A = True, B = False unless stated otherwise name (required) _____

You must turn in both this hard copy (with your name on it) and your scantron to receive credit for this exam.

One answer and only one answer per question. Leaving a question blank or filling in 2+ answers will be incorrect no matter what. When you are given a list of options for a set of questions, some options may not be a correct answer for any of the questions.

Correlations, Causation & Hidden variables

1-7 Recall the hypothetical table in which each cell gives the accident rate per 1000 cars of that type per year. Also remember that this table does not give the number of vehicles that occur in each cell. There are now 4 different tables, and you are asked to indicate which tables (could) have a specific property. Ignore the possibility of other variables (besides color and type of car) affecting accident rates – assume the tables give you all relevant variables.

Table		W		X		Y		Z	
		Type of car		Type of car		Type of car		Type of car	
		risky	safe	risky	safe	risky	safe	risky	safe
Car color	**red**	**5**	**6**	**5**	**5**	**10**	**2**	**4**	**2**
	not red	4	2	5	5	5	4	3	1

1-4 (8pts) Which tables indicate that:

(A) red always has the higher accident rate when controlling for type

(B) red never has the higher accident rate when controlling for type

(C) red has the higher accident rate for one type of car but not the other

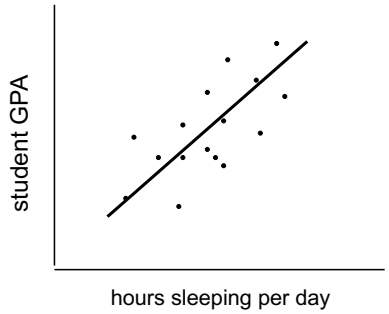
1. (A)(B)(C) Table W **2.** (A)(B)(C) Table X **3.** (A)(B)(C) Table Y **4.** (A)(B)(C) Table Z

5-7 (6pts) For which tables could the overall correlation between accident rate and color go in the opposite direction from the effect of red when controlling for type of car (Simpson's paradox)?

(A) Could go in opposite direction
(B) could not go in opposite direction

5. (A)(B) Table W **6.** (A)(B) Table X **7.** (A)(B) Table Z

8-12 (10 pts) Consider the following graph of data, showing that students who sleep longer also have higher GPAs in college; each point is the average for a different university. Which of the following models can be rejected?



(A) Can be rejected (B) Cannot be rejected

8(A)(B) GPA is positively correlated with sleep

9 (A)(B) GPA is negatively correlated with sleep

10 (A)(B) Sleeping longer lowers GPA

11 (A)(B) Sleeping shorter raises GPA

12 (A)(B) Sleeping longer has no effect on GPA

13-16 (8 pts) The average tooth decay of people living in different towns decreases with the amount of fluoride in the town's water supply (which comes from groundwater). Which of the following models use(s) a 3rd variable to explain the cause of this correlation? *One option refers to enamel -- enamel is a mineralized protein in teeth.*

A = 3rd variable is used, B = no 3rd variable

Choose (A) if third variable is used	Causal model
13. (A)(B)	Teeth with harder enamel have less tooth decay. Fluoride in the diet – including water consumed – causes enamel to be harder.
14. (A)(B)	Fluoride inhibits the growth of bacteria in the mouth that cause tooth decay. With fewer decay-causing bacteria in the mouth, there is less tooth decay.
15. (A)(B)	High fluoride in the water comes from underground rock formations that also have oil. Thus towns with fluoride in the water also have high levels of income from oil revenue. The higher the local income levels provide better dental care, and better dental care lowers tooth decay.
16. (A)(B)	Water with more fluoride also has more magnesium. Tooth decay is reduced by higher levels of magnesium.

17-22. **(12 pts)** Which of the following options is indicated? Base your answer only on the information provided.

(A) no correlation or causation is described.
(B) causation only – the statement describes one or more causal models with no correlation
(C) correlation only – the statement merely describes one or more non-zero correlations,
(D) correlation is used to infer/argue causation (i.e., a correlation leads people to infer the causal basis of the correlation)
(E) causation is used to explain a correlation (both correlation and causation must go in the same direction)
(F) correlation and causation are described but go in opposite directions (Simpson's paradox)

17. (A)(B)(C)(D)(E)(F) Young drivers have higher insurance rates than old drivers

18 (A)(B)(C)(D)(E)(F) A study reveals that athletic teams wearing red uniforms have higher winning rates than teams wearing other colors. To offset a losing streak, the Guthrie Panthers change their uniform from green to red.

