AP Released Items:

1. The dentists in a dental clinic would like to determine if there is a difference between the number of new cavities in people who eat an apple a day and in people who eat less than one apple a week. They are going to conduct a study with 50 people in each group.

Fifty clinic patients who report that they routinely eat an apple a day and 50 clinic patients who report that they eat less than one apple a week will be identified. The dentists will examine the patients and their records to determine the number of new cavities the patients have had over the past two years. They will then compare the number of new cavities in the two groups.

a. Why is this an observational study and not an experiment?

b. Explain the concept of confounding in the context of this study. Include an example of a possible confounding variable.

c. If the mean number of new cavities for those who ate an apple a day was statistically significantly smaller than the mean number of new cavities for those who ate less than one apple a week, could one conclude that the lower number of new cavities can be attributed to eating an apple a day? Explain.
2.

High cholesterol level in people can be reduced by exercise or by drug treatment. A pharmaceutical company has developed a new cholesterol-reducing drug. Researchers would like to compare its effects to the effects of the cholesterol-reducing drug that is currently available on the market. Volunteers who have a history of high cholesterol and who are currently not on medication will be recruited to participate in a study.

(a) Explain how you would carry out a completely randomized experiment for the study.

(b) Describe an experimental design that would improve the design in (a) by incorporating blocking.

(c) Can the experimental design in (b) be carried out in a double blind manner? Explain.
3.

Students are designing an experiment to compare the productivity of two varieties of dwarf fruit trees. The site for the experiment is a field that is bordered by a densely forested area on the west (left) side. The field has been divided into eight plots of approximately the same area. The students have decided that the test plots should be blocked. Four trees, two of each of the two varieties, will be assigned at random to the four plots within each block, with one tree planted in each plot.

The two blocking schemes shown below are under consideration. For each scheme, one block is indicated by the white region and the other block is indicated by the gray region in the figures.

Blocking Scheme A

Blocking Scheme B

Key

- Grey Block 1
- White Block 2

(a) Which of the blocking schemes, A or B, is better for this experiment? Explain your answer.

(b) Even though the students have decided to block, they must randomly assign the varieties of trees to the plots within each block. What is the purpose of this randomization in the context of this experiment?
4.

A preliminary study conducted at a medical center in St. Louis has shown that treatment with small, low-intensity magnets reduces the self-reported level of pain in polio patients. During each session, a patient rested on an examining table in the doctor’s office while the magnets, embedded in soft pads, were strapped to the body at the site of pain. Sessions continued for several weeks, after which pain reduction was measured.

A new study is being designed to investigate whether magnets also reduce pain in patients suffering from herniated disks in the lower back. One hundred male patients are available for the new study.

(a) Describe an appropriate design for the new study. Your discussion should briefly address treatments used, methods of treatment assignment, and what variables would be measured. Do not describe how the data would be analyzed.

(b) Would you modify the design above if, instead of 100 male patients, there were 50 male and 50 female patients available for the study? If so, how would you modify your design? If not, why not?
5.
A manufacturer of boots plans to conduct an experiment to compare a new method of waterproofing to the current method. The appearance of the boots is not changed by either method. The company recruits 100 volunteers in Seattle, where it rains frequently, to wear the boots as they normally would for 6 months. At the end of the 6 months, the boots will be returned to the company to be evaluated for water damage.

(a) Describe a design for this experiment that uses the 100 volunteers. Include a few sentences on how it would be implemented.

(b) Could your design be double blind? Explain.
There have been many studies recently concerning coffee drinking and cholesterol level. While it is known that several coffee-bean components can elevate blood cholesterol level, it is thought that a new type of paper coffee filter may reduce the presence of some of these components in coffee.

The effect of the new filter on cholesterol level will be studied over a 10-week period using 300 nonsmokers who each drink 4 cups of caffeinated coffee per day. Each of these 300 participants will be assigned to one of two groups: the experimental group, who will only drink coffee that has been made with the new filter, or the control group, who will only drink coffee that has been made with the standard filter. Each participant’s cholesterol level will be measured at the beginning and at the end of the study.

(a) Describe an appropriate method for assigning the subjects to the two groups so that each group will have an equal number of subjects.

(b) In this study, the researchers chose to include a group who only drank coffee that was made with the standard filter. Why is it important to include a control group in this study even though cholesterol levels will be measured at the beginning and at the end of the study?

