1. What is the most important way in which a research methods course differs from a course in social psychology?
   (a) the social psychology course will emphasize process, while the research methods course will emphasize content
   (b) the methods course will emphasize how research occurs, while the social psychology course will focus on the research outcomes themselves
   (c) the methods course will have a focus on ethics, while the social psychology course will not consider ethics
   (d) the social psychology course will have a greater emphasis on statistics
2. Which of the following is true about belief perseverance?
   (a) it is the basis for Peirce's way of knowing called the a priori method
   (b) it refers to an unwillingness have one's opinions changed, even by solid scientific evidence
   (c) it is a tendency for events to stand out in our minds because we keep seeing them on the news
   (d) it refers to the fact that most of our strong beliefs are formed in childhood, and last throughout adulthood
3. Ed believes he is in telepathic communication with Sally because it seems like every time he thinks of her, she calls him on the phone. He ignores all the times he is thinking of her and she doesn't call. That is, he is being affected by
   (a) a confirmation bias
   (b) belief perseverance
   (c) the unavailability heuristic
   (d) statistical determinism
4. Of the following questions, only one is an empirical question. Which one?
   (a) Will males or females be more likely to give blood?
   (b) How does the mind exert its influence of the physical body?
   (c) Are people basically good, but corrupted by society?
   (d) Can people be truly evil?
5. After research shows that a form of behavior therapy can reduce phobic responses, therapists begin using the technique in their practices. Which of the goals of research in psychology is being reflected here?
   (a) explanation
   (b) application
   (c) description
   (d) prediction
6. Which of the following provide the best reasons why publishing scientific results is important for science?
   (a) publication allows reviewers to know about the results of research before the public; publication allows authors to present their research to their peers.
   (b) publication is a method for creating public knowledge of the results of research; publication allows for science to correct itself

(c) publication allows journals to select the most interesting new research; publication allows granting agencies to evaluate proposals

(d) publication allows researchers to establish their credentials; publication allows universities to gain government grants for research

7. According to Karl Popper, scientific claims must be:

   (a) proven true

   (b) novel

   (c) as simple as possible

   (d) falsifiable

8. Psychology is a science because:

   (a) It uses systematic theoretical constructs to help the public understand ideas about their own psychology

   (b) It uses applications of research to help people

   (c) Many universities classify Psychology under natural and behavioral sciences

   (d) It uses systematic observations to answer empirical questions and distributes the results of research publicly

9. Many people who believe in extrasensory perception (ESP) and other psychic powers claim that such powers can disappear when they are observed too closely. This is an example of pseudo-science because:

   (a) Observations showing no evidence of ESP would not count as evidence against ESP.

   (b) Observations showing no evidence of ESP would count as evidence against ESP.

   (c) ESP believers never discuss the results of their studies publicly.

   (d) ESP believers always discuss the results of their studies publicly.

10. Scientist's are often skeptical about ideas and claims. Being skeptical refers to:

    (a) Thinking that the ideas and claims are probably wrong.

    (b) Believing in ideas and claims when they are presented by other scientists.

    (c) Requiring evidence to back up ideas and claims.

    (d) Being cynical about the motivation behind ideas and claims.

11. Scientist's should reject a claim when:

    (a) The claim flies in the face of facts about life that everyone knows are true.

    (b) There is credible evidence in support of the claim, and there is no credible evidence against the claim.

    (c) There is no credible evidence in support of the claim, and there is a lot of credible evidence not supporting the claim.

    (d) A dominant scientific theory says the claim is not true, so it should be rejected.

Original study effect size versus replication effect size (correlation coefficients). Diagonal line represents replication effect size equal to original effect size. Dotted line represents replication effect size of 0. Points below the dotted line were effects in the opposite direction of the original. Density plots are separated by significant (blue) and nonsignificant (red) effects.

12. The graph shows the results from the open science collaboration replication studies discussed in lecture. If the replication studies produces larger effects than the original studies then the dots would be _____; if the replication studies produced the same effects as the original studies, then the dots would be _____; if the replication studies produced smaller effects as the original studies, then the dots would be _____.

    (a) above the line; on the line; below the line.
    (b) above the line; below the line; on the line.
    (c) on the line; above the line; below the line.
    (d) below the line; on the line; above the line.