19 (A)(B)(C)(D)(E)(F) In attempting to correct decades of gender discrimination, the University of Idalou intentionally accepts a higher fraction of women than of men in each of its programs. Despite this effort, the average acceptance rate of women to the University still remains less than that of men.

20. (A)(B)(C)(D)(E)(F) Eating a high-carbohydrate diet causes heart disease. Frequent exercise reduces heart disease.

21. (A)(B)(C)(D)(E)(F) Students who study abroad have higher graduation rates than students who don't. Despite this, studying abroad lowers graduation rates.

22. (A)(B)(C)(D)(E)(F) College football team winning rates are higher for coaches with high salaries. Coach salaries are higher for teams with higher winning rates.

Controls, controlled variables and experiments

23-24 (2.5 pts each) *We observe a correlation between two variables where we suspect a possible causal relationship*. Why do we try to control for hidden/3rd variables in determining whether X causes Y or Y causes X? **(A) True** (B) False

23. (A)(B) The 3rd variables can be the actual cause of the correlation. If a 3rd variable is causal and we successfully control for it, then the correlation between X and Y will disappear and will no longer be considered as possibly causal.

24. (A)(B). 3rd variables can be imbalanced in their associations between X and Y. We want to control for 3rd variables to destroy those imbalances

25-32. (10 pts) You wish to test claims that an expensive study aid improves student performance in class. This aid is a tasteless dietary supplement. You also have a tasteless alternative that is not expected to have an effect. You have 100 students (half of which are male, half female). At the beginning of the semester, you tell all of them that you are giving them a performance enhancer that is a food additive; the additive is a pill they will take every day for a month. Students are led to believe that everyone gets the same pill, and they cannot taste the difference.

Design 1. You give the true aid to all the women and the alternative to all the men. After a month you give a test to compare performance scores between random subsets of half the male and half the female students.

Which are true of Design 1? (**A = TRUE**, B = false)

25 (A)(B) The study controls for the study aid

26 (A)(B) The study included blind

27 (A)(B) Except for gender, the study controls for all possible third variables

28 (A)(B) The study controls for gender

29 (A)(B) The study manipulates student expectation of performance as a treatment variable.

30-32.

Design 2. You randomly assign the true additive to half the class, the alternative to the other half. After a month, you compare performance scores those given the supplement and those given the tasteless alternative.

Which are true of Design 2? (**A = TRUE**, B = false)

30 (A)(B) The study controls for gender

31 (A)(B) The study controls for the study aid

32 (A)(B) The study controls for all possible third variables, including those differing between men and women

33-36. (**9 pts**) Jules is testing the effect of different mineral supplements on goldfish color brightness. He mixes different combinations of supplements that are added to the water in which the fish live. The different supplements are denoted 1, 2, 3, 4, 5. A + indicates the supplement is present in the water, - is absent. He has no expectation of whether these supplements will have an effect, and if they do, whether they will increase or decrease brightness. He then finds out how bright the fish are 4 weeks after being kept in the water; average brightness (B) is given in the right-most column; you only know that brightness is measured, not its value. Which statements in the following questions are true?

	mineral supplement					**Brightness**
Mix	1	2	3	4	5	
(A)	-	-	-	-	+	B_A
(B)	-	-	-	+	+	B_B
(C)	-	-	+	+	+	B_C
(D)	-	+	+	+	+	B_D
(E)	+	-	-	-	-	B_E
(F)	+	+	+	+	-	B_F
(G)	+	+	+	+	+	B_G
(H)	-	-	-	-	-	B_H

A = TRUE, B = false

33. (A) (B) Only one pair of mixes allows you to assess the effect of supplement 2 when all other supplements are controlled.

34. (A)(B) The fish in mix (G) are expected to be brighter than the fish in mix (H) – because (G) received all the supplements.

35. (A) (B) A comparison of fish exposed to mix (B) with those exposed to mix (C) controls for 4 of the 5 supplements.

