AP | CollegeBoard

AP® Statistics
Sample Student Responses and Scoring Commentary

Inside:

Free-Response Question 2

☑ **Scoring Guidelines**

☑ **Student Samples**

☑ **Scoring Commentary**

Question 2: Focus on Collecting Data   4 points

General Scoring Notes

- Each part of the question (indicated by a letter) is initially scored by determining if it meets the criteria for essentially correct (E), partially correct (P), or incorrect (I). The response is then categorized based on the scores assigned to each letter part and awarded an integer score between 0 and 4 (see the table at the end of the question).
- The model solution represents an ideal response to each part of the question, and the scoring criteria identify the specific components of the model solution that are used to determine the score.

Model Solution	Scoring
(a) Treatments: New drug, placebo. Experimental units: The 72 people who receive the new drug or placebo. Response variable: Improvement in acne severity	**Essentially correct (E)** if the response satisfies the following three components: 1. Identifies the treatments as new drug and placebo 2. Identifies the experimental units as the 72 people (subjects, participants, twins) in the experiment 3. Identifies the response variable as the improvement in acne severity **Partially correct (P)** if the response satisfies only two of the three components. **Incorrect (I)** if the response does not satisfy the criteria for E or P.

Additional Notes:
- To satisfy component 1, identification of the treatments must include both the placebo and the new drug.
- To satisfy component 2, the response must indicate that the experimental units are individual people. The response could refer to participants, subjects, twins, or members of the pairs of twins without explicitly mentioning the number 72. However, a response that states or implies that there are 36 experimental units (e.g., "the pairs of twins") does not satisfy component 2.
- To satisfy component 3, the response must include the context of "acne" and "improvement" (e.g., "improvement in acne severity," "acne improvement score"), but it does not need to include a reference to the scale, the dermatologist, two-week time periods, or treatments. Reasonable synonyms for improvement can be used, such as using "reduction" or "change" or by including the verbal descriptions of the scale ("no improvement" to "complete cure"). However, a description of a binary outcome (e.g., "whether or not the acne improves") does not satisfy component 3.
- For responses that indicate the 36 pairs of twins are the experimental units, component 3 may be satisfied by indicating that the response variable is the improvement in acne severity or by indicating that the response variable is the difference in improvement in acne severity.
- If the response provides parallel solutions (i.e., two or more complete solutions without choosing or indicating which is to be scored), the response is scored based on the weaker of the two solutions. For example, if a response says that the experimental units are "the 72 participants and the scores from 0 to 100," component 2 is not satisfied.

Model Solution	Scoring
(b) Improvement scores will vary due to many factors, including initial acne severity, what treatment is received, and other variables such as diet and genetics. Because the pairs of twins are similar in initial acne severity, pairing allows for the variation in improvement scores due to the treatment received to be distinguished from variation due to initial acne severity, unlike in a completely randomized design. Consequently, using the matched-pairs design will provide a more precise estimate of the mean difference in improvement in acne severity for the new drug compared to the placebo and make it easier to find convincing evidence that the new drug is better, if it really is better.	**Essentially correct (E)** if the response describes a statistical advantage of a matched-pairs design AND satisfies the following three components: 1. The advantage pertains to an inference made after collecting the data (e.g., the ability to distinguish between the effects of the treatments or the precision of the estimate of the drug effect) 2. Indicates that the matched-pairs design is better by using a comparative word (e.g., easier, clearer, greater) or by making an explicit comparison to a completely randomized design 3. Includes context (e.g., "drug," "improvement," "acne," or "twins") **Partially correct (P)** if the response describes a statistical advantage of a matched-pairs design AND satisfies one or two of the three components. **Incorrect (I)** if the response does not satisfy the criteria for E or P.

