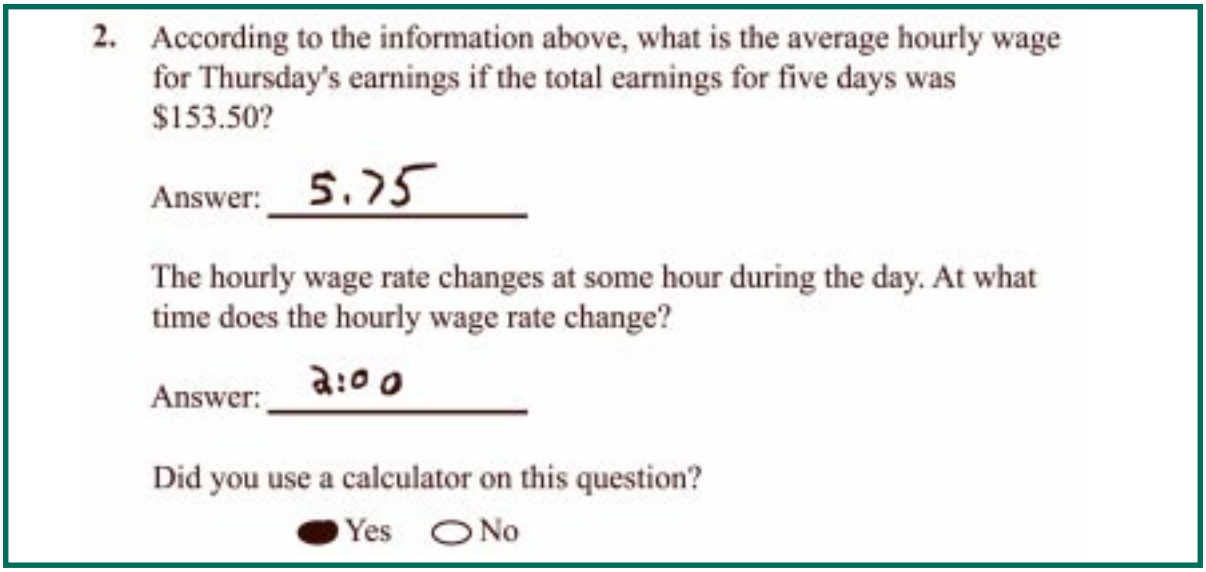
The hourly wage rate changes at some hour during the day. At what time does the hourly wage rate change?

Answer: 34.50

Did you use a calculator on this question?

● Yes ○ No

Sample "partial" response 2



2. According to the information above, what is the average hourly wage for Thursday's earnings if the total earnings for five days was $153.50?

Answer: 5.75

The hourly wage rate changes at some hour during the day. At what time does the hourly wage rate change?

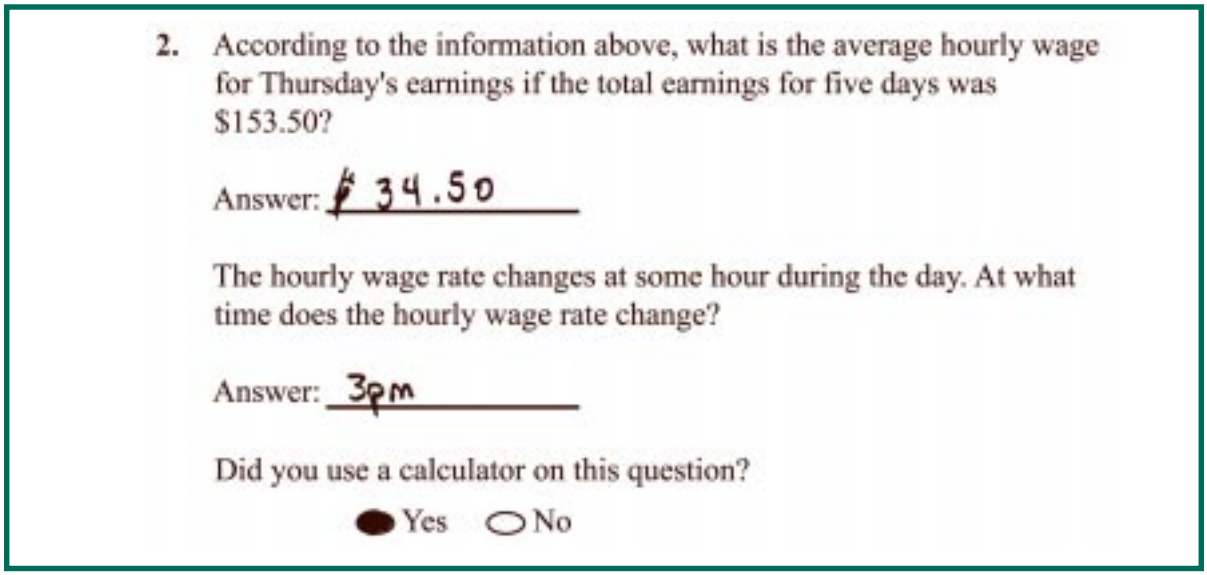
Answer: 2:00

Did you use a calculator on this question?

● Yes ○ No

The third sample response was rated "incorrect." This student entered the total daily earnings for the fourth day as a response to the first question and identified 3:00 p.m. as the time of the rate change.

Sample "incorrect" response



2. According to the information above, what is the average hourly wage for Thursday's earnings if the total earnings for five days was $153.50?

Answer: $ 34.50

The hourly wage rate changes at some hour during the day. At what time does the hourly wage rate change?

Answer: 3pm

Did you use a calculator on this question?

● Yes ○ No

Performance data are presented in Tables 6.7 and 6.8. Thirteen percent of students answered both parts of the question correctly, whereas 43 percent responded correctly to one of the two parts. Students who had taken calculus were more likely to respond correctly than students who had not taken calculus, and students who had taken at least third-year algebra or pre-calculus were more likely than those who had taken less mathematics to provide at least a partially correct response.

Table 6.7 ***Score Percentages for "Use Data in Table to Compute Average Hourly Wage and Determine When Wage Rate Changes"***

THE NATION'S REPORT CARD

	Correct	Partial	Incorrect	Omit
Grade 12				
Overall	13	43	40	3
Males	13	43	40	3
Females	13	43	40	3
White	16	46	36	2
Black	4	38	51	7
Hispanic	11	36	46	7
Asian/Pacific Islander	15	44	37	4
American Indian	***	***	***	***
Geometry Taken	14	46	36	3
Highest Algebra-Calculus Course Taken:				
Pre-Algebra	***	***	***	***
First-Year Algebra	8	36	51	4
Second-Year Algebra	15	45	37	3
Third-Year Algebra/Pre-Calculus	14	53	31	2
Calculus	25	59	14	2

NOTE: Row percentages may not total 100 due to rounding. Responses that could not be rated were excluded.
*** Sample size is insufficient to permit a reliable estimate.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 1996 Mathematics Assessment.

The question mapped at a score of 420. Twenty-three percent of the students classified as *Proficient* and 14 percent of those classified as *Basic* were credited with a fully correct response.

Table 6.8 ***Percentage Correct Within Achievement-Level Intervals for "Use Data in Table to Compute Average Hourly Wage and Determine When Wage Rate Changes"***

THE NATION'S REPORT CARD

	NAEP Grade 12 Composite Scale Range			
Overall	Below Basic	Basic	Proficient	Advanced
13	4	14	23	***

*** Sample size is insufficient to permit a reliable estimate.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 1996 Mathematics Assessment.

Sampling and statistics

Questions in this area, which appeared primarily on the instruments for eighth- and twelfth-grade students, assessed students' understanding of and ability to apply sampling theory and statistical analyses. Students were asked questions regarding sampling, data representation, and data summarization. They were evaluated on their understanding of the various measures of central tendency as well as on their ability to calculate these measures. Students also needed to understand concepts related to correlation. Three questions are presented here. Two were eighth-grade multiple-choice questions, and one was a twelfth-grade extended constructed-response question.