13. In the Monte Hall problem there is a prize behind one of the three doors. You pick one of the doors, then the host opens another. You're given the option to switch your choice or keep the initially chosen door; what do you do next?

    (a) Switch your choice because you likely picked the door without the prize to begin with.
    (b) Stay with the initial door, there is a 50% chance you chose the right door.
    (c) Switch the doors, there is no difference which door you choose there is still a 50% chance you'll land the prize.
    (d) Stick with the initial choice because, first choice is usually the right one.

14. Your grandfather lived to be 96 years old even thou he smoked three packs of cigarette per day. You therefore argue that smoking cigarette is not bad for your health. This is an example:

    (a) Operational definition
    (b) Pluralistic ignorance

(c) Confirmation bias

(d) Availability heuristic

15. What is the difference between applied research and basic research?

(a) applied research refers to research that can be used on the basis of how applicable it is to a certain topic; and basic research refers to how basic the research is

(b) applied research refers to research that can be used to apply to other areas of science; and basic research refers to research that can be applied to the science of psychology only

(c) applied research refers to research conducted in an uncontrolled setting outside of a lab; and basic research refers to research conducted in a controlled setting of a lab

(d) applied research refers to research conducted with the intention of trying to solve real world issues; and basic research refers to research conducted with the intention of predicting or describing principles of behavior

16. One important general aim of the scientific method is to:

(a) produce unverifiable, subjective knowledge

(b) produce verifiable, subjective knowledge

(c) produce unverifiable, objective knowledge

(d) produce verifiable, objective knowledge

17. People who believe that the earth is flat and that the sun revolves around the earth provide an extreme example of:

(a) a group of people who do not exist

(b) Endlessly holding on to a belief, in the face of overwhelming evidence

(c) The availability heuristic

(d) the tendency to estimate the frequency of an event based on how easy it is to bring to mind

18. The concept in scientific experiment that is a critically important attribute of a good theory so that it can be disproven, at least in principle is referred to as:

(a) Falsification

(b) Empirical question

(c) Statistical determinism

(d) Data-driven

19. A categorical variable is:

(a) has single a non-binary scalar value

(b) has different qualities

(c) has a continuous quantity

(d) uses an interval scale, where 0 does not mean complete absence

20. A quantitative variable:

(a) has different qualities

(b) has a numeric range

(c) has single a non-binary scalar value

(d) uses an interval scale, where 0 does not mean complete absence

21. The following measured variables: reaction time, nationality, height, and species, are examples of which sequence of variable types.

    (a) quantitative, quantitative, qualitative, qualitative
    (b) qualitative, quantitative, qualitative, quantitative
    (c) quantitative, qualitative, quantitative, qualitative
    (d) qualitative, qualitative, quantitative, quantitative

22. Imagine a population contains four different kinds of people. Four different researchers each take large samples from the population, using different sampling strategies. The answers below show the proportions of each kind of person that each researcher found in their sample. Which researcher's pattern is most consistent with a random sampling strategy:

    (a) .25, .25, .25, .25
    (b) .6, .1, .1, .1
    (c) .85, .05, .05, .05
    (d) .5, .5, .5, .5

23. A researcher at a university asks the students in their class to participate in an experiment. The textbook describes this kind of sample as:

    (a) convenience sample
    (b) random sample
    (c) young adult sample
    (d) stratified sample

24. A statistical relationship occurs between two variables when:

    (a) the pattern in one variable does not predict any of the pattern in the other variable
    (b) the pattern in one variable predicts at least some of the pattern in the other variable
    (c) the two variables are entered into a statistics program and formally related by conducting the analysis
    (d) the pattern in one variable is self-predictive

25. A difference between groups in an experiment could indicate:

    (a) there is an effect of dependent variable
    (b) there is an effect of the independent variable
    (c) there is not an effect of the dependent variable
    (d) there is an effect of the measured variable

26. There is a positive correlation between variable A and B, this means that:

    (a) as variable A gets bigger, so does variable B. And, as variable A gets smaller, so does variable B
    (b) as variable A gets bigger, so does variable B. And, as variable A gets smaller, variable B gets bigger
    (c) as variable A gets bigger, variable B gets smaller. And, as variable A gets smaller, variable B gets larger
    (d) changes in variable A do not predict changes in variable B.