Additional Notes:

- To be considered an advantage of a matched-pairs design, the advantage described must be true for a matched-pairs design and not be true for a completely randomized design. For example, saying that "random assignment allows us to conclude cause-and-effect" is true of both designs. Similarly, "this allows the dermatologist to make conclusions about people with differing acne severity" is true of both designs. Also, "reduces bias" and "reduces variability in the estimates of the individual treatment means" is true of neither design.
- Responses that describe only the set-up of a matched-pairs experiment do not satisfy the requirement to describe an advantage of a matched-pairs design. For example, the response "in a matched-pairs design, the members of each pair will be similar in terms of acne severity" does not describe an advantage. However, "in a matched-pairs design, we can compare two people with similar acne severity" does describe an advantage.
- Advantages of a matched-pairs design that satisfy component 1 include "makes it easier to determine if the drug is effective," "gives a better estimate of the effect of the new drug," "reduces variability in the estimate of the drug effect," "makes the difference between the drug and the placebo more easily distinguishable," and "gives a clearer picture of how well the drug works."
- Advantages of a matched-pairs design that don't satisfy component 1 include "accounts for a source of variability," "controls for potentially confounding variables," "allows you to distinguish variation due to severity from variation due to treatment," "each person can be compared to someone similar," "reduces variability," "more balanced treatment groups," and "more accurate results."

- It is acceptable to provide a disadvantage of a completely randomized design rather than an advantage of the matched-pairs design (e.g., "The completely randomized design will make it harder to find convincing evidence that the new drug is better").
- It is acceptable to use the term "blocking" as a synonym for "pairing."
- A response that states that a matched-pairs design requires a smaller sample size to get power or precision equal to that in a completely randomized design and describes this advantage in context should be scored E.

Model Solution	Scoring
(c) For each pair of twins, label one person as twin A and label the other person as twin B. For each pair of twins, toss a coin. If the coin lands on heads, twin A gets the placebo and twin B gets the active drug. If the coin lands on tails, twin A gets the active drug and twin B gets the placebo. *OR* Label the members of each pair of twins as "Twin 1" and "Twin 2." Using a random number generator, generate an integer from 1 to 2. Give the drug to the twin whose number is selected and the placebo to the twin whose number is not selected. Repeat for all pairs of twins. *OR* Label 1 notecard "A" and another notecard "B." For each pair of twins, shuffle the cards and give one card to each twin. The twin who gets "A" receives the drug and the twin who gets "B" receives the placebo.	**Essentially correct (E)** if the response randomly assigns the two treatments within pairs of twins AND satisfies the following three components: 1. Uses a random process (e.g., flipping a coin, using a random number generator, shuffling cards) that gives each twin in a pair a 50% probability of getting the drug and a 50% probability of getting the placebo 2. Describes how to use the random process to assign one specific twin in each pair to the drug and the other twin to the placebo 3. Indicates that the random assignment process will be completed for each pair of twins **Partially correct (P)** if the response randomly assigns the two treatments within pairs of twins AND satisfies only two of the three components for E. **Incorrect (I)** if the response does not satisfy the criteria for E or P.

Additional Notes:

- A response that does not randomly assign both treatments within pairs of twins should be scored incorrect (I). Examples include a response that describes a completely randomized design, describes a crossover design where each person receives both treatments, uses pairs other than twins, does not use random assignment, or indicates that both twins in a pair receive the same treatment.
- For responses that use slips of paper or selecting items from a hat, the slips must be shuffled (or blindly drawn) or the hat mixed or shaken to have a random process and satisfy component 1.
- To satisfy component 2, the response must describe what to do for each possible outcome of the random process and specify which treatment each twin receives. For example, none of the following descriptions satisfy component 2:
  o "Roll a die. If it is 1–3, give the first twin the drug and the second twin the placebo." (Response doesn't describe what to do if the die is 4–6.)
  o "Have one member of each pair flip a coin. If it is heads, that twin gets the drug. If it is tails, that twin gets the placebo." (Response doesn't indicate what treatment the other twin will receive.)
  o "Flip a coin. If it is heads, give one twin the drug and the other twin the placebo. If it is tails, do the reverse." (Response doesn't specify which twin is getting the drug.)
  o "Label one slip of paper "A" and a second slip of paper "B." Mix them in a hat and have each member of the pair choose one slip." (Response doesn't specify if A represents the new drug or the placebo.)

- Ignore any discussion about randomly selecting 36 pairs of twins to obtain subjects for the experiment. Likewise, ignore any discussion about how to perform the analysis for a paired design (e.g., "subtract the improvement scores for each pair of twins").
- It is acceptable to refer to each pair of twins as a block.

