Psychology 205, Fall 2016
Research Methods in Psychology
Midterm #1

Name: _____

Short Answer (12 questions, 45 points):

1. *What two appropriate logical deductions are needed to verify the conditional statement: if P, then Q? Give an example of a rule and how you would check it. (3 pts)

   Affirm the antecedent, deny the consequent. If P, then Q. Not Q, then not P.
   Good example of this, such as beer -> 21. Check each person who is drinking beer is 21, and if a person is not 21 is not drinking beer.

2. List and define three measures of central tendency. What are some advantages and disadvantages of these measures relative to each other (give at least two). Why do we refer to some statistics as 'robust'? (4 pts)

   Mean sensitive to outliers. Median doesn't represent all the data. Mode only tells the frequent number, not good measure of central tendency. Robust statistics are statistics with good performance for data drawn from a wide range of probability distributions, especially for distributions that are not normal.

3. Regarding the relationship between the data we collect and the data's uncertainty:

   a. (2 pts) Data = _____ + _____

   b. (2 pts) The research process revolves around the concept of variability. What variance is the good kind and what kind of variance is bad?

Data = Model + Error/Residual
Systematic variability related in a predictable fashion to the variable the researcher is interested in is good. Error variance that is unaccounted for is bad.

4. In general, if we draw a very large number of samples from a population, how will the mean, variance, and shape of the resulting distribution compare to the mean and variance of the parent population? (3 pts)

4 points: By the central limit theorem, the (i) sampling distribution mean = mean of parent distribution and (ii) variance of sampling distribution is less than variance of parent distribution. (Note: "sigma-squared/N" was also acceptable for sampling distribution variance). (iii) sampling distribution is normal but parent population does not have to be.

How does the shape of the resulting distribution vary with sample size n? (1 pt)

As n gets large, the shape tends towards normal and the curve gets skinnier (standard dev. of the sampling distribution gets smaller).

5. How do psychologists define 'Type I' and 'Type II' errors? What relevance do these have for the analysis of experimental data? How are they related to each other? (4 pts)

   (i) Type 1: rejecting the null hypothesis when it is true; Type 2: failing to reject the null hypothesis when it is false. (ii) Various correct statements about the utility of alpha and beta settings, used in setting significance, tolerance for false positives (alpha level), power (enlarge alpha), etc. were accepted. (iii) Relationship: something along the lines of "inversely related" or "as you decrease the probability of making a Type 1 error, you increase the probability of making a Type 2 error and vice versa"

6. Give an example of a within-subjects design. What are the potential advantages and disadvantages of this type of design? (Describe 2 of each). (6 pts)

   2 advantages (various)
   Adv – controls subject variability, less subjects needed (they serve as their own control group or they go through all necessary conditions)

   2 disadvantages (various)
   Disadvantage – changes over time (fatigue, carryover, practice effects), order effects

7. Define reliability and compare it to validity. Give an example of when a measure could be valid but not reliable. (4 pts)

   Reliability is the consistency or dependability of a measurement technique.
   Validity is the extent to which a measurement procedure actually measures what it is intended to measure.
   Example: a weight scale that gives a different result every time the same person stands on it repeatedly. Another example: a scale that actually measures hunger but has poor test-retest reliability. [Other examples were accepted.]

8. Describe what an interaction effect is. Draw an example of an interaction, using a line graph (and make up the variables). Make sure to label your graph. (4 pts)

   An interaction between two or more IVs is present if the effect of one IV is different under one level of another IV than it is under another level of that IV. [Many students got this wrong because they said something to the effect of "when two IVs have a relationship with each other." This answer is incorrect because it is not specific enough. Two IVs can have a relationship (e.g., a positive correlation)

9. In class, we discussed Madsen and McGaugh's (1961) experiment on rats exploring the effects of electroconvulsive shock on memory.

   a) State the IV and DV for this experiment. (2 pt)

   IV – ECS (5 secs after stepping off plate) or no ECS (control). DV – step down avoidance

   b) What specific hypotheses were being tested? (2 pt).

   ECS induces fear or amnesia. Fear leads to not stepping off plate. Amnesia inhibits learning and leads to stepping off plate.

   c) Explain what elements of this experiment are congruent with what Platt refers to as "touchstone(s) of strong inference"? (2 pts)

   Contains a test that pit two alternative hypotheses against each other. The confirmation of one hypothesis disproves the alternate.

10. What are operational definitions and how do they compare to conceptual definitions? (2 pts)

    Operational definition defines concept by how it is measured or manipulated in a particular study; a conceptual definition is a definition like one finds in the dictionary—it is usually too broad to be used for research, whereas an operational definition is precisely how a construct in measured in research and may change depending on study.