36. (A) (B) The average brightness of fish exposed to one mix is always expected to differ from the average brightness of fish exposed to a different mix.

37-41 (10 pts). Which of the following studies describe(s) experiments, regardless of whether the experiment was designed well or poorly and regardless of ethics? In each problem, the goal is given. The question is: does the option describe an experiment for the goal.

(A) = is an experiment (B) is not

37. (A)(B) You normally eat meat, but for the next 6 weeks you fail to eat meat because you are on holiday with a vegetarian family who only feeds you plants, and you do not wish to offend this family.

38. (A)(B) You randomly assign exam grades to each student in your class to see if they are paying attention to their scores.

39 (A)(B) You compare cancer rates between people who have cell phones and people who lack cell phones to see if cell phones cause cancer.

40. (A)(B) A researcher compares per capita chocolate consumption with the per capita rate of millionaires across different cities. To their surprise, the data show that the rate at which people are millionaires increases with chocolate consumption.

41. (A)(B) A psychic who makes predictions according to standard protocol gets confused and starts giving false predictions.

42. (**3 pts**) **Prisoners of Silence video (FC = facilitated communication).** *The video showed tests of FC suggesting that the facilitator, not the child, was the author of the typed responses.* The following questions require you to address and interpret the features of this experiment.

(A) The test was NOT an experiment of the type in which the relevant 3rd variable(s) was/were known in advance.

(B) The test WAS an experiment in which the relevant 3rd variable was known in advance: Facilitator knowledge of the answers was the relevant 3rd variable

(C) The test WAS an experiment in which the relevant 3rd variable was known in advance: Child knowledge of the answers was the relevant 3rd variable

(D) The test WAS an experiment in which the relevant 3rd variable was known in advance: Observer influence was the relevant 3rd variable

43. (**3p**ts) **Prisoners of Silence.** What design feature was specifically required to enable the manipulation critical to this test of FC?

(A) Controls (B) Replication (C) Blind (D) Standards

44. (3pts) James Randi video: Randomization was apparently absent in the horoscope test. What can we say about the benefit of randomization in this study -- randomly assigning horoscopes to students?

(A) Randomization was not relevant in this experiment

(B) Full randomization (to all students) would have enabled a clearer interpretation of the results.

(C) Randomization could have been applied – was relevant -- but probably would not have made any difference

(D) Randomization would best have been applied to half the class but not the other half – to serve as a control

45-47. (6pts) In a properly controlled horoscope experiment (otherwise similar to that shown), if the goal is to compare the two models

Model (1) students respond to personalized accuracy versus

Model (2) students respond to the expectation of personalized accuracy,

what design feature(s) specifically allow(s) you to control for the expectation of personalized accuracy given the way the experiment was done?

(A) controls for expectation of personalized accuracy **(B) does not** specifically control

45. (A)(B) Replication **46. (A)(B)** Randomization **47. (A)(B)** Blind

48-52 (10pts) We want to control for as many 3rd variables as possible in asking whether smoking causes cancer. What can we say about the following potential study designs? Questions apply to the design immediately above them. **(A) = True**

Design 1: From thousands of people for which you have histories of cancer and smoking or not smoking (by their choice), compare cancer rates in a randomly chosen subset of smokers to cancer rates in a randomly chosen subset of non-smokers.

48 (A)(B) The randomization controls for third variables that may differ between smokers and non-smokers

49 (A)(B) The extent to which the random data control for 3rd variables between smokers and non-smokers depends on how big the randomly chosen subsets are.

Design 2: Now imagine that, for the people for which you have histories of smoking, cancer, you also have information on lifestyles, gender and age. You compare cancer rates between smokers and non-smokers when controlling for all recorded characteristics.

50 (A)(B) This approach would control for all possible unwanted third variables.

51 (A)(B) This approach is correlational.

52 (A)(B) Randomly choosings subsets of smokers and of non-smokers combined with an analysis of differences in recorded characteristics would allow you to discover if the causal variable differed between smokers and non-smokers.

53. (2 pts) (C) **Key code, name, and ID number**. Fill in (**C**) in scantron question 53 to indicate your key for this version of the exam. Be sure your name and EID number are correctly bubbled in on the scantron.