Scoring for Question 2	Score
Complete Response Three parts essentially correct	**4**
Substantial Response Two parts essentially correct and one part partially correct	**3**

Developing Response
Two parts essentially correct and no part partially correct
    OR
One part essentially correct and one or two parts partially correct
    OR
Three parts partially correct
	2

Minimal Response One part essentially correct and no part partially correct *OR* No part essentially correct and two parts partially correct	**1**

Question 2

Begin your response to **QUESTION 2** on this page.

2. A dermatologist will conduct an experiment to investigate the effectiveness of a new drug to treat acne. The dermatologist has recruited 36 pairs of identical twins. Each person in the experiment has acne and each person in the experiment will receive either the new drug or a placebo. After each person in the experiment uses either the new drug or the placebo for 2 weeks, the dermatologist will evaluate the improvement in acne severity for each person on a scale from 0 (no improvement) to 100 (complete cure).

(a) Identify the treatments, experimental units, and response variable of the experiment.

- **Treatments:** There are two treatments: the placebo and the real, new drug.

- **Experimental units:** The experimental units refer to the individuals; each person in the experiment.

- **Response variable:** The response variable is the improvement in acne severity, measured on a scale from 0 to 100.

Each twin in the experiment has a severity of acne similar to that of the other twin. However, the severity of acne differs from one twin pair to another.

(b) For the dermatologist's experiment, describe a statistical advantage of using a matched-pairs design where twins are paired rather than using a completely randomized design. While a completely randomized design would help ensure that the two resulting groups — the placebo-receiving and the drug-receiving groups — are similar, it does not help reduce the variability between each individual or twin pair, and so these differences could confound with the treatment of the drug in determining effectiveness of the new drug. The matched pairs design ensures that multiple treatments are given to the same — in this case, very similar — individuals, minimizing variability amongst individuals and enabling a comparison of the effectiveness of the new drug vs. placebo among different twin pairs.

Use a pencil or a pen with black or dark blue ink. Do NOT write your name. Do NOT write outside the box.
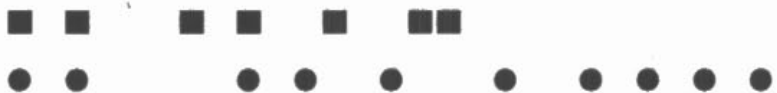
Q5149/4

Question 2

Continue your response to **QUESTION 2** on this page.

(c) For the dermatologist's experiment, describe how the treatments can be randomly assigned to people using a matched-pairs design in which twins are paired. Assign the number 0 to the placebo, and the number 1 to the new drug. Choose one twin from each twin pair. For each chosen twin, use a random number generator to generate an integer between 0 and 1, inclusive. If the twin receives a "0", assign to them the placebo and their twin the new drug. If the twin gets a "1", they are assigned to the new drug, their twin the placebo. Repeat for each twin pair

Use a pencil or a pen with black or dark blue ink. Do NOT write your name. Do NOT write outside the box.

0051525   ■ ■ ■   ■ ■ ■   ■■

Q5149/5

Question 2

Begin your response to **QUESTION 2** on this page.

2. A dermatologist will conduct an experiment to investigate the effectiveness of a new drug to treat acne. The dermatologist has recruited 36 pairs of identical twins. Each person in the experiment has acne and each person in the experiment will receive either the new drug or a placebo. After each person in the experiment uses either the new drug or the placebo for 2 weeks, the dermatologist will evaluate the improvement in acne severity for each person on a scale from 0 (no improvement) to 100 (complete cure).

(a) Identify the treatments, experimental units, and response variable of the experiment.

- Treatments: the type of drug received (new or placebo)

- Experimental units: The people participating in this experiment

- Response variable: Level of improvement after two weeks

Each twin in the experiment has a severity of acne similar to that of the other twin. However, the severity of acne differs from one twin pair to another.

(b) For the dermatologist's experiment, describe a statistical advantage of using a matched-pairs design where twins are paired rather than using a completely randomized design.

A matched pairs design is statistically advantageous in this experiment because it reduces the effect of initial acne severity as a confounding variable. Since each twin in a pair has a similar acne severity to the other, it is more effective to determine the drug's effectiveness when comparing between twins, as opposed to a randomized design where different acne severities may make it harder to determine how much of an effect the drug actually had.