The first eighth-grade question assessed students' understanding of what can happen when a sample is taken. Students were told that a bag contained two red candies and one yellow candy and that each of two persons took one candy out of the bag, without replacement. The question then listed four combinations of candy colors and asked which combinations could have been drawn by these two people, given the candies in the bag. Only the fourth combination, both picking yellow candies, was not possible.

4. A bag contains two red candies and one yellow candy. Kim takes out one candy and eats it, and then Jeff takes out one candy. For each sentence below, fill in the oval to indicate whether it is possible or not possible.

Possible	Not Possible	
A	B	Kim's candy is red and Jeff's candy is red.
A	B	Kim's candy is red and Jeff's candy is yellow.
A	B	Kim's candy is yellow and Jeff's candy is red.
A	B	Kim's candy is yellow and Jeff's candy is yellow.

Table 6.9 presents the percentages of students responding correctly to none, one, two, three, or all four of the statements of sampling possibilities. Nearly 80 percent responded correctly to all four statements, and another 10 percent responded correctly to at least three of the statements. More females than males responded correctly to all four questions. When the question was anchored to the NAEP scale, the categories of none, one, or two correct responses to statements were collapsed.

Table 6.9 ***Score Percentages for "Reason About Sample Space"***

THE NATION'S REPORT CARD

	Number Correct					
	4	3	2	1	None	Omit
Grade 8						
Overall	79	10	3	6	1	0
Males	75	11	4	8	2	0
Females	85	9	2	4	1	0
White	86	8	2	2	0	0!
Black	61	14	7	14	4	0!
Hispanic	65	13	3	16	3	0!
Asian/Pacific Islander	– –	– –	– –	– –	– –	– –
American Indian	***	***	***	***	***	***
Mathematics Course Taking:						
Eighth-Grade Mathematics	76	12	5	6	1	1
Pre-Algebra	82	11	2	5	1	0
Algebra	86	7	1	4	2	0

NOTE: Row percentages may not total 100 due to rounding. Responses that could not be rated were excluded.
*** Sample size is insufficient to permit a reliable estimate.
– – Data for grade 8 Asian/Pacific Islanders are not reported due to concerns about the accuracy and precision of the national estimates. See Appendix A for further detail.
! Statistical tests involving this value should be interpreted with caution. Standard error estimates may not be accurately determined and/or the sampling distribution of the statistics does not match statistical test assumptions (see Appendix A).
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 1996 Mathematics Assessment.

This question was very easy for eighth-grade students. Table 6.10 shows that at least three-quarters of the students who performed below the *Basic* level on the NAEP mathematics assessment gave the correct response to at least three statements, as did nearly all of the students in the other achievement level categories. The question mapped at a score of 235.

Table 6.10 ***Percentage with at Least Three Correct Within Achievement-Level Intervals for "Reason About Sample Space"***

THE NATION'S REPORT CARD

	NAEP Grade 8 Composite Scale Range			
Overall	Below Basic	Basic	Proficient	Advanced
89	75	97	99!	100!

! Statistical tests involving this value should be interpreted with caution. Standard error estimates may not be accurately determined and/or the sampling distribution of the statistics does not match statistical test assumptions (see Appendix A).
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 1996 Mathematics Assessment.

The second example for this area is also an eighth-grade multiple-choice question. This question assessed students' understanding of what constitutes a representative sample. Students were told that a poll was being taken at a junior high school to determine the school mascot and were asked where they could find a sample of students to interview that was most representative of the students in the school.

6. A poll is being taken at Baker Junior High School to determine whether to change the school mascot. Which of the following would be the best place to find a sample of students to interview that would be most representative of the entire student body?

- (A) An algebra class
- (B) The cafeteria
- (C) The guidance office
- (D) A French class
- (E) The faculty room

The correct option is B.

This question also was fairly easy for eighth-grade students. Table 6.11 shows that 65 percent of the students answered correctly. One percent of the students chose Option D, the French class, while approximately 10 percent of the students chose each of the remaining three options. The question mapped at a score of 289, and most of the students whose performance was classified as *Basic* or above chose the correct option.

Table 6.11

Percentage Correct for "Identify Representative Sample"

THE NATION'S REPORT CARD

	Percentage Correct
Grade 8	
Overall	65
Males	64
Females	66
White	73
Black	48
Hispanic	47
Asian/Pacific Islander	– –
American Indian	***
Mathematics Course Taking:	
Eighth-Grade Mathematics	59
Pre-Algebra	67
Algebra	74

*** Sample size is insufficient to permit a reliable estimate.

– – Data for grade 8 Asian/Pacific Islanders are not reported due to concerns about the accuracy and precision of the national estimates. See Appendix A for further detail.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 1996 Mathematics Assessment.

Table 6.12

Percentage Correct Within Achievement-Level Intervals for "Identify Representative Sample"

THE NATION'S REPORT CARD

	NAEP Grade 8 Composite Scale Range			
Overall	Below Basic	Basic	Proficient	Advanced
65	43	72	87	96!

! Statistical tests involving this value should be interpreted with caution. Standard error estimates may not be accurately determined and/or the sampling distribution of the statistics does not match statistical test assumptions (see Appendix A).

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 1996 Mathematics Assessment.

The final example for this area is an extended constructed-response question for grade 12 in which students were asked to determine whether the mean or median better represented the typical daily attendance in each of two theaters and to justify their answers. They were provided with data on each theater's daily attendance over a 5-day period, along with the median and the mean of the 5 days. Theater A had a nontypical, or outlier, value for attendance on day 4. The attendance for Theater B was bimodal.

This question requires you to show your work and explain your reasoning. You may use drawings, words, and numbers in your explanation. Your answer should be clear enough so that another person could read it and understand your thinking. It is important that you show all of your work.

10. The table below shows the daily attendance at two movie theaters for 5 days and the mean (average) and the median attendance.

	Theater A	Theater B
Day 1	100	72
Day 2	87	97
Day 3	90	70
Day 4	10	71
Day 5	91	100
Mean (average)	75.6	82
Median	90	72

(a) Which statistic, the mean or the median, would you use to describe the typical daily attendance for the 5 days at Theater A? Justify your answer.

(b) Which statistic, the mean or the median, would you use to describe the typical daily attendance for the 5 days at Theater B? Justify your answer.

Did you use the calculator on this question?