(c) Which test would you conduct to determine whether the change in cholesterol level would be greater if people used the new filter rather than using the standard filter?

(d) Why would the researchers choose to use only nonsmokers in the study?
Researchers who are studying a new shampoo formula plan to compare the condition of hair for people who use the new formula with the condition of hair for people who use the current formula. Twelve volunteers are available to participate in this study. Information on these volunteers (numbered 1 through 12) is shown in the table below.

<table>
<thead>
<tr>
<th>Volunteer</th>
<th>Gender</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Male</td>
<td>21</td>
</tr>
<tr>
<td>2</td>
<td>Female</td>
<td>20</td>
</tr>
<tr>
<td>3</td>
<td>Male</td>
<td>47</td>
</tr>
<tr>
<td>4</td>
<td>Female</td>
<td>60</td>
</tr>
<tr>
<td>5</td>
<td>Female</td>
<td>62</td>
</tr>
<tr>
<td>6</td>
<td>Male</td>
<td>61</td>
</tr>
<tr>
<td>7</td>
<td>Male</td>
<td>58</td>
</tr>
<tr>
<td>8</td>
<td>Female</td>
<td>44</td>
</tr>
<tr>
<td>9</td>
<td>Male</td>
<td>44</td>
</tr>
<tr>
<td>10</td>
<td>Female</td>
<td>24</td>
</tr>
<tr>
<td>11</td>
<td>Male</td>
<td>23</td>
</tr>
<tr>
<td>12</td>
<td>Female</td>
<td>46</td>
</tr>
</tbody>
</table>

(a) These researchers want to conduct an experiment involving the two formulas (new and current) of shampoo. They believe that the condition of hair changes with age but not gender. Because researchers want the size of the blocks in an experiment to be equal to the number of treatments, they will use blocks of size 2 in their experiment. Identify the volunteers (by number) that would be included in each of the six blocks and give the criteria you used to form the blocks.

(b) Other researchers believe that hair condition differs with both age and gender. These researchers will also use blocks of size 2 in their experiment. Identify the volunteers (by number) that would be included in each of the six blocks and give the criteria you used to form the blocks.

(c) The researchers in part (b) decide to select three of the six blocks to receive the new formula and to give the other three blocks the current formula. Is this an appropriate way to assign treatments? If so, describe a method for selecting the three blocks to receive the new formula. If not, describe an appropriate method for assigning treatments.
At a certain university, students who live in the dormitories eat at a common dining hall. Recently, some students have been complaining about the quality of the food served there. The dining hall manager decided to do a survey to estimate the proportion of students living in the dormitories who think that the quality of the food should be improved. One evening, the manager asked the first 100 students entering the dining hall to answer the following question.

Many students believe that the food served in the dining hall needs improvement. Do you think that the quality of food served here needs improvement, even though that would increase the cost of the meal plan?

_____ Yes  _____ No  _____ No opinion

(a) In this setting, explain how bias may have been introduced based on the way this convenience sample was selected and suggest how the sample could have been selected differently to avoid that bias.

(b) In this setting, explain how bias may have been introduced based on the way the question was worded and suggest how it could have been worded differently to avoid that bias.
Answers:
1.

Holistic grading is essential for evaluating student responses to this question. In judging the responses for each individual question as correct, weak, or incorrect, the total response to all questions (a), (b), and (c) should be read and evaluated. If the reader can construct an unambiguous correct response to (a), for example, from the student's total response, then question (a) should be counted as "correct."

Solution
This problem has three parts.

a. The student can appeal to any of three reasons in judging this study not an experiment:
   1. there is no random assignment of subjects to treatments;
   2. there are no treatments imposed;
   3. existing data is being used.

b. Two variables are confounded if their effect on the number of new cavities cannot be distinguished from one another. The student must mention not only that the confounding variables may affect the outcome but that they have differential effects within the two groups. For instance: confounding would occur if patients who eat an apple a day differ from those who eat less than one apple a week on some variable that is related to dental health. In this example, diet or general level of health are examples of what might be confounding variables. For example, it is possible that people who eat an apple a day are more nutrition conscious and have a more healthy diet in general than those who eat one or fewer apples per week, and this might explain the observed difference in dental health.
Solution

(a) Describes an experimental design that includes:
1. Random assignment of volunteers to the treatment groups
2. Identification of treatment groups as old drug and new drug
3. Indication that a comparison or measurement of cholesterol levels should be made

OR

The student may give a detailed diagram that addresses the three parts:

1. Random assignment of subjects
   - Group 1
   - Group 2
2. Treatment 1
   - 2. (old drug)
3. Compare Cholesterol Levels
   - 2. (new drug)

Note: In part (a), it is incorrect to use the terminology “treatment” and “placebo” for the treatment groups. It is considered correct to use “old drug” and “new drug”, and “placebo,” if a third group is used, for the treatment groups.