11. What is the relationship between an independent variable and a variable that is confounding? How can we reduce the effect that a confounding variable has? (2 pts)

    It is correlated to the IV and affects the DV. Randomization.

12. Name two advantages that block randomization has over complete randomization? (2 pts)

    Block randomization leads to equal cell sizes and reduces chance of end effects and effects of ordering.

Long Answer and Thought Problems (7 questions, 55 points):

13. An experimenter is interested in whether memory can be improved if people use visual imagery. Participants (all female) are placed in one of two groups—some are trained in imagery techniques, others are trained to use rote repetition. The imagery group is given a list of concrete nouns (easier to form images than abstract nouns) to study and the other group is given 20 abstract words (ones that are especially easy to pronounce so repetition will be easy), matched with the concrete words for frequency of general use. To match the method of presentation with the method of study, subjects in the imagery group are shown the words visually (on a computer screen). To control for any "techno-phobia," rote subjects also sit at the computer terminal, but for them the computer is programmed to read the lists to them. After hearing their respective word lists, subjects have 60 seconds to recall as many words as they can, in any order that occurs to them.

    (4 pts) What are the constructs of interest in this study? How were they operationalized?

    (i) Visual imagery and memory; (ii) training in imagery vs. rote techniques, number of words recalled in 60 second recall task (respectively)

    (4 pts) Is anything wrong with this study? Elaborate on any problems you see.

    Confounds are (i) noun type and (ii) presentation method. All correct elaborations accepted.

14. A clinical psychologist interested in treating flying phobia administered a pretest of anxiety to 50 phobic participants. The pretest involved a simulated flight aboard a functional aircraft. To increase the effect of the manipulation, the researcher chose not to inform participants that the flight was simulated. While 'flying,' participants filled out an anxiety questionnaire, but five participants became too anxious to complete the measure. Afterward, individuals with completed measures were matched on anxiety and randomly assigned to either an "Implosion Therapy" group or a "Systematic Desensitization" group. Four of the participants who panicked during the simulation were offered spots in the implosion group; the fifth was offered a spot in the desensitization group. All five declined and withdrew from the study.

    Treatment groups met regularly for four weeks. Three participants dropped out of the implosion group before treatment was complete. Following treatment, remaining participants took part in a flight simulation identical to the first and completed the anxiety measure again. Anxiety levels for both groups had declined, but declines were greatest for the implosion group. The psychologist concluded that implosion therapy is better for treating individuals with flying phobia.

    The pretest and posttest means for the two groups are shown in the table below:

Group	Pretest	Posttest	Average
Desensitization (n=24)	43.70	33.54	38.63
Implosion (n=18)	34.17	18.44	26.31
Average	39.62	27.07	

    (4 pts) What can you conclude about the researcher's findings on the basis of these data? Can you think of at least two alternative explanations for the results obtained?

    Any two good reasons for the above results were accepted, as long as they were stated correctly (e.g., attrition, familiarity/practice, state changes in anxiety, placebo effect, etc.)

    (5 pts) Leaving aside obvious ethical considerations, describe two ways in which you would you make this study better.

    5 points: (i) an untreated control group is needed, (ii) further reasonable discussion of how to solve problems with attrition, (iii) at least one other valid improvement (e.g. how to fix clumsy assignment of the five participants who panicked)
    of control group, (ii) attrition discussion, (iii) at least one other valid improvement

15. A developmental psychologist suspected that parents' divorce has a greater impact on sons' aggressive behavior than on daughters'. To test this, 45 boys (ages 13-15) were recruited from several middle class junior high schools. Participants' divorced parents had been living apart for between one and five years.

    Each boy was seated alone in an experimental room and given four mazes to complete. Experimenters told the boys that the first two mazes (which were moderately difficult) would be for practice; however, the boys would win a small cash prize if their performance on the remaining mazes was above average. Unbeknownst to participants, the second two mazes were unsolvable. While boys worked on each of the mazes, the experimenter rated their behavior for signs of aggression. After working with the mazes (but before debriefing), boys were asked to rate how angry they had felt while attempting to solve each maze.

    Although pre- and posttest self-ratings of anger did not differ significantly, boys' observer-rated behavior was significantly more aggressive in the posttest than in the pretest condition. Based on these findings, the psychologist concluded that anger is a distinguishing characteristic of sons of divorced parents.

    (5 pts) What constructs did this study investigate? How were they operationalized?

    (i) divorce affecting aggression, anger; (ii) observer ratings of aggressive behavior, self-ratings of recalled anger (okay to mention frustration, too.)

    (5 pts) Do you think that, given the hypothesis being tested, the researcher's choice of design was a good one? If not, what would you have done to make it better?