○ Yes ○ No

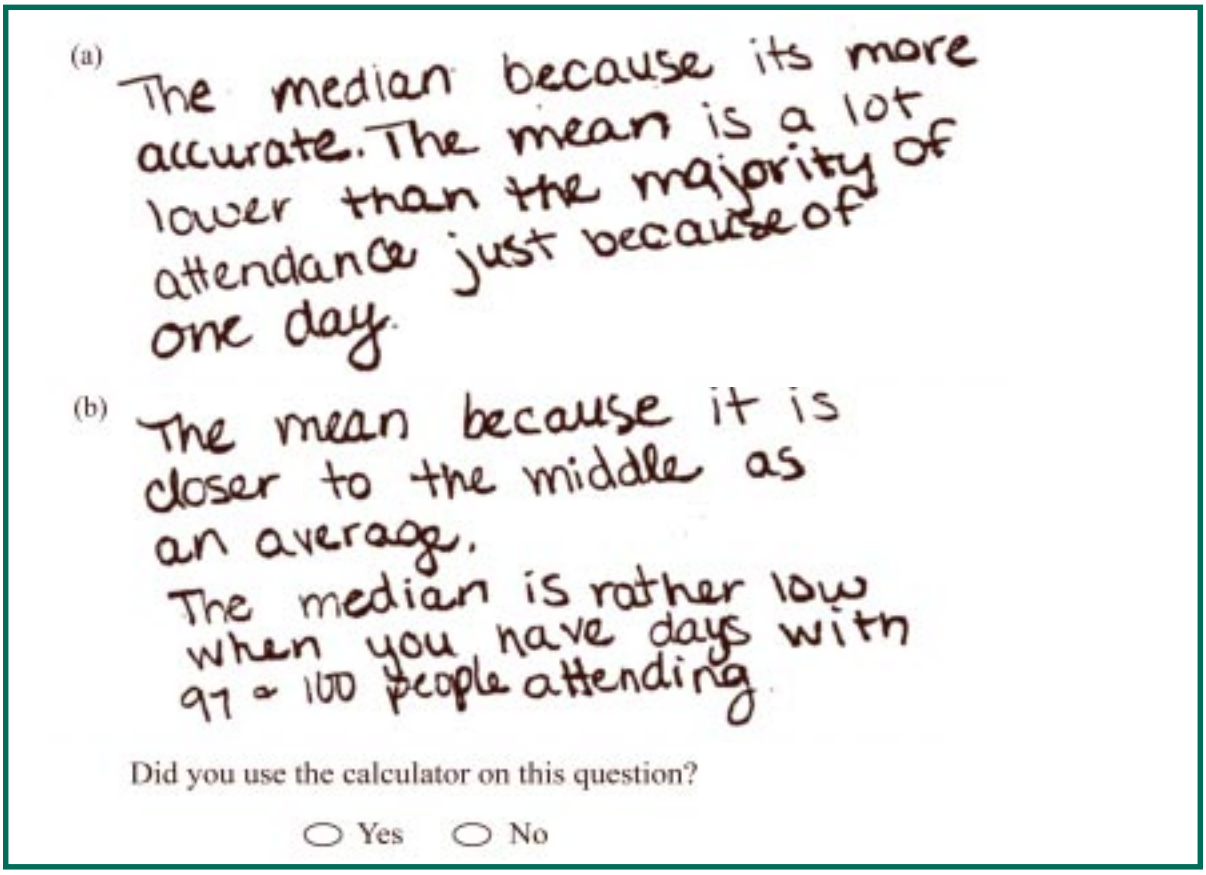
The correct answer for Theater A was the median, and the correct answer for Theater B was the mean. The appropriate explanation for the choice of the median for Theater A conveyed the idea that the attendance on day 4 was significantly different from attendance on the other days. Appropriate explanations for the choice of the mean for Theater B were variations on the following:

- There are two clusters of data;
- The median is representative of one of the clusters, while the mean is representative of both; and
- 82 is a better indicator of where the "center" of the five data points is located.

Responses were rated as "extended," "satisfactory," "partial," "minimal," and "incorrect." However, when the question was anchored to the NAEP scale, the "extended" and "satisfactory" rating categories were collapsed. A description of the ratings and sample responses for each rating category follow.

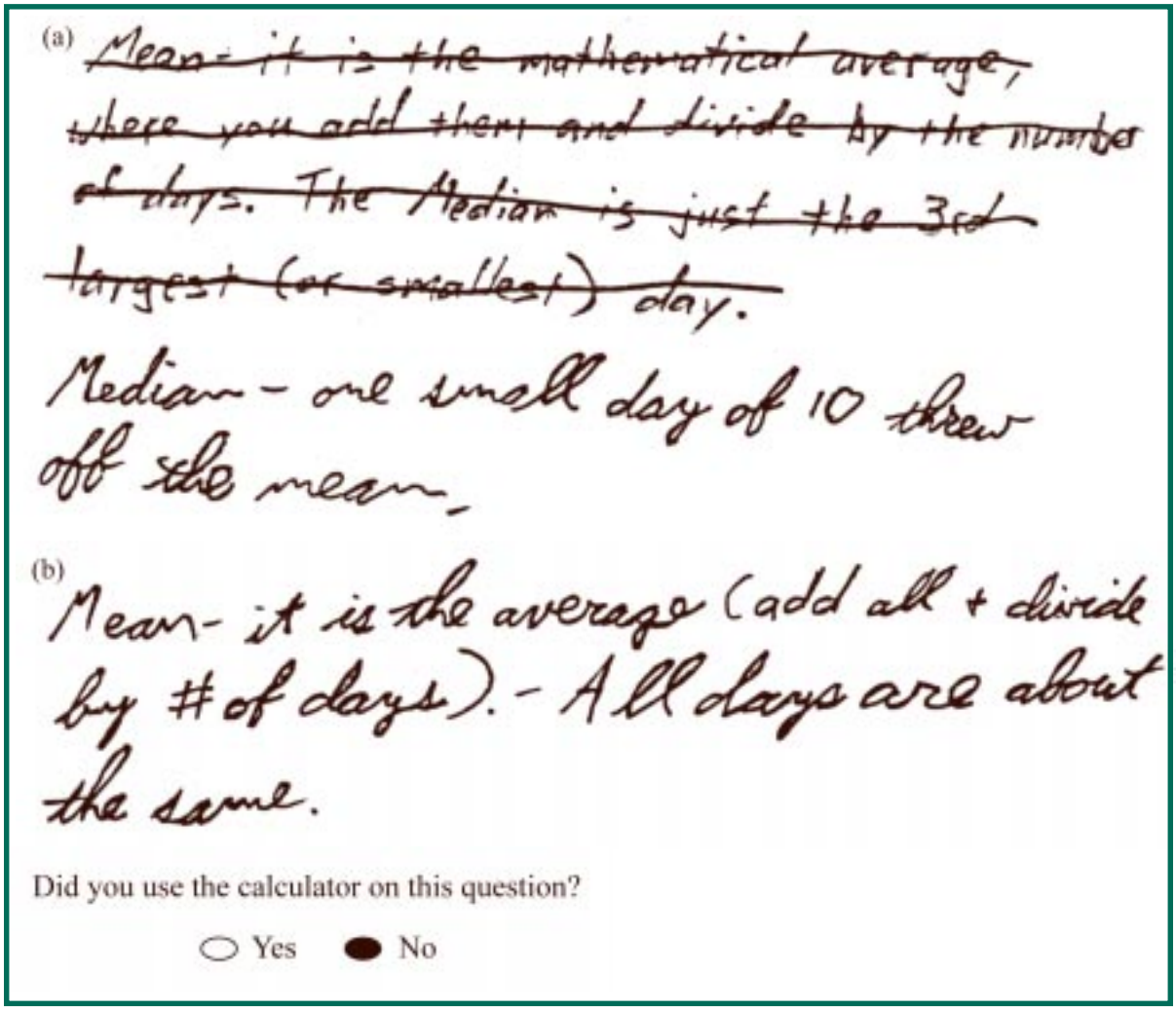
In order to have been rated as "extended," a student's response had to identify the appropriate measure for each theater and provide a correct explanation for at least one of the choices. The following is a sample of an "extended" response. After correctly identifying each statistic, the student explained that the median is better for Theater A because the mean is pulled down by 1 day, and that the mean is better for Theater B because it is closer to the middle.

Sample "extended" response



In the following sample, rated as "satisfactory," the student indicated the better measure for both theaters, but only provided a complete explanation for Theater A.

Sample "satisfactory" response


(a) ~~Mean- it is the mathematical average, where you add them and divide by the number of days. The Median is just the 3rd largest (or smallest) day.~~

Median - one small day of 10 threw off the mean.