27. There is a negative correlation between variable A and B, this means that:
    (a) as variable A gets bigger, variable B gets smaller. And, as variable A gets smaller, variable B gets larger
    (b) as variable A gets bigger, so does variable B. And, as variable A gets smaller, variable B gets bigger
    (c) as variable A gets bigger, so does variable B. And, as variable A gets smaller, so does variable B
    (d) changes in variable A do not predict changes in variable B.
28. There is no correlation between variable A and B, this means that:
    (a) changes in variable A do not predict changes in variable B.
    (b) as variable A gets bigger, so does variable B. And, as variable A gets smaller, variable B gets bigger
    (c) as variable A gets bigger, variable B gets smaller. And, as variable A gets smaller, variable B gets larger
    (d) as variable A gets bigger, so does variable B. And, as variable A gets smaller, so does variable B
29. Which statistic is a common measure of correlation?
    (a) variance
    (b) cohen's d
    (c) Pearson's r
    (d) standard deviation
30. Imagine there is a U-shaped relationship between variables A and B. This would tend to produce what kind of result as measured by Pearson's r?
    (a) $r > 0$
    (b) $r < .05$
    (c) $r < 0$
    (d) $r = 0$
31. In lecture we discussed an experiment by Carter, Ferguson, and Hassin (2011) who showed that a single exposure to the American flag causes people to shift their political views toward Republicanism. What was their main explanation for this effect?
    (a) The American flag happens to be more associated with the Democratic party, and therefore primes people to endorse Democratic political views.
    (b) The American flag does not have a priming influence and the people who were shown the flag happened to express more Republican views than Democratic views.
    (c) The American flag happens to be more associated with the Republican party, and therefore primes people to endorse Republican political views.
    (d) The American flag does not have a priming influence and the people who were shown the flag happened to express more Democratic than Republican views.
32. In the flag-priming study, the author's mentioned a plausible alternative account of their findings that does not assume the American flag is more associated with the Republican party, which option best represents this plausible alternative?

(a) Flags are a symbol of national unity and may prime people to be more open to endorsing political views from another political party that they would not normally support.

(b) Flags are meaningful symbols, but they are not associated with political views, and the results they found were due to random chance.

(c) Flags are not meaningful symbols and they do not have a priming influence on political decision-making.

(d) Flags are a symbol of national pride and may prime people to more strongly endorse their existing political views.

33. In terms of increasing complexity, which order below of the four measurement scales is correct?

(a) nominal, ordinal, ratio, interval

(b) ordinal, nominal, interval, ratio

(c) ordinal, nominal, ratio, interval

(d) nominal, ordinal, interval, ratio

34. What are Nominal Scales?

(a) Numbers do not represent order, they represent different categories

(b) Numbers represent order (rank) with equal interval, zero mean complete absence

(c) Numbers used to make ranks, or show the order of smallest to largest, most to least, best to worst

(d) Numbers represent order (rank) with equal interval, zero is just another value not the absences of something

35. What is the advantage of using operational definitions?

(a) there are no advantages to using operational definitions; no two persons can ever agree on the best definition

(b) they force researchers in different laboratories to all use the exact same definition

(c) it's easy to agree on a universal definition for a concept like aggression

(d) they allow researchers to understand the details and logic of an experiment, and facilitate replication and communication of research

36. Operational definitions

(a) are used by researchers to be clear about the terms of their studies

(b) are needed to force researchers in different laboratories to all use the exact same definition of a construct

(c) differ from one study to another, which means that using operational definitions hinders the replication process

(d) are seldom needed because of modern advances in behavioral technology

37. The results of an inkblot test might be quite different when given to the same person on two different occasions. If this is the case, then based on this fact alone, the inkblot test is

(a) not reliable

(b) not valid

(c) not reliable but probably valid

(d) neither reliable nor valid

38. On a reaction time test, which of the following factors could contribute to measurement error?

    (a) all of the options
    (b) subject attentiveness
    (c) increased boredom if the task lasts too long
    (d) equipment irregularities

39. Which of the following is true about measures of behavior?

    (a) if a measure has content validity, it is almost certain to be reliable
    (b) measurement error can be eliminated completely by careful researchers
    (c) they are more likely to be valid than reliable
    (d) they all include some degree of measurement error

40. A test is said to be reliable if _____, and valid if it _____.

    (a) its results are repeatable; measures what it is supposed to measure
    (b) its results are repeatable; is low in measurement error
    (c) measures what it is supposed to measure; is low in measurement error
    (d) has a sufficiently high amount of measurement error; measures what it is supposed to measure

41. Classification is the major purpose of a(n) _____ scale of measurement.

    (a) ratio
    (b) nominal
    (c) interval
    (d) ordinal

42. When considering a student's overall standing in a class (first, second, third, etc.), which measurement scale is being used.