Page 4   **GO ON TO THE NEXT PAGE.**

Use a pencil or a pen with black or dark blue ink. Do NOT write your name. Do NOT write outside the box.

Q5149/4

Question 2

Continue your response to **QUESTION 2** on this page.

(c) For the dermatologist's experiment, describe how the treatments can be randomly assigned to people using a matched-pairs design in which twins are paired.

For each pair of identical twins flip a fair coin to determine which twin gets the experimental treatment. Whichever twin does not get the experimental treatment, gets the placebo instead.

    Page 5   **GO ON TO THE NEXT PAGE.**

Use a pencil or a pen with black or dark blue ink. Do NOT write your name. Do NOT write outside the box.

0003294

Q5149/5

Question 2

Begin your response to **QUESTION 2** on this page.

2. A dermatologist will conduct an experiment to investigate the effectiveness of a new drug to treat acne. The dermatologist has recruited 36 pairs of identical twins. Each person in the experiment has acne and each person in the experiment will receive either the new drug or a placebo. After each person in the experiment uses either the new drug or the placebo for 2 weeks, the dermatologist will evaluate the improvement in acne severity for each person on a scale from 0 (no improvement) to 100 (complete cure).

(a) Identify the treatments, experimental units, and response variable of the experiment.

- Treatments: New Drug OR placebo (no drug)

- Experimental units: pairs of identical twins (36 pairs)

- Response variable: improvement of acne on scale from 0 - 100.

Each twin in the experiment has a severity of acne similar to that of the other twin. However, the severity of acne differs from one twin pair to another.

(b) For the dermatologist's experiment, describe a statistical advantage of using a matched-pairs design where twins are paired rather than using a completely randomized design.

It would be statistically advantageous to use a matched pairs design as this would lead to more accurate results, being each twin has only similar severity of acne to each other and not another pair of twins. If this were a completely randomized design, two pairs of twins that DO NOT have same acne severity are likely to be compared which would not led to fair and accurate results.

Page 4   **GO ON TO THE NEXT PAGE.**

Use a pencil or a pen with black or dark blue ink. Do NOT write your name. Do NOT write outside the box.

Q5149/4

Question 2

Continue your response to **QUESTION 2** on this page.

(c) For the dermatologist's experiment, describe how the treatments can be randomly assigned to people using a matched-pairs design in which twins are paired.

In order to conduct a matched pairs design within this experiment, I would take each individual from the pair of twins and give them ~~each~~ both treatments (new drug and placebo). From this, you are able to obtain the results from each individual (in pair of twins) and evaluate the improvement of ~~the~~ acne of the scale of 0 - 100. In addition to this, to assure the treatments are randomly assigned you can use ~~the~~ "double blinded experiment" to assure the participants nor dermatologists are aware of what treatment they are administrating

Page 5   **GO ON TO THE NEXT PAGE.**

Use a pencil or a pen with black or dark blue ink. Do NOT write your name. Do NOT write outside the box.

0051549

Q5149/5

Question 2

Note: Student samples are quoted verbatim and may contain spelling and grammatical errors.

Overview

The primary goals of the question were to assess a student's ability to (1) identify the treatments, the experimental units, and the response variable from a description of an experiment; (2) identify a statistical advantage of a matched pairs design (such as increased ability to detect a treatment effect, reduced variability of the difference in treatment means, or more precise estimation of the treatment effect) relative to a completely randomized design; and (3) describe a correct procedure for randomly assigning two treatments to experimental units in a matched pairs experiment.

This question primarily assesses skills in skill category 1: Selecting Statistical Methods. Skills required for responding to this question include (1.B) Identify key and relevant information to answer a question or solve a problem, and (1.C) Describe an appropriate method for gathering and representing data.

This question covers content from Unit 3: Collecting Data of the course framework in the AP Statistics Course and Exam Description. Refer to topics 3.5 and 3.6 and learning objectives VAR-3.A, VAR-3.D, and VAR-3.B.

Sample: 2A
Score: 4

The response earned the following: Part (a) – E; Part (b) – E; Part (c) – E.

In part (a) the response correctly identifies the treatments ("the placebo and the real, new drug") satisfying component 1, correctly identifies the experimental units ("each person in the experiment") satisfying component 2, and correctly identifies the response variable ("the improvement in acne severity"), satisfying component 3. Part (a) was scored essentially correct (E).

