What's new: Under the new, online delivery of tests and answers, you should deliver your answers by the end of exam time to the class email (cors236@uidaho.edu) or to BbLearn (as a short answer to the quiz labeled 'exam 3 answers').

Begin your set of answers with:

Last name, First name, last 4 digits of your Vandal ID.

Then provide your answers separated by commas. You may do this in any of various formats, but keep the answers in order:

A, A, F, H, … for all 63 answers, or

1. A, 2. A, 3. F, 4. H, … or even

1-5: A, A, F, H, C 6-10: B, B, B, B, A

I don't necessarily care how you provide the answers, other than putting commas between them, but I recommend some occasional numbering to be sure you have not gotten your answers out of order or skipped some.

Note: you will be providing 63 answers.

Italicized phrases and sentences should be considered true.

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.

What follows in red text are explanations for how to reason the answers. After reading the explanations, if you are not sure of what the answer is, refer to the key.

Language and concepts

1-5. (8 pts). (Evidence of absence vs. absence of evidence) Which statements either:

(A) Can legitimately be made only if we have some data to support a conclusion or reject some models

(B) Can legitimately be made in the absence of data

(disregard the fact that some of these statements are not true)

1. (A)(B) Masks are not effective against the new corona virus

This statement requires some data because it indicates we have ruled out/rejected masks being effective.

2. (A)(B) There is no evidence that genetically modified plants are unsafe to eat

The 'no evidence' gives it away – we can make the statement if we have no evidence. The statement may MAKE YOU FEEL as if it is OK to eat GMO plants, and there MIGHT BE data, but the statement can be made if you have no evidence either way.

3. (A)(B) There is no evidence that genetically modified plants are safe to eat

The 'no evidence' gives it away – we can make the statement if we have no evidence. Ditto as in (2).

4.(A)(B) Psychics have no predictive power.

This statement requires some data because it indicates we have ruled out/rejected psychics having predictive power.

5.(A)(B) Despite extensive testing, we have observed no cases of covid-19 in Latah County

The statement basically says data exist, so you would not make the statment without the data from 'extensive testing.'

6-8. **(5 pts)** A few days ago, the World Health Organization (WHO) revised its advice about ibuprofen and covid-19 infections to:

"We do not currently believe there is any proven scientific evidence linking over-the-counter use of ibuprofen to the aggravation of COVID-19."

What does this statement mean – how should it be interpreted? **(A) = True** (B) = False

You have to think about the statement (which WHO did make). These questions do not ask if WHO's statement requires data, because WHO probably did have data.

6. (A)(B). Ibuprofen is known to be safe for treating COVID-19

The statement says they don't know of evidence to suggest ibuprofen (IBP) worsens covid. That effectively means they cannot reject the model that IBP is harmless. But, even though you may not be able to reject the model that IBP is harmless, that is not the same as saying it is safe. This is a bit like the legality of marketing herbal remedies without FDA approval – you market without evidence that the herbal remedy is harmful, but it could turn out to be harmful when you do the studies.

7. (A)(B). Ibuprofen is harmful for treating COVID-19

The statement says they don't know of evidence to suggest ibuprofen (IBP) worsens covid, so they certainly are not saying it is harmful.

8. (A)(B). We do not yet know whether ibuprofen is safe, harmful, or has no effect – it could be any if those possibilities.

This is certainly the implication of the WHO statement. The WHO statement rules out nothing.

Correlations, Causation & 3rd variables

9-12. **(7 pts)** Say that it has been observed that people who take high doses of ibuprofen more often die from covid-19 than people who don't take ibuprofen. Which of the following are valid conclusions from this observation? **(A) = True**, (B) = False

This is similar to the following question in that you are given a correlation in the problem preamble. You must then understand what it means.

9. (A)(B). Ibuprofen aggravates (contributes to) death from covid-19 infections

This is a causal model and to assume that it follows from the preamble would be to commit the error of inferring causation from correlation.

10. (A)(B).. Serious covid-19 infections cause people to take ibuprofen

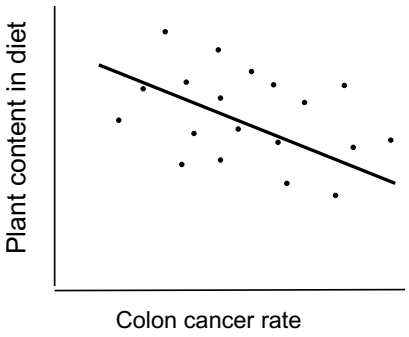
This too is a causal model and to assume that it follows from the preamble would be to commit the error of inferring causation from correlation.