    (a) ordinal
    (b) interval
    (c) nominal
    (d) ratio

43. When using a(n) _____ measurement scale, the most that can be said is that one score is greater than another.

    (a) interval
    (b) ordinal
    (c) ratio
    (d) nominal

44. The main difference between an interval and a ratio scale is that an interval scale

    (a) is used only for placing participants into categories
    (b) has equal intervals between numbers
    (c) does not preserve a rank order in the assignment of numbers
    (d) does not have a true zero point

45. Which of the following statements are both true?

    (a) in a ratio scale, a score of zero means the absence of the phenomenon being measured, whereas zero is just another value for interval scale

    (b) equal intervals exist in ratio scales, but such is not the case in interval scales

    (c) in an interval scale, a score of zero means complete absence, whereas zero is just another value for the ratio scale

    (d) equal intervals exist in interval scales, but such is not the case in ratio scales

46. Remember the experiment where people were asked to judge the quality of different stockings. Based on this experiment what did Nisbett and Ross conclude about the use of verbal reports?

    (a) Introspection has its drawbacks but can provide potentially interesting insight into the inner mental world

    (b) People have the ability to know and verbally report facts about their own behavior

    (c) Introspection is a classic psychological technique that is well known to provide valid and reliable results

    (d) People often make up stories that do not relate to their actual behavior

47. The internal validity of a study is high when

    (a) potential confounds are properly controlled

    (b) the results generalize to other situations

    (c) external validity is also high (they go together)

    (d) the results apply to other groups of people

48. If you develop a test to measure performance on some task, but everyone who takes the test scores between 95 and 100%, what has happened?

    (a) The test results were not very similar indicating different levels of performance

    (b) The test results were very similar and participants were on the same floor

    (c) The test was too easy creating a ceiling effect

    (d) The test was too easy creating a floor effect

49. Internal validity is determined by:

    (a) The longevity of results.

    (b) The degree to which research findings generalize beyond the specific context of the experiment being conducted.

    (c) The subject pool or the participant pool.

    (d) The degree to which an experiment is methodologically sound and confound free.

50. In lecture we discussed a study by Bargh, Chen, and Burrows (1996) who showed that priming participants with the concept of old age caused them to walk out of the experiment room more slowly than a group of participants who were not primed. What was one of the most important major flaws in the study?

    (a) A confederate in the study used a hand-held stopwatch to measure the walking times of participants, and potentially introduced a bias into the measurement.

    (b) The were not aware of the old-age primes and these words could not have caused them to behave as if they were elderly.

(c) The words used in the sentence scramble task were not strongly associated with the concepts of old age, and did not prime this concept in the participants

(d) The participants were all undergraduates in a psychology program and were too young to be primed with the concepts of old age.

51. Researcher A measures 10 scores from a distribution with mean = 100, and standard deviation = 25, and Researcher B measures 10 scores from a distribution with mean = 200 and standard deviation = 10. Which researcher's measure is more reliable?

(a) There is no way to answer this question with the information provided

(b) The measures have equal reliability.

(c) Researcher B's measure

(d) Researcher A's measure

52. Researcher A measures 10 scores from a distribution with mean = 100, and standard deviation = 10, and Researcher B measures 10 scores from a distribution with mean = 200 and standard deviation = 50. Which researcher's measure is more reliable?

(a) The measures have equal reliability.

(b) There is no way to answer this question with the information provided

(c) Researcher B's measure

(d) Researcher A's measure

53. How is the process of sampling from a distribution related to the concept of reliability?

(a) The consistency of a sample is directly related to the variance of distribution. Distributions with higher variances lead to more consistent and more reliable samples, compared to distributions with lower variances which lead to more inconsistent less reliable samples.