    The study needs three more control groups before the researchers can justify their claims; specifically, groups needed are sons of nondivorced parents and daughters of both divorced and nondivorced parents. (Other valid responses were acceptable if given in addition to these)

16. An investigator hypothesizes that IQ score is related to the number of siblings that a child has. He decides to test his idea and randomly picks 1000 families with young children. He measures the IQ scores of all children in these families and records the highest IQ, which he then correlates with the number of siblings in each family. He finds correlation of .68, so he concludes that the more siblings you have, the smarter you are.

    (4pts) Is this a sound reasoning, and if not, where is the problem?

    Experimenter picks only the highest IQ in a family, but the more siblings you have in a family, more likely it is one of them to be better than average.

    (4 pts) How can this study be improved?

    You can improve the design by comparing the mean IQ in the families rather than picking up the highest one.

17. One measure of a dog's intelligence is whether he knows and responds to his name. (Assume that this is a scientifically valid measure.) You decide to try this out on your dog. You call his name from across the room and he comes running, performing consistently across many trials. You take this to be confirmation that he knows his name.

    (4 Pts) What other test should you run before you make this conclusion? Why?

    Need a test of denying the consequent. Should try calling out words other than the dog's name to make sure he doesn't come for them.

18. A professor wanted to test to see if overall experience with college courses leads to an improved ability to learn new math concepts. The professor noted the number of years each participant had been in college and then taught them about non-Euclidean geometry. After two hours of instruction, the professor measured the participants' comprehension of non-Euclidean geometry. The professor found that there was a strong relationship between number of years in college and level of comprehension. She concluded that exposure to a greater number of college courses led to a greater ability to understand new concepts.

    (3 pts) There are several alternative explanations for the professor's findings. Name at least two.

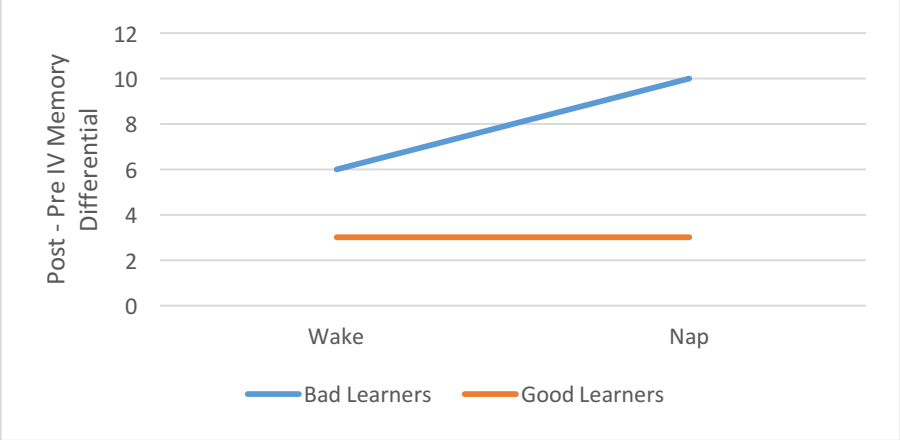
    Exposure to more math classes is the key, not more classes in general. Individuals who have had more classes are older, they clearly have more initiative than individuals who drop out. In other words, older college students are a group stripped (mostly) of the lazy or not-bright. Older students in higher level courses may have run across non-Euclidean geometry before.

    (3 pts) If you were to design a study to investigate this area of interest, what controls would you put in place to eliminate confounds?

    Take note of the number of math courses each P has taken, their exposure to non-Euclidean geometry in the past, the IQs maybe, their GPAs definitely.

19. Researchers wanted to study how spatial memory can be improved by sleeping, with the hypothesis that sleep helps consolidation of memory. A large number of participants were recruited and separated into sleep and wake groups. Participants who indicated that they are habitual nappers on a survey were placed in the sleep group, with others placed in the wake condition. The wake group engaged in a task that occupied their attention and working memory. Participants learn the locations of objects on a grid and take a pre-test where they place the objects as close to the learned location as possible. After a 90 minute nap or wake session, participants take the location test again and a post-nap – pre-nap accuracy differential score is calculated.

    The results indicate that the group that napped had significantly better memory of spatial locations than the wake group and there was a difference between the group that performed well on the task pretest vs ones that did not perform well.



(3 pts) What are the independent and dependent variables explored?

(DV) Spatial memory accuracy differential. (IV) Nap/wake and good/bad learners.

(4 pts) List 2 alternate explanations that could explain the results.

Selection bias for nappers. Regression to the mean for bad learners/ceiling effects for good learners. Also accept differences in the task occupying the wake group as a source of findings.

(3 pts) Is there an interaction? If so, explain the effect.

The worst learners before the nap benefitted most from the nap session. They conclude that good learners do not reap as much of the benefits of enhanced memory consolidation during sleep as bad learners do.