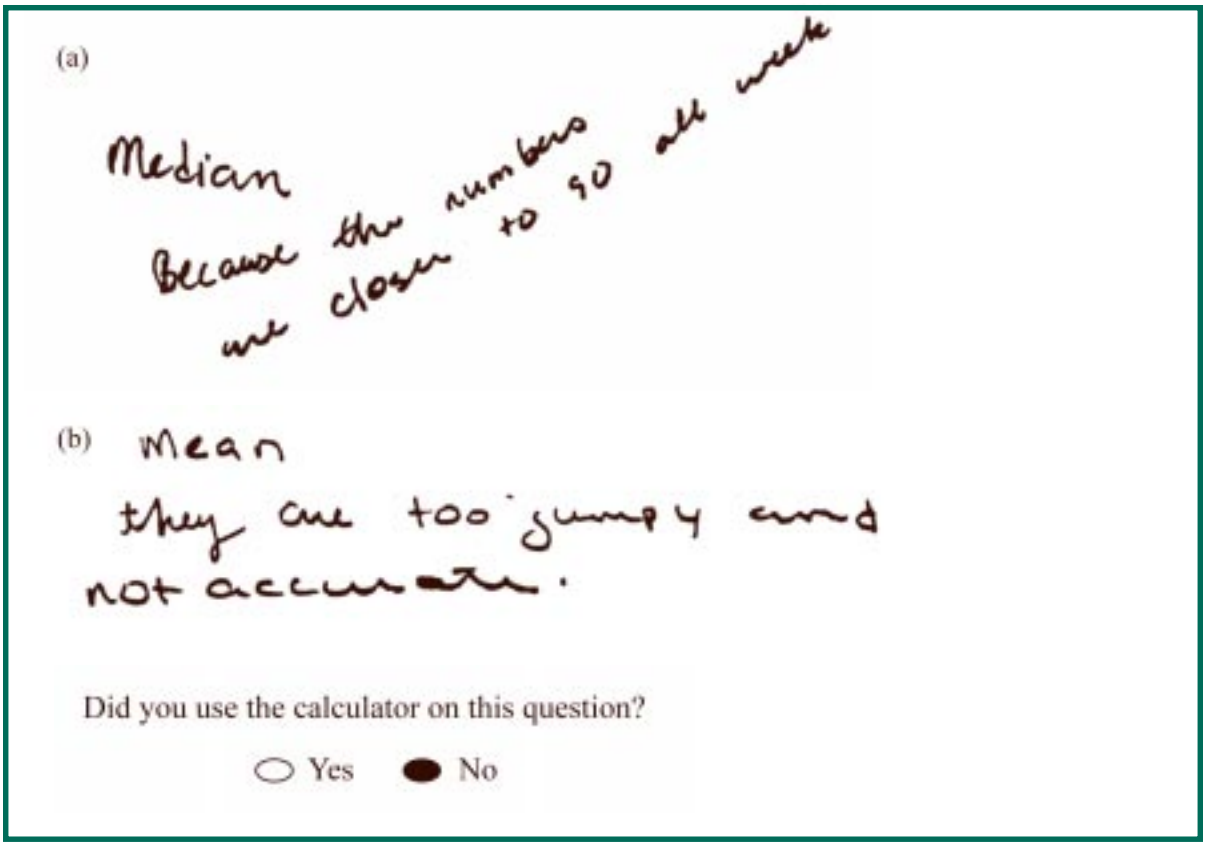
(b) Mean- it is the average (add all + divide by # of days). - All days are about the same.

Did you use the calculator on this question?

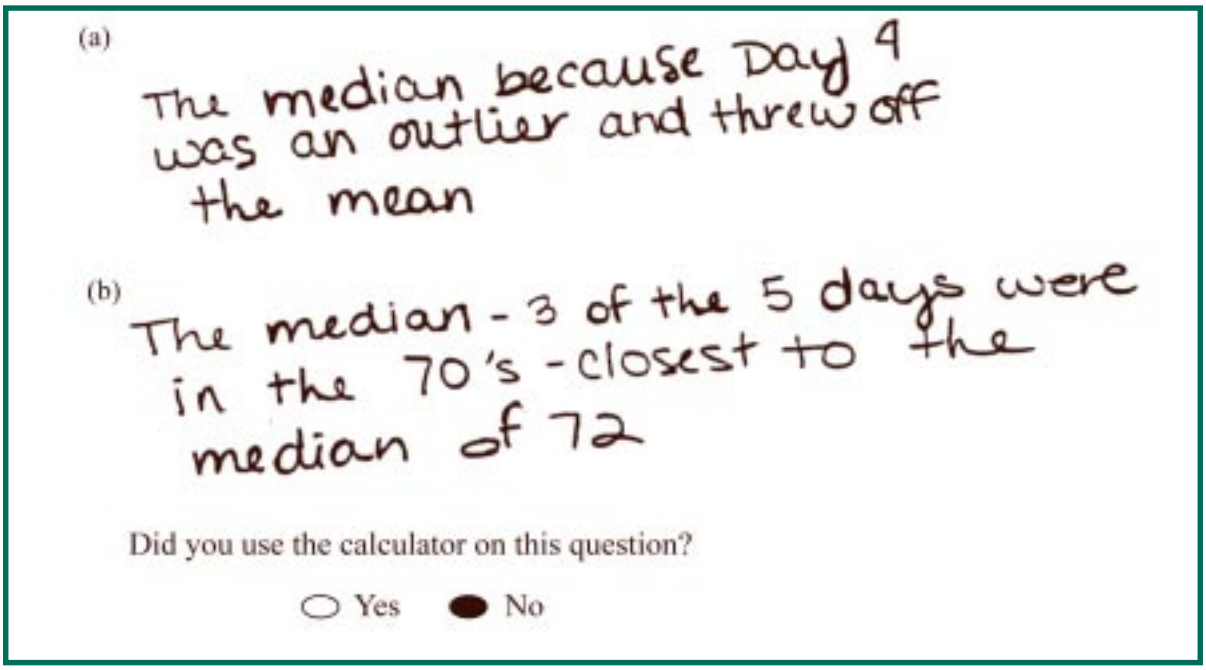
○ Yes ● No

Students' responses also could be rated as "partial" or "minimal." "Partial" responses either correctly identified the better measure for both theaters but did not provide appropriate explanations for either, or correctly identified and explained only one measure. Two examples of "partial" responses follow.

Sample "partial" response 1

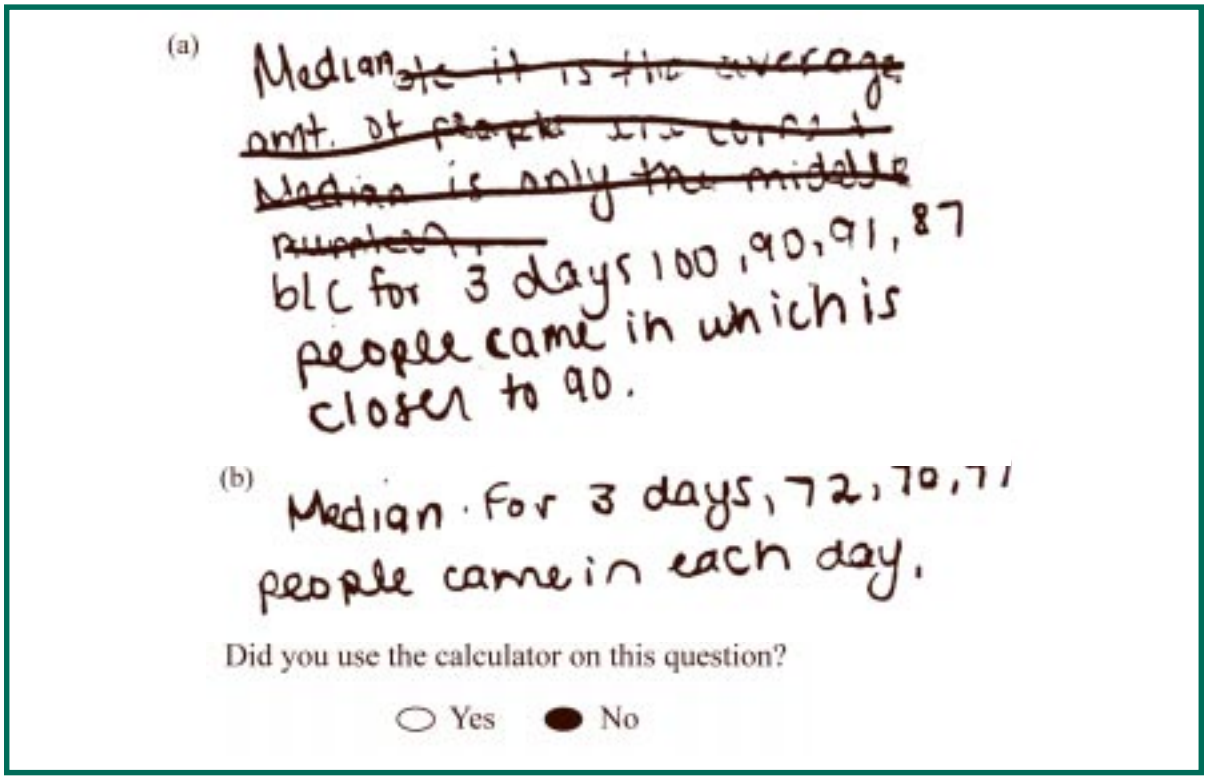


Sample "partial" response 2



The next two sample responses were rated as "minimal." Both students correctly identified the better measure for only one of the theaters (the first for Theater A and the second for Theater B); however, neither had an appropriate explanation. The first student came close with the explanation of the median for Theater A, but failed to complete the thought that the attendance for day 4 was an outlier compared to the other 4 days.