(b) The consistence of a sample is directly related to the mean of a distribution, and we expect that the sample mean will generally be close to the population mean, thereby increasing reliability.

(c) The consistency of a sample is directly related to the variance of a distribution. Distributions with higher variances lead to more inconsistent and less reliable samples, compared to distributions with lower variances which lead to more consistent more reliable samples.

(d) These two concepts are not related.

54. If we computed the sampling distribution of the mean for a set of samples taken from a population, and then divided those scores by the standard deviation of the sampling distribution of the mean, this would be equivalent to finding:

(a) the distribution of population means

(b) the distribution of p-values

(c) the distribution of t-scores

(d) the distribution of degrees of freedom values

55. The sample mean is the mean of a set of sample scores. The sampling distribution of the mean is the distribution of all means taken from all possible samples of a particular size (n=sample size). What happens to the variance of the sampling distribution of the mean as we increase sample-size?

(a) The variance increases

(b) The standard deviation increases but the variance decreases

(c) The variance decreases

(d) Sample-size has no effect on the variance of the sampling distribution of the mean

56. All other things being equal, should you be more confident that a sample mean estimates the population mean when it is taken from a large sample or small sample?

(a) A small sample, because the sample mean approaches the population mean as n increases.

(b) A small sample, because the sample mean approaches the population mean as n decreases.

(c) A large sample, because the sample mean approaches the population mean as n decreases.

(d) A large sample, because the sample mean approaches the population mean as n increases.

57. Imagine that two researchers conducted different studies with similar sample-sizes. Researcher A found a t-value of 2, and Researcher B found a t-value of 4. Based on these numbers alone, which result was less likely to have been produced by chance alone.

(a) t-values give you no information about the likelihood that a result was due to chance.

(b) Researcher A's result.

(c) Researcher B's result.

(d) only the degrees of freedom can tell you if a result is due to chance.

58. The above graph shows a sampling distribution of the mean for a particular population. 2.5% of the sample means fall to the left of the first bar, and 2.5% of the sample means fall to the right of the second bar. Imagine you measured a sample, and found a mean of 97. What should you conclude?

(a) The sample with mean 97 could easily have come from this distribution

(b) The sample with mean 97 only occurs with a specific probability, so is unlikely to have come from this distribution

(c) The sample with mean 97 was produced by random chance, so the likelihood it came from this distribution can not be determined.

(d) The sample with mean 97 was highly unlikely to come from this distribution

59. You've run an experiment where one group of professional swimmers' individual lap-speeds were clocked after each swimmer took a placebo to enhance performance, and another group was clocked after they each took a performance enhancing supplement that you are researching in order to possibly enter the market soon. Which statistical test should you run to help determine whether any changes in swimmers' lap speeds was due to confounds (e.g. chance) alone?

(a) Single sample t-test

(b) Independent sample t-test

(c) Paired sample t-test

(d) Repeated measures ANOVA

60. Which keywords (definition) corresponds best to Independent Sample, Paired Sample, and One sample test, respectively?

(a) Within subject design, to see difference between sample mean from population mean, and between subject design

(b) Between subject design, to see difference between sample mean from a population mean, and within subject design

(c) Within subject design, Between subject design, and to see difference between sample mean from a population mean

(d) Between subject design, Within subject design, and to see difference between sample mean from a population mean

61. The independent variable refers to

(a) the variable which is only used in the control condition.

(b) the variable that is being measured in an experiment

(c) the variable which is being viewed in its natural state.

(d) the variable being manipulated or varied in some way by the researcher.

62. The difference between a within-subjects design and a between-subject design is that:

(a) within-subject designs have the same subjects participate in both experimental groups

(b) within-subject design is the same as a between-subject design

(c) between subjects designs have the same subjects participate in both experimental groups

(d) within-subject designs have different subjects in each experimental group

63. Which one of the following t-tests compares the same group of participants on two different conditions (e.g., performance before training vs. after training) in order to determine whether a significant difference exists?

(a) One-sample t-test

(b) Paired-samples t-test

(c) Between-subjects t-test

(d) Independent-samples t-test

64. Which T-test is used to test whether a sample mean is different from a population mean?

(a) Independent-samples T-test

(b) All of the above

(c) Paired- samples T-test

(d) One- sample T-test

65. What's the difference between a Between-Subjects Design and a Within-Subjects Design?

(a) There is no difference, they are just alternative terms for the same thing.