11. (A)(B). We cannot rule out that ibuprofen contributes to death from covid-19 infections

This simply says the preamble cannot be used to rule out a causal model. This is a specific case of avoiding the 'correlation does not imply causation' error, so true (A).

12. (A)(B). We cannot rule out that serious covid-19 infections cause people to take ibuprofen

This also says the preamble cannot be used to rule out a causal model. This is a specific case of avoiding the 'correlation does not imply causation' error, so true (A).



13-16 (8 pts). Across a set of US cities we observe the pattern (shown in the figure) between dietary plant consumption and colon cancer rate. Which models are *consistent* with these data? This question is the same as asking which models cannot be rejected.

A = consistent, thus cannot be rejected, B = not consistent – can be rejected

I covered this type of problem in lecture today (8 III 2021). The graph is a (negative) correlation, and you cannot use correlational data to reject or rule out any causal model. So step 1 is to find the causal models below and realize that they cannot be rejected. Of the remaining (correlational) models, they can be rejected only if the correlation in the problem goes in the opposite direction as the correlation in the figure.

13 (A) (B) Colon cancer rates are increased by eating more plants

Causal model – cannot be rejected.

14. (A) (B) Colon cancer rates are reduced by eating more plants

Causal model – cannot be rejected.

15. (A) (B) Plant consumption has no causal effect on colon cancer rates

Effective a type of causal model – cannot be rejected.

16. (A) (B) Cities with higher dietary plant consumption have older populations than cities with low plant consumption, and age is the cause of colon cancer.

Depending on how far lecture has proceeded with this topic, this probably may or may not make sense. The first part describes a correlation (age and dietary plant content); the second part is a causal model. The second part of the statement, being a causal model, is thus compatible with the correlation in the figure. It is possible that the correlation part of the question might not be compatible with the figure. But the figure shows a correlation between dietary plant and cancer rate, whereas the statement is about a correlation between dietary plant and age. The two compare the same variables, so the figure cannot be used to reject this compound statement.

17-20 (8 pts) The death rate of covid-19 patients is higher in people taking ibuprofen than in people not taking ibuprofen. The preceding sentence describes a correlation. Which of the following models explain(s) the cause of this correlation by some means **other than** ibuprofen causes higher death rates from covid-19? The second statement is asking for a causal model that introduces a third variable to explain the correlation. If class has not covered third variables yet, you would not worry about answering this question.

A = a model in which ibuprofen is NOT the cause of higher death, **B = ibuprofen causes higher death**

Choose (A) if ibuprofen is not the cause of higher death	Causal model
17. (A)(B)	Ibuprofen lowers immunity and thereby makes people more susceptible to dying from any infection, including covid-19. By the way we have defined 'causation,' ibuprofen would be considered the cause. In the model, ibuprofen works through immunity, but it is basically a model that, by taking ibuprofen, you are more likely to die from covid.
18. (A)(B)	Older people are more inclined to take ibuprofen than are young people. Higher death rates from covid-19 are due to age. Here, age is the cause of higher death, not ibuprofen.
19. (A)(B)	The more serious a covid-19 infection becomes, the more likely the patient is to take ibuprofen to alleviate the symptoms. Death rate increases with how serious the infection is, and ibuprofen actually reduces the death rate over what it would be otherwise. This is a model in which ibuprofen and severity are causally linked, but in the wrong direction for the problem – disease severity is the cause of higher death and ibuprofen use.
20. (A)(B)	People who take ibuprofen tend to have underlying health issues. The underlying health issues are what lead to the increase in covid-19 deaths. Ibuprofen use is invoked as caused by the same thing that causes increased covid deaths, so ibuprofen is not the cause.

21-26. **(10 pts)** Which of the following options is indicated? Base your answer only on the information provided.

(A) no correlation or causation is indicated.

(B) correlation only – the statement merely describing one or more non-zero correlations,

(C) correlation and causation are described but go in opposite directions (Simpson's paradox)

(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)

21. (A)(B)(C)(D)(E) Athletic teams whose uniforms are red have higher win rates than teams whose uniforms are not red.

describes data and 2 variables – winning rates and different athletic teams (or you could think of this as 1 variable and different groups). Since winning rates are different for red than for other colors, you have a correlation. There is no causal statement.