Sample "minimal" response 1

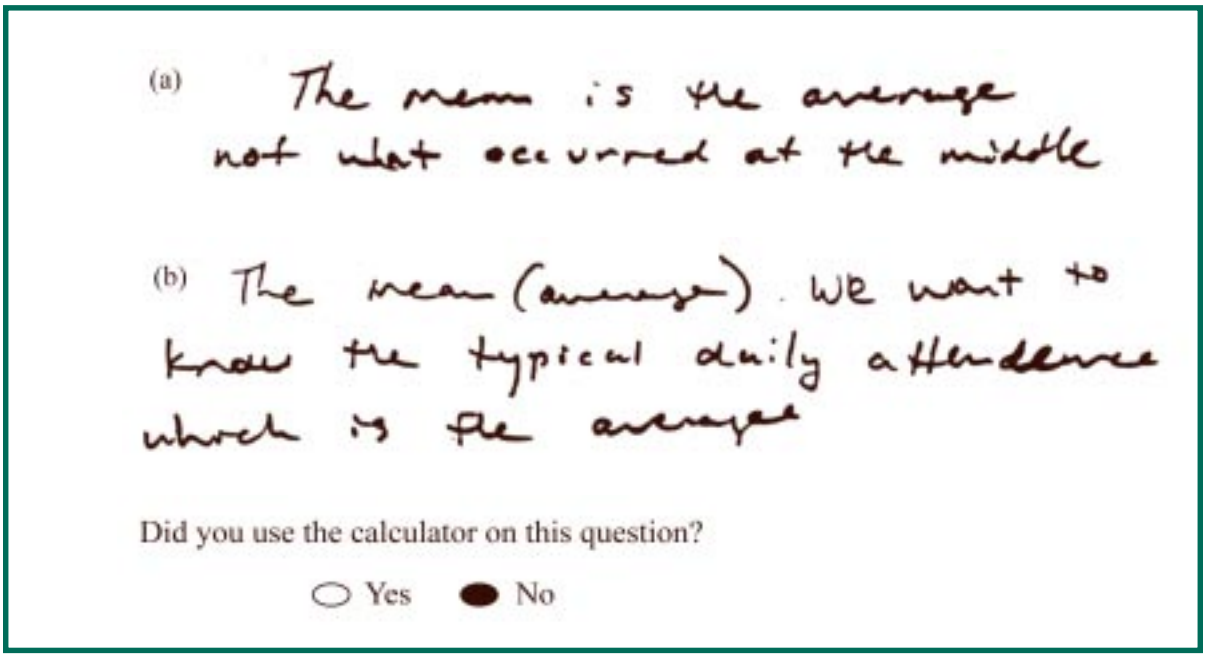

(a) Median ~~the it is the average amt. of people it's correct Median is only the middle number~~ blc for 3 days 100, 90, 91, 87 people came in which is closer to 90.

(b) Median · For 3 days, 72, 70, 71 people came in each day.

Did you use the calculator on this question?

○ Yes ● No

Sample "minimal" response 2


(a) The mean is the average not what occurred at the middle

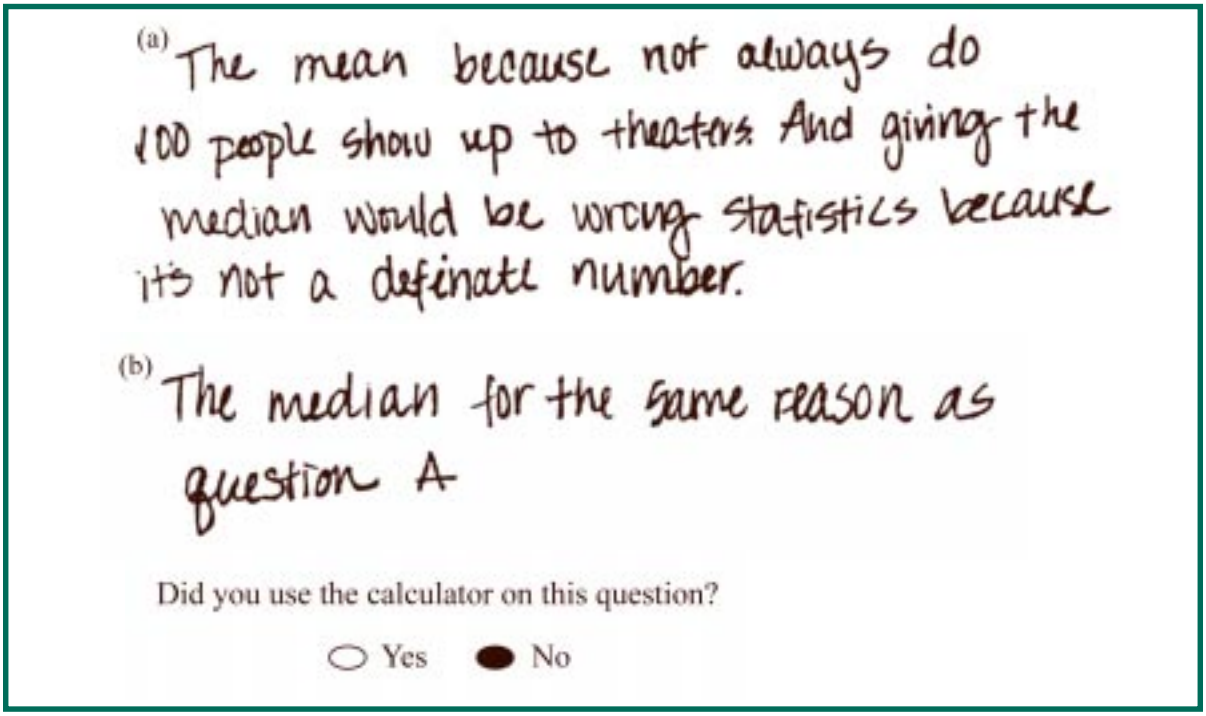
(b) The mean (average). We want to know the typical daily attendence which is the average

Did you use the calculator on this question?

○ Yes ● No

All other responses were considered "incorrect." The final example is of an "incorrect" response.

Sample "incorrect" response


(a) The mean because not always do 100 people show up to theaters And giving the median would be wrong statistics because it's not a definate number.

(b) The median for the same reason as question A

Did you use the calculator on this question?

○ Yes ● No

Tables 6.13 and 6.14 present student performance data for this question. The question was fairly difficult, and only four percent of the students chose the better measure for both theaters and gave a complete explanation for at least one of their choices. Slightly more than 30 percent of the students omitted the question, and over half of the students produced responses that were rated "incorrect" or "minimal." Students who had taken at least third-year algebra or pre-calculus were more likely than other students to choose the better measures for both theaters and offer at least one complete explanation; however, even among this group the percentage of responses that were at least "satisfactory" was small.