(b) Describes an experimental design that includes:
1. Creating blocks based on level of exercise or cholesterol level, or creating blocks using age, diet, gender, or any other factor plausibly related to cholesterol level with explanation (i.e., block on gender because males and females may respond differently)
2. Random assignment of subjects to treatments within blocks

OR

The student may give a detailed diagram that addresses the two parts as long as the blocking factor is described.

Note: No credit will be given in part (b) if a student does not use blocking in his/her design even though they randomize correctly.

(c) Clearly explains a double blind experiment—neither the subjects nor those administering the drugs or monitoring results know which of the two drugs is being used.

An answer of yes without explanation receives no credit.

An answer of no could receive credit if the design described in part (b) does not allow for double-blinding.
3.

**Part (a):**

Blocking scheme A is preferable because it creates homogeneous blocks with respect to forest exposure. That is, plots in the same block have similar exposure to the forest.

**Part (b):**

Randomization of varieties of trees to the plots within each block should reduce any possible bias due to confounding variables, such as fertility or moisture, on the productivity of the two types of dwarf trees.

OR

Randomization of varieties of trees to the plots within each block should even out (or equalize) the effect of other characteristics of the plots that might be related to the productivity of the trees.
Part (a):

1. Two treatments: magnets and no magnets (or magnets and placebo). Subjects in the no magnet group would be handled in the same way as the magnet group, but there would be no magnets embedded in the pads used.
2. There must be random assignment of subjects to treatments (or treatments to subjects). How the randomization would be carried out does not need to be specified, but it must be clear what is being randomized.
3. Variable measured: Self-reported level of pain or reduction in pain.

The design may be described by a diagram, but the treatments and the variable measured must be included and the randomization must be very clear.

Part (b): Either one of the following approaches is acceptable.

1. Saying yes and indicating how they would alter the design: Separating the subjects into the two gender groups and then randomizing subjects to treatments within each group. This may also be described using a diagram, as shown below, but the blocking factor and randomization must be clearly indicated.

OR

2. Saying no and describing why. For example, indicating that the randomization in (a) should equalize the effects of gender in the two groups or assuming gender does not have a strong effect and since the sample size is large

OR providing a good explanation for why gender does not have a strong differential effect on the outcome.
Part (a):

A *paired design* is used in which each subject receives a pair of boots where one boot is treated with the new method and the other with the current method. Subjects should be randomly assigned to one of two groups. Group 1 would have the new method applied to the right boot; group 2 would have the new method applied to the left boot. OR For each subject, whether the new method is applied to the right or left boot is determined at random.

OR

A *crossover design* is used in which each subject receives a pair of boots, both of which were treated with one treatment. The boots are used for three months and then exchanged for a second pair of boots, both of which were treated with the other treatment. These boots are then used for the next three months. Subjects should be randomly assigned to one of two groups. One group receives boots with the new treatment first and the other group receives boots with the current method first.

NOTE: Additional appropriate blocking schemes are considered extraneous.

Part (b):

The design could be double blind, as long as both the *subjects* and the person *evaluating* the boots for water damage do not know which boots were treated with the new method and which were treated with the current method.

NOTE: If the student does something unexpected in part (a) and gives a design that actually cannot be double blind, then part (b) could be considered correct provided the response explains why the design could not be double blind.
6.

Part (a):

Assign each subject a number from 001 to 300 and then use a random number table or a random number generator to select 150 of the 300 for the new filter group. The other 150 would be assigned to the standard filter group.

OR

For each subject, flip a coin. If the coin lands H, assign the subject to the new filter group; otherwise assign the subject to the standard filter group. Continue in this way until one of the groups has 150 subjects. Assign all remaining subjects to the other group.