(b) Between-subject design is where there are the same subjects in each experimental group, whereas a Within-subject design is where there are different subjects in each experimental group.

(c) Between-subject designs are only used in paired-sample t-tests, whereas a Within-subject designs are only used in independent-samples t-test.

(d) Between-subject design is where there are different subjects in each experimental group, whereas a Within-subject design is where there are the same subjects in each experimental group.

66. When a confound exists

    (a) it is the confounding factor not the independent variable which causes the behavior to occur

    (b) there will be at least 2 different ways of interpreting the results

    (c) the researcher failed to include a controlled group in the study

    (d) it simply means that the dependent variable hasn't been define precisely enough

67. If you were to conduct an experiment on whether people write more clearly with their dominant hand compared to using their feet, the difference in legibility between conditions would be:

    (a) impossible to interpret without a statistical test to establish significance

    (b) completely random and impossible to interpret

    (c) obvious to anyone looking at the writing samples

    (d) will need stem and leaf plots to measure variance in the spread of the biconditional modal distribution

68. A researcher shows tragic sad films to one group of subjects and violent films to another group of subjects. The researcher then assesses the participants actions after watching each film. The independent variable in this study is

    (a) the level of aggressiveness in subjects.

    (b) the type of film seen.

    (c) the level of depression in subjects.

    (d) the level of cooperativeness to watching the film.

69. For any experiment investigating the effect of X on Y,

    (a) X is the independent variable, Y is the dependent variable

    (b) Y is the independent variable

    (c) X is the dependent variable, Y is the independent variable

    (d) X is the dependent variable

70. A researcher tests four different groups of participants. Each group is given a different dosage of caffeine, and reaction time is measured for each subject. Which of the following is true?

    (a) the independent variable has four levels

    (b) reaction time is an independent variable

    (c) dosage level is the dependent variable

    (d) the independent variable is a subject variable

71. To see if the time of day has an effect on helping behavior, experimenters ask passersby for directions in New York's Central Park, either at 8:00 am or 8:00 pm. Which of the following is true?

    (a) the independent variable is an instructional variable

    (b) this is an example of field research that lacks an independent variable

    (c) this is an example of a field experiment

    (d) this is a study with two independent variables (the times)

72. What do all experiments have in common?
    (a) there must be at least two dependent measures
    (b) at least one independent variable will be an instructional variable
    (c) there must be a control group
    (d) there are at least two different ways in which participants are treated
73. Experimental group is to control group as _____ is to _____.
    (a) untreated; treated
    (b) dependent variable; independent variable
    (c) treated; untreated
    (d) independent variable; dependent variable
74. In a study finding effects of X on Y, variable Z is confounded with X. Which of the following is true?
    (a) Neither X or Z could be causing Y because of the confound
    (b) X could be causing Y
    (c) Z could be causing Y
    (d) X and Z could be causing Y
75. When a confound exists,
    (a) it cannot be determined whether the confounding variable or the independent variable is causing the results to occur
    (b) there will be at least two different ways of interpreting the results
    (c) some uncontrolled factor covaries with the independent variable
    (d) all of the answers
76. A researcher wants to know if children prefer cereal that has colors in it. Preschoolers compare plain-colored Wheatios with multicolored Rice Chrunchios. The children preferred the latter. What can be concluded?
    (a) children prefer colored cereal
    (b) whether the cereal is composed of wheat or rice is the independent variable
    (c) cereal color is the dependent variable
    (d) cereal type (wheat vs. rice) is confounded with cereal color
77. In any experiment, the dependent variable is
    (a) the factor being manipulated by the experimenter
    (b) some behavior being measured
    (c) the factor that is being controlled
    (d) usually selected randomly
78. When a subject variable is used as an independent variable and differences occur between groups on the dependent measure, what can be concluded?
    (a) in this case, the subject variable must be a dependent variable, not an independent variable
    (b) the two groups performed differently

(c) the independent variable caused the differences to occur
(d) because a subject variable is being used, nothing at all can be concluded

79. In a study about self-esteem, self-esteem is

(a) could be any of the options
(b) the dependent variable
(c) a subject variable
(d) a manipulated variable