22 (A)(B)(C)(D)(E) A company that makes face masks finds that people who wear masks have lower infection rates of covid-19 than people who do not wear masks. = a correlation People who learn of this fact buy masks to protect themselves. = a case of people responding to the correlation as if causal.

23 (A)(B)(C)(D)(E) Studying improves a student's exam scores = causation. Yet students who study more have lower exam scores than students who study less = data, and a correlation. In this case, the correlation goes in one direction, causation in the other – Simpson's paradox.

24. (A)(B)(C)(D)(E)) Having a pet at home increases the survival rates of the elderly = causation (now underlined, but not in the original test). Consequently, old people who have pets live somewhat longer than old people without pets = a correlation. The use of 'consequently' means that the causal model is used as the basis of the correlation.

25. (A)(B)(C)(D)(E) People who eat lots of sugar have high levels of tooth decay. People who avoid dental checkups have high levels of tooth decay. People who eat sugar and avoid dental checkups have the highest levels of tooth decay. Three statements of correlation – all are data with no reason suggested (your dentist would probably infer causation behind these behaviors, but the problem offers no causation).

26. (A)(B)(C)(D)(E) 50% of BSU students attend their university's football games. One variable, one population. Not enough for a correlation to possibly exist.

Controls, controlled variables and Experiments

27-28 (2.5 pts each) When we observe a correlation between two variables (X, Y) where we suspect a possible causal relationship, *why* do we try to control for 3rd variables in determining whether X causes Y or Y causes X? **(A) True** (B) False

27. (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.

28. (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 and thus destroy the ability of 3rd variables to influence the correlation.

29-34. (10 pts) An associate of yours markets a dietary supplement for weight loss. You consider the following experimental design to test whether it works. You are asked about which variables are controlled or manipulated.

Which are true (A = TRUE, B = false)

Design. For 200 overweight male patients enrolled in a diet study, all are put on a vegetarian diet and told that the diet will enable them to lose 30 pounds. In addition to the diet, you will give the patients either the supplement in a capsule/pill or an empty capsule indistinguishable from the pill but which has no supplement. You randomly assign the supplements versus empty capsules to your 200 patients and tell the patients all capsules are the same digestive aid. You compare weight change among the two groups after 2 months to see if there is an effect of the supplement.

Which are true of this Design with respect to testing the effect of the supplement?

29 (A)(B) It controls for vegetarian diet

30 (A)(B) It controls for body weight of the patient

31 (A)(B) It controls for educational background of the patient

32 (A)(B) It controls for patient expectation of weight change.

33 (A)(B) Patient expectation of weight change is a treatment variable.

34 (A)(B) Supplement is a treatment variable.

35-39. (**9 pts**) Holly is testing the effect of different chemical additives on tomato production/output. She mixes different combinations of chemicals and adds them to the soil of the plants. The different chemical additives are calcium, nitrogen, phosphate, and potassium. The different plants are grown separately to be sure that the chemicals given one plant do not leach to other plants. Different mixes are indicated by rows (A)-(H) in the following table, a + indicating the additive is present in the mix, a - indicating absence. There is no *a priori* expectation that these additives will have any effect on tomato production. After 2 months growth, she measures plant output given each mix; output is given in the right-most column. Which statements in the following questions are true?

A = TRUE, B = false

		chemical additive				
Mix		calcium	nitrogen	phosphate	potassium	**plant output**
	(A)	+	-	+	-	O_A
	(B)	-	+	-	+	O_B
	(C)	-	+	+	+	O_C
	(D)	-	-	-	-	O_D
	(E)	+	+	+	+	O_E
	(F)	+	+	-	+	O_F
	(G)	-	+	+	+	O_G
	(H)	+	+	-	-	O_H

35. **(A) (B)** The output of plants given mix (E) is expected to be higher than the output of plants given mix (D).

36. **(A) (B)** A comparison of output of plants given mix (A) with plants given mix (B) controls for all additives.

37. **(A) (B)** A comparison of mix (E) with mix (F) controls for 3 of the 4 chemical additives.

38. **(A) (B)** The average output of plants given one mix is always expected to differ from the average output of plants given a different mix.

39. **(A) (B)** At least one pair of comparisons allows you to assess the effect of phosphate when all other chemicals are controlled.