Table 6.13

Score Percentages for "Compare Mean and Median"

THE NATION'S REPORT CARD

	Extended	Satisfactory	Partial	Minimal	Incorrect	Omit
	Better Measure Both Theaters; Complete Explanation	Better Measure Both Theaters; Complete Explanation for 1 Theater	Better Measure and Complete Explanation 1 Theater; or Better Measure Both Theaters with No or Incomplete Explanation	Better Measure 1 Theater; No or Incomplete Explanation		
Grade 12						
Overall	1	3	10	28	25	31
Males	1	3	12	25	23	32
Females	0!	2	9	31	27	30
White	1	4	12	30	25	27
Black	0!	0!	7	25	24	42
Hispanic	0!	1!	6	18	24	48
Asian/Pacific Islander	1!	4	7	25	24	34
American Indian	***	***	***	***	***	***
Geometry Taken	1	3	10	29	27	30
Highest Algebra-Calculus Course Taken:						
Pre-Algebra	***	***	***	***	***	***
First-Year Algebra	0	2	8	28	26	33
Second-Year Algebra	0	2	10	29	26	30
Third-Year Algebra/Pre-Calculus	0	6	18	31	18	25
Calculus	6	10	11	26	25	21

NOTE: Row percentages may not total 100 due to rounding. Responses that could not be rated were excluded.
*** Sample size is insufficient to permit a reliable estimate.
! Statistical tests involving this value should be interpreted with caution. Standard error estimates may not be accurately determined and/or the sampling distribution of the statistics does not match statistical test assumptions (see Appendix A).
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 1996 Mathematics Assessment.

Table 6.14 shows that few of the students within any of the achievement level classifications received full credit on this question, and when the question was anchored to the NAEP scale, the "extended" and "satisfactory" rating categories were collapsed. The question mapped at 463 on the NAEP composite mathematics scale.

Table 6.14 ***Percentage at Least Satisfactory Within Achievement-Level Intervals for "Compare Mean and Median"***

THE NATION'S REPORT CARD

	NAEP Grade 12 Composite Scale Range			
Overall	Below Basic	Basic	Proficient	Advanced
4	0!	2	13	***

*** Sample size is insufficient to permit a reliable estimate.
! Statistical tests involving this value should be interpreted with caution. Standard error estimates may not be accurately determined and/or the sampling distribution of the statistics does not match statistical test assumptions (see Appendix A).
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 1996 Mathematics Assessment.

Probability

This area included questions measuring students' understanding of probabilistic events and their ability to determine the probability of simple and compound events. Questions for fourth-grade students used less advanced terminology than those for older students, and probabilities were simpler to calculate. Questions for older students required them to predict outcomes given two or more dependent events. Some questions also involved percents and proportions. Two questions are presented as examples for this area. One is a fourth-grade multiple-choice question, and the other is a twelfth-grade extended constructed-response question.

The following example question asked fourth-grade students to determine the chances that the person randomly chosen to be the captain of a swim team would be a fifth grader, given that the membership of the swim team was divided between fifth- and sixth-grade students in a specified manner. The language used in this example is typical of the probability questions presented in the fourth-grade assessment.

9. There are 3 fifth graders and 2 sixth graders on the swim team. Everyone's name is put in a hat and the captain is chosen by picking one name. What are the chances that the captain will be a fifth grader?

(A) 1 out of 5

(B) 1 out of 3

(C) 3 out of 5

(D) 2 out of 3

The correct option is C.

Tables 6.15 and 6.16 display student performance data on this question. Approximately one-third of the students responded correctly. Twenty-two percent of the students chose Option A, the probability that any individual student would be chosen, whereas 16 percent chose Option B, and 28 percent chose Option D. The appeal of the latter option may have been that it contained both of the numbers specified in the stem of the question. Clearly, many fourth-grade students did not know how to determine chance.

Table 6.15 ***Percentage Correct for "Determine a Probability"*** THE NATION'S REPORT CARD naep

	Percentage Correct
Grade 4	
Overall	31
Males	32
Females	30
White	34
Black	22
Hispanic	26
Asian/Pacific Islander	35
American Indian	***

*** Sample size is insufficient to permit a reliable estimate.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 1996 Mathematics Assessment.

Table 6.16 shows the percentage of students within each of the achievement level intervals who responded correctly to this question. Half of the students classified as *Proficient* responded correctly compared with approximately one-quarter of the students at each of the lower two levels. The question mapped at 278.

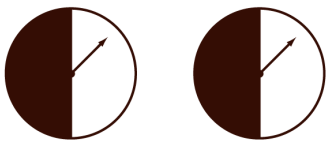
Table 6.16 ***Percentage Correct Within Achievement-Level Intervals for "Determine a Probability"***

THE NATION'S REPORT CARD

	NAEP Grade 4 Composite Scale Range			
Overall	Below Basic	Basic	Proficient	Advanced
31	23	24	54	***

*** Sample size is insufficient to permit a reliable estimate.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 1996 Mathematics Assessment.

The final example is a twelfth-grade extended constructed-response question on joint probabilities. The question showed two spinners that were half black and half white, and students were told that to "win," both arrows had to land on black when the spinner was spun once. They then were asked whether they agreed that there was a 50-50 chance of this happening and instructed to justify their answers. The correct response was "no" because the possibility of either event happening was 1 in 2; therefore, the possibility of both happening was 1 in 4, or 25 percent.



9. The two fair spinners shown above are part of a carnival game. A player wins a prize only when both arrows land on black after each spinner has been spun once.

James thinks he has a 50-50 chance of winning. Do you agree?

Ⓐ Yes Ⓑ No

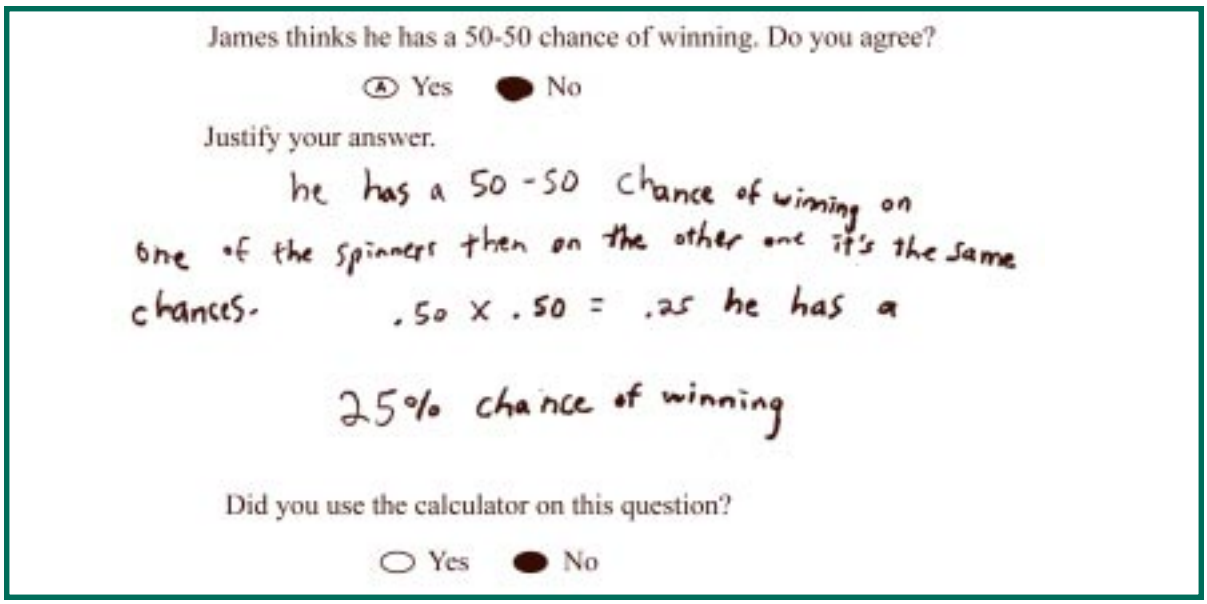
Justify your answer.

Did you use the calculator on this question?

◯ Yes ◯ No

Students' explanations were rated "correct," "partial," or "incorrect." A "correct" explanation had to indicate that the actual chances were 1 in 4, or 25 percent, and correctly justify this conclusion. Both of the following responses were considered "correct."

Sample "correct" response 1


James thinks he has a 50-50 chance of winning. Do you agree?