In part (b) the response provides several advantages of a matched-pairs design, including "minimizing variability amongst individuals" and "enabling a comparison of the effectiveness of the new drug vs. placebo." The second advantage pertains to an inference about the effectiveness of the drug, satisfying component 1. The response indicates that the matched-pairs design is better by contrasting the matched pairs design with the completely randomized design ("while a completely randomized design … The matched pairs design ensures") and stating that it "minimizes variability" in comparison to a completely randomized design, satisfying component 2. By using terms such as "twin" and "drug," the response is in context, satisfying component 3. Part (b) was scored essentially correct (E).

In part (c) the response uses a random process (randomly generating an integer between 0 and 1) that gives each twin a 50 percent probability of getting the drug and a 50 percent probability of getting the placebo, satisfying component 1. The response specifies how to use the random process to assign twins to treatments ("Choose one twin … If the twin receives a '0,' assign to them the placebo and their twin the new drug. If the twin gets a '1,' they are assigned to the new drug and their twin the placebo"), satisfying component 2. The response indicates that the random assignment process will be completed for each pair of twins ("Repeat for each twin pair"), satisfying component 3. Part (c) was scored essentially correct (E).

Question 2 (continued)

Sample: 2B
Score: 2

The response earned the following: Part (a) – P; Part (b) – E; Part (c) – P.

In part (a) the response correctly identifies the treatments ("type of drug received (new or placebo)"), satisfying component 1. The response correctly identifies the experimental units as the "people participating in this experiment," satisfying component 2. The response does not include the word "acne" in the description of the response variable, so component 3 is not satisfied. Part (a) was scored partially correct (P).

In part (b) the response provides two advantages of a matched-pairs design: "reducing the effect of initial acne severity as a confounding variable" and stating that the matched-pairs design "is more effective to determine the drug's effectiveness." The response also provides a disadvantage of a completely randomized design ("make it harder to determine how much of an effect the drug actually had"). The second advantage of the matched-pairs design and the disadvantage of the completely randomized design both pertain to an inference about the effectiveness of the drug, each satisfying component 1. The response indicates that the matched-pairs design is better by contrasting it to a completely randomized design ("as opposed to"), satisfying component 2. By using terms such as "twin," "drug," and "acne," the response is in context, satisfying component 3. Part (b) was scored essentially correct (E).

In part (c) the response then uses a random process ("flip a fair coin") that gives each twin a 50 percent probability of getting the drug and a 50 percent probability of getting the placebo, satisfying component 1. The response does not specify how to use the random process to assign twins to treatments. In other words, there is no description of how to match the outcome of the coin flip with the assignment of a treatment to a specific twin. Component 2 is not satisfied. The response indicates that the random assignment process will be completed for each pair of twins ("For each pair"), satisfying component 3. Part (c) was scored partially correct (P).

Sample: 2C
Score: 1

The response earned the following: Part (a) – P; Part (b) – P; Part (c) – I.

In part (a) the response correctly identifies the treatments ("new drug or placebo") satisfying component 1. The response does not correctly identify the experimental units as the individual people in the experiment ("pairs of identical twins"), so component 2 is not satisfied. The response correctly identifies the response variable ("improvement of acne"), satisfying component 3. Part (a) was scored partially correct (P).

In part (b) the response provides an advantage of a matched-pairs design ("more accurate results") and a disadvantage of a completely randomized design ("not lead to fair and accurate results.") However, these advantages do not clearly pertain to an inference made after collecting the data. For example, "accurate results" could be describing the accuracy of the data collected and not a conclusion about the effectiveness of the drug. Component 1 is not satisfied. The response does indicate that the matched-pairs design is better by using a comparative word ("more") and contrasting the matched-pairs design to a completely randomized design ("if this were a completely randomized design"), satisfying component 2. The response is in context ("twin," "severity of acne"), satisfying component 3. Part (b) was scored partially correct (P).

In part (c) the response indicates that each individual will receive both treatments ("I would take each individual from the pair of twins and give them both treatments"). Because the response does not randomly assign the two treatments within pairs of twins, with one twin receiving each treatment, part (c) was scored incorrect (I).