40-44 (9 pts). Which of the following studies describe(s) experiments, regardless of whether the experiment was designed well or poorly and regardless of ethics. The goal is given (but not underlined). The question is whether the option describes an experiment for obtaining data with respect to the goal. **(A) = is an experiment** **(B) is not**

40. **(A)(B)** You normally study for tests only the night before an exam, but this time, you study 2 days in advance to see if your score improves.

41 **(A)(B)** You analyze surveys from 1,000 college students (the surveys have data on lifestyles and health issues) to identify possible correlates of health issues.

42. **(A)(B)** You wish to see if your students respond more favorably to horoscopes that are individualized to them personally than to horoscopes that are generic and not personalized. You give half your students the identically same horoscope, the other half get individualized descriptions, but you tell everyone that they were getting individualized descriptions. You collect their responses to the horoscopes to see if there is a difference between the two groups.

43. **(A)(B)** During sequestration from covid-19, you have run out of your favorite brand of spaghetti sauce. To avoid starving, you try a new brand because you have no choice.

44. **(A)(B)** A researcher wants to know the effect of maternal diet on birth defects. For mothers who have recently given birth and for whom the status of their baby is known (with or without defects), she has the mothers fill out surveys to describe their diet in the last year (without knowing the purpose). From these surveys, she analyzes the data by sorting the mothers into age groups. She finds that mothers in their 20s who took thrice the recommended daily dose of vitamin A have a 1 in 570 chance of a child with birth defects, but mothers in their 30s have a 1 in 400 chance.

45-52. (**10 pts**) **Prisoners of Silence video (FC = facilitated communication).** *The video showed tests of FC suggesting that the facilitator, not the child with autism, was the author of the typed responses.* The following questions require you to address and interpret the features of this experiment. *Tthis one is straight out of a lecture)*

A = TRUE **B = false**

45. **(A)(B)** Models. An experiment was done to test the *null* model that the child knew the answers *(the question is whether this experiment used the null model approach)*,

46. **(A)(B)** Manipulation. The *manipulation* (what makes the test an experiment) consisted of showing the facilitator something different than what was shown to the child and not letting the facilitator know what the child saw.

47.(A)(B) Variables. The 'third variable' in FC that the experiment needed to eliminate was child knowledge of the answer.

48. **(A)(B)** Controls. The controls in this experiment consisted of showing the facilitator and child the same picture/object.

49. **(A)(B)** Ideal data. An explicit protocol for the experiment was evident from the way the study was conducted.

50. **(A)(B)** The study is considered an experiment specifically <u>because</u> it included controls, replication and randomization.

51. **(A)(B)** Type of experiment. This experiment was the type in which the relevant 3rd variable was known in advance

52.(A)(B) Blind. The experiment included blind in some respects but not in others. For example, the facilitators knew they were being tested and so were not blinded in that respect.

53-58. (**8 pts**) **Nova video – *Secrets of the psychics*.** *Class was shown 2 psychic experiments in the video (bogus horoscope and palm reading) and, as an online exercise, was subjected to a mock personality experiment whose design was similar to the horoscope experiment in the video. (Neglect the Moscow psychics part of the video.)* The following questions require you to address and interpret the features of these three tests/experiments.

A = TRUE **B = false**

53. **(A)(B)** Models. In at least two of these tests, the models being compared were:

'responses reflect the personalized accuracy of a reading/description'

versus

'responses reflect the expectation of personalized accuracy.'

54. **(A)(B)** All three tests required that the subjects were blinded in some fashion.

55. **(A)(B)** Null model. The type of analysis used in the horoscope experiment was a null model approach.

56. **(A)(B)** Controlled variables. The horoscope/personality designs maintained an expectation of personalized accuracy for all subjects.

57. **(A)(B)** Controls. A control group was evident in the palm reading 'study'

58-62 (8 pts) 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 versus 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.

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

59 (A)(B) The extent to which the random choice controls 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.

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

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

62 (A)(B) Randomly choosing subsets of smokers and of non-smokers combined with an analysis of differences in recorded characteristics would guarantee that you can discover any possible causal third variable that might between smokers and non-smokers.

63. (2 pts) (A) **Key code, name, and ID number**. Provide answer (**A**) on question 63 to indicate your key for this version of the exam. Be sure your name and last 4 digits of your Vandal ID are included at the front of your string of answers you upload for this exam.