Part (b):

Without a comparison group, the cholesterol level could change overall, but we would not be able to determine whether the observed change was due to some other extraneous variable that changed during the 10-week period. For example, diet might change with time of the year, and the diet might result in changes in cholesterol changes. So a change in cholesterol would not be attributable to the new coffee filter. The addition of a control group enables the researchers to assess the mean change in cholesterol level due to the coffee filter, as opposed to just determining if the cholesterol level changed. The control group eliminates the confounding variable of another change that might have occurred over the 10-week period.

Part (c):

The two-sample \( t \) test for means or mean differences would be used (or the two-sample \( z \) test for means).

Part (d):

If it is known that smoking is related to changes in cholesterol level, it would be best to control for smoking by using only nonsmokers. This eliminates smoking as a source of variability, creating more homogenous groups, enabling more direct comparisons between the treatment and control groups and more precise estimates of the treatment effects (though we will only be able to generalize the results to nonsmokers).
Part (a):

<table>
<thead>
<tr>
<th>Block</th>
<th>Volunteers</th>
<th>Ages</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1, 2</td>
<td>20, 21</td>
</tr>
<tr>
<td>2</td>
<td>10, 11</td>
<td>23, 24</td>
</tr>
<tr>
<td>3</td>
<td>8, 9</td>
<td>44, 44</td>
</tr>
<tr>
<td>4</td>
<td>3, 12</td>
<td>46, 47</td>
</tr>
<tr>
<td>5</td>
<td>4, 7</td>
<td>58, 60</td>
</tr>
<tr>
<td>6</td>
<td>5, 6</td>
<td>61, 62</td>
</tr>
</tbody>
</table>

Since these researchers believe that the condition of hair changes with age but not gender, the volunteers are sorted from youngest to oldest. The volunteers in the sorted list are paired to form six blocks of size two. More specifically, the youngest two volunteers are placed in the first block. The next two volunteers in the sorted list are placed in the second block. This pairing continues until all six blocks of two are formed, with the oldest two volunteers in the sixth block.

Part (b):

<table>
<thead>
<tr>
<th>Block</th>
<th>Volunteers</th>
<th>Ages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female 1</td>
<td>2, 10</td>
<td>20, 24</td>
</tr>
<tr>
<td>Female 2</td>
<td>8, 12</td>
<td>44, 46</td>
</tr>
<tr>
<td>Female 3</td>
<td>4, 5</td>
<td>60, 62</td>
</tr>
<tr>
<td>Male 1</td>
<td>1, 11</td>
<td>21, 23</td>
</tr>
<tr>
<td>Male 2</td>
<td>3, 9</td>
<td>47, 44</td>
</tr>
<tr>
<td>Male 3</td>
<td>6, 7</td>
<td>61, 58</td>
</tr>
</tbody>
</table>

Since these researchers believe that the condition of hair changes with both age and gender, the women are sorted from youngest to oldest and then the men are sorted from youngest to oldest. The women (men) in the sorted list are paired to form the blocks of size two. More specifically, the youngest two women (men) are placed in a block. The next two youngest women (men) are placed in another block. Finally, the oldest two women (men) are placed in another block.

Part (c):

No, the researchers in part (b) should not randomly select three blocks to receive the new formula and then give the current formula to the other three blocks. They blocked on both age and gender to form homogeneous groups because they believe hair condition differs with both age and gender. Giving the youngest or oldest women (men) the same formula defeats the purpose of blocking. In a block design, randomization should be carried out separately within each block. That is, for each block, two random numbers are generated (via a random number generator or a table of random digits) and assigned to the two volunteers. The volunteer with the smallest random number is given the new formula and the other volunteer is given the current formula.
8.

**Part (a):**

Since the manager used a convenience sample (the first 100 students entering the cafeteria), bias may have been introduced. Students who arrive at the cafeteria early may have opinions of food quality that differ in some important way from other students who live in the dormitories.

This bias could be avoided by selecting a random sample of 100 dormitory residents instead of just asking the first 100 students entering the cafeteria.

**Part (b):**

The way the question is worded may be leading. The first part of the question included a statement that many students think the food needs improvement. This may lead people to support this view by responding that the food does need improvement.

The inclusion of the phrase "even though that would increase the cost of the meal plan" may lead students to say the food is OK only because they do not want to pay more.

A better wording might be to simply ask "Do you think that the quality of the food served in the cafeteria needs improvement?"