(A) Yes ● No

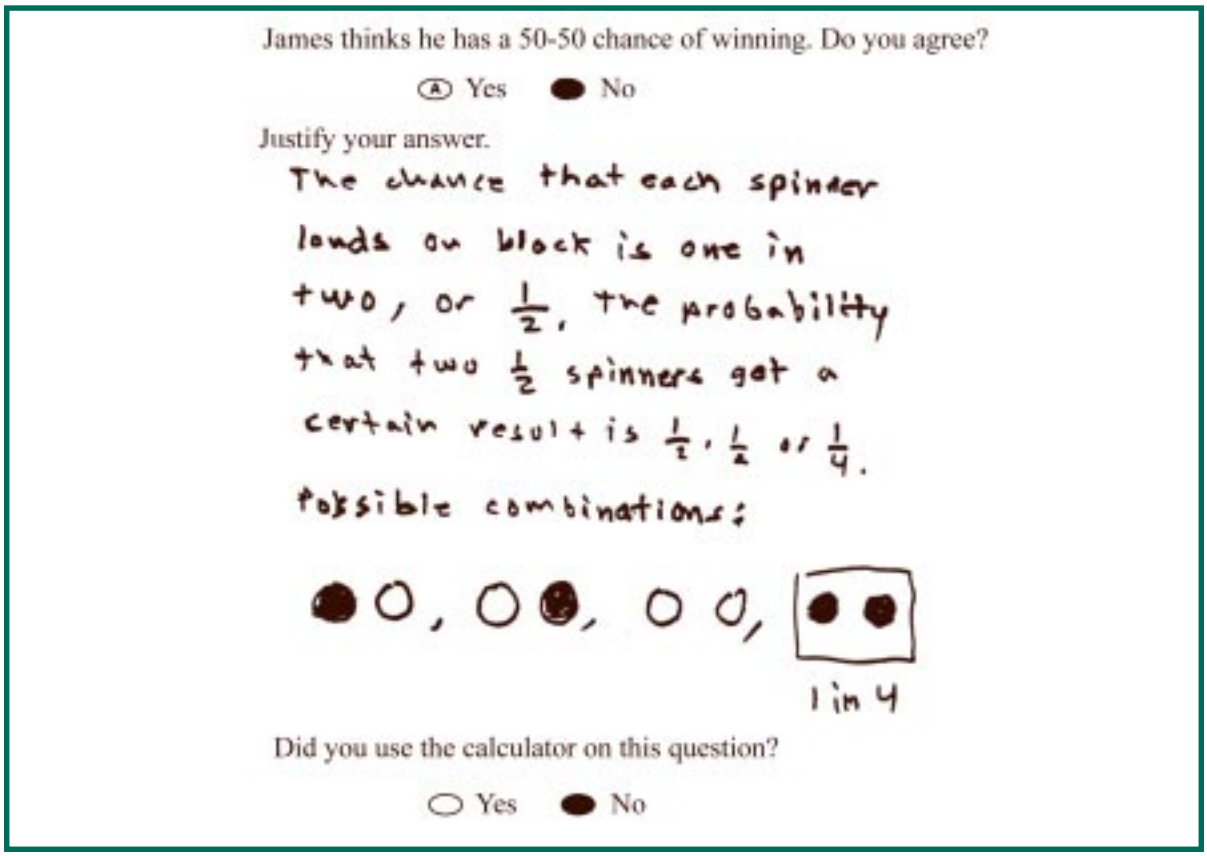
Justify your answer.

he has a 50-50 chance of winning on one of the spinners then on the other one it's the same chances. .50 x .50 = .25 he has a 25% chance of winning

Did you use the calculator on this question?

○ Yes ● No

Sample "correct" response 2


James thinks he has a 50-50 chance of winning. Do you agree?

(A) Yes ● No

Justify your answer.

The chance that each spinner lands on black is one in two, or $\frac{1}{2}$, the probability that two $\frac{1}{2}$ spinners get a certain result is $\frac{1}{2}$, $\frac{1}{2}$ or $\frac{1}{4}$.

Possible combinations:

●○, ○●, ○○, ●● 1 in 4

Did you use the calculator on this question?

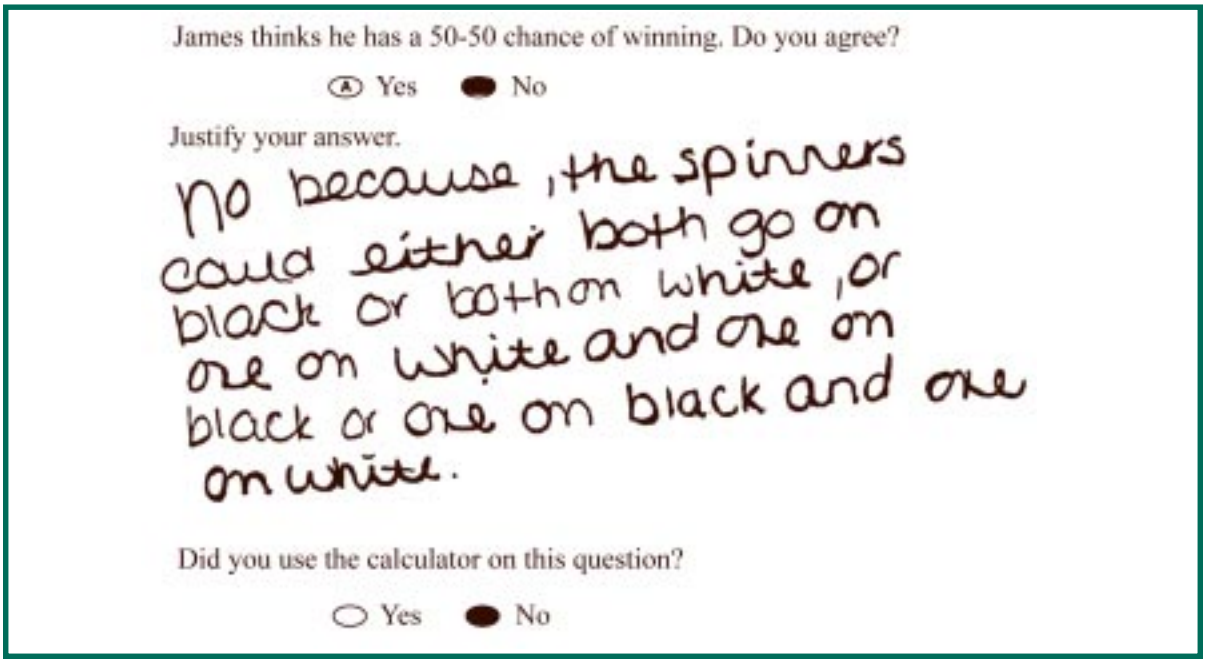
○ Yes ● No

Students also could have simply drawn a diagram similar to the one presented in the second sample above and still have been considered as giving a "correct" response. For a response to have been rated as "partial," students had to do one of the following:

- list the sample space correctly, but with less than a complete explanation;
- draw a correct tree diagram, but with less than a complete explanation; or
- simply state that the chance would be 1 in 4.

In the following example of a "partial" explanation, the student described the sample space but did not tell what the actual chances of winning were.

Sample "partial" response


James thinks he has a 50-50 chance of winning. Do you agree?

Ⓐ Yes ● No

Justify your answer.