80. At the very least, the independent variable should have:

(a) at least 2 IVs
(b) at least 2 levels
(c) at least more than 2 IVs
(d) more than 2 levels

81. A between-subjects design

(a) must deal with the problem of sequence effects
(b) tests the same group of participants at each level of the independent variable
(c) requires fewer subjects than a comparable within-subjects design
(d) includes at least two different groups of participants

82. A within-subjects design

(a) requires more subjects than a comparable between-subjects design
(b) usually must deal with the problem of order effects
(c) includes at least three different groups of subjects
(d) tests different groups of participants at each level of the independent variable

83. All of the following generally characterize between-subjects designs except

(a) concern over sequence effects
(b) concern over how to create equivalent groups
(c) random assignment frequently used
(d) requires larger N than comparable within-subjects designs

84. All of the following generally characterize within-subjects designs except

(a) requires smaller N than comparable between-subjects designs
(b) concern over sequence effects
(c) researcher will use matching to control for the subject variable
(d) some form of counterbalancing will be used

85. In _____, each subject volunteering for the study has an equal chance of being placed into group A or group B.

(a) counterbalancing
(b) random assignment
(c) using a Latin square
(d) matching

86. If a between-subjects design uses random assignment, the design will be called a(n)
    (a) independent groups design
    (b) matched groups design
    (c) repeated-measures design
    (d) nonequivalent groups design
87. If a between-subjects design uses the subject variable of sex and has just one independent variable, which of the following is true?
    (a) the design is a multilevel design
    (b) the design will be analyzed with a t test for related samples
    (c) the design is a repeated-measures design
    (d) the design is a nonequivalent groups design
88. What does every single-factor, two level design have in common with single-factor, multilevel designs?
    (a) continuous dependent variable
    (b) t test for all analyses
    (c) random assignment
    (d) one independent variable
89. If you ran a single factor independent groups design with 3 levels, which T-test would you use, and how many separate comparisons would you need to make to test all of the possible differences?
    (a) one sample T-test, 6 comparisons
    (b) independent samples T-test, 3 comparisons
    (c) independent samples T-test, 6 comparisons
    (d) one sample T-test, 3 comparisons
90. Teachers had a theory that starting school at a later time would produce better performance from each student. The teachers tested the performance levels for students who began their school day at 12 pm and those who began at 7:45 am. Which test should be performed given the structure of the experiment?
    (a) Paired sample t-test
    (b) One-sample t-test
    (c) Trick question, you could use all the above,
    (d) Independent t-test
91. If there is a difference in the experimental data, what can it be caused by
    (a) confounds
    (b) manipulation
    (c) chance
    (d) all of the options
92. In within-subjects designs
    (a) There are two independent variables.

(b) There are two separate groups of subjects.

(c) The same subjects take part in both experimental groups.

(d) There are different subjects in each experimental group.

93. What is the difference between a paired- samples test and a independent- samples test?

(a) A paired-samples test results in a P-value and an independent-samples test results in a T-value.

(b) A paired-samples test is a within subject design and an independent-samples test is a between subjects design.

(c) A paired-samples test is with 2 people and an independent-samples test is for a single person.

(d) A paired-samples test is a between subject design and an independent-samples test is a within subject design.

94. Which of the following statements about a Latin square is true?

(a) All of the options

(b) It ensures that each condition precedes and follows each other condition exactly one time.

(c) It is a form of partial counterbalancing.

(d) It ensures that each condition of the study occurs equally often in each sequential position.

95. Out of a set of four numbers I tell you that you can rearrange and change all the numbers except one in order for the mean to stay the same. The amount of numbers you are not allowed to change is your:

(a) Alpha level

(b) t-value

(c) Degrees of freedom

(d) Standard Deviation

96. What is not an advantage to within-subject designs?

(a) determine changes within individuals and not between groups

(b) Less subjects are needed because everyone participates in each level of the design

(c) Sequence or order effects

(d) rules out the confound of experimental groups

97. What is the correct way to properly statistically report a t-test?

(a) p < or > .05(df)= t-value

(b) t(t-value)= df, p < or > .05

(c) t-value(df)=t, p < or > .05

(d) t(df)= t-value, p < or > .05

98. A Double-Blind study is when:

(a) The participants and experimenters are aware of what the purpose of the study is