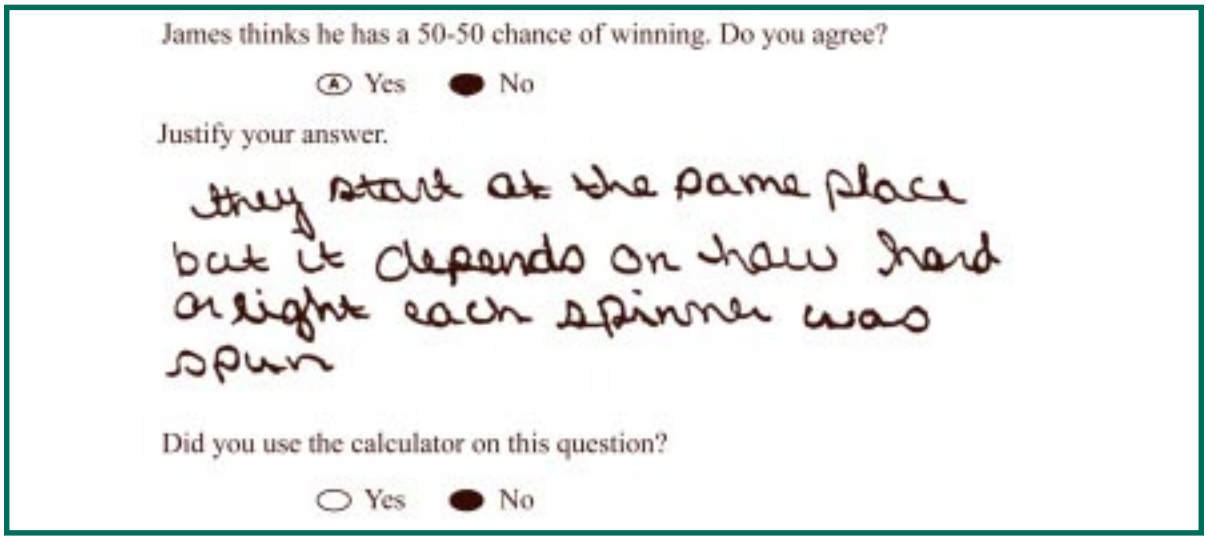
no because, the spinners could either both go on black or both on white, or one on white and one on black or one on black and one on white.

Did you use the calculator on this question?

○ Yes ● No

"Incorrect" explanations included all explanations that did not meet the criteria stated above. Note that students who responded correctly to the initial "yes/no" question but were not able to provide at least a partially adequate explanation received a rating of "incorrect." Two examples of responses that were rated as "incorrect" follow.

Sample "incorrect" response 1


James thinks he has a 50-50 chance of winning. Do you agree?

(A) Yes ● No

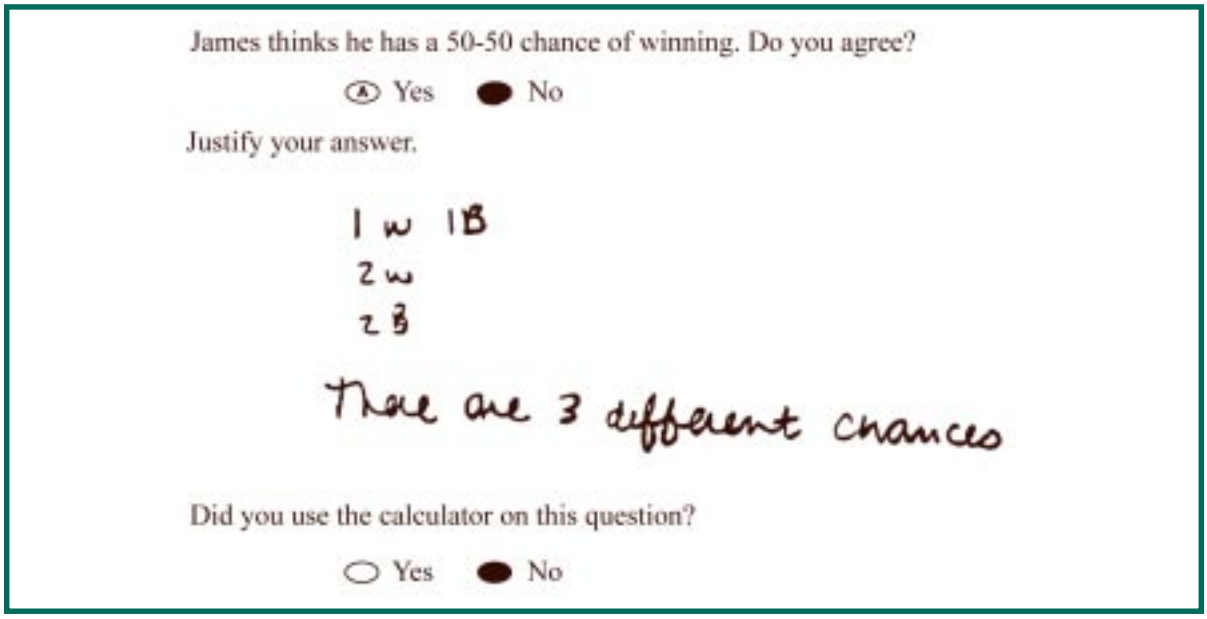
Justify your answer.

they start at the same place but it depends on how hard or light each spinner was spun

Did you use the calculator on this question?

○ Yes ● No

Sample "incorrect" response 2


James thinks he has a 50-50 chance of winning. Do you agree?

(A) Yes ● No

Justify your answer.

1 w 1B
2 w
2 B

There are 3 different chances

Did you use the calculator on this question?

○ Yes ● No

Student performance data are presented in Tables 6.17 and 6.18. Forty-four percent of the students provided an "incorrect" response to the initial question, meaning that they did not answer "no" to whether there was a 50-50 chance of the spinners both landing on black. Approximately one-quarter of the students answered the initial question correctly but provided an incorrect explanation. The remainder were able to give an explanation that was at least partially correct. Students whose highest course was calculus were substantially more likely than other students to provide a fully correct explanation: 34 percent of these students provided a response rated "correct." Males were more likely than females to provide at least a partial explanation to the question.

Table 6.17 ***Score Percentages for "Compare Probabilities"*** THE NATION'S REPORT CARD

	Correct	Partial	Incorrect		Omit
	Correct Answer to "Yes/No" Question; Correct Explanation	Correct Answer to "Yes/No" Question; Partial Explanation	Correct Answer to "Yes/No" Question; Incorrect Explanation	Incorrect Answer to "Yes/No" Question	
Grade 12					
Overall	8	20	24	44	4
Males	9	23	25	40	3
Females	7	16	24	49	4
White	9	23	26	39	3
Black	4	8	18	62	7
Hispanic	5	7	18	65	4
Asian/Pacific Islander	11	20	29	38	1!
American Indian	***	***	***	***	***
Geometry Taken	8	22	25	41	4
Highest Algebra-Calculus Course Taken:					
Pre-Algebra	***	***	***	***	***
First-Year Algebra	5	12	24	56	3
Second-Year Algebra	7	21	28	42	3
Third-Year Algebra/Pre-Calculus	8	37	21	26	8
Calculus	34	24	16	19	6

NOTE: Row percentages may not total 100 due to rounding. Responses that could not be rated were excluded.
*** Sample size is insufficient to permit a reliable estimate.
! Statistical tests involving this value should be interpreted with caution. Standard error estimates may not be accurately determined and/or the sampling distribution of the statistics does not match statistical test assumptions (see Appendix A).
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 1996 Mathematics Assessment.

The question mapped at a score of 411. Almost two-thirds of the students who were classified as being *Proficient* on the NAEP mathematics assessment responded with at least a partial explanation. However, only 27 percent of those classified as *Basic* and 5 percent of those classified as below *Basic* performed as well.

Table 6.18 ***Percentage Correct Within Achievement-Level Intervals for "Compare Probabilities"***

THE NATION'S REPORT CARD

	NAEP Grade 12 Composite Scale Range			
Overall	Below Basic	Basic	Proficient	Advanced
28	5	27	65	***

*** Sample size is insufficient to permit a reliable estimate.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 1996 Mathematics Assessment.

Summary

This content strand included questions that assessed students' understanding of data, including how to best collect, display, and interpret data. Questions also assessed students' understanding of statistics and probability and their competence in calculating statistics and determining probabilities. Statistics included mean, median, mode, and standard deviation of distributions, and probabilities could be simple, dual, or conditional.

As might be expected, straightforward interpretations of graphs, charts, and tables were easier for students than questions that asked them to perform calculations with displayed data. Students also had difficulty explaining why one method of reporting or displaying data was better than another, even though they may have recognized which was the better method. Questions asking students to determine chance or probability were generally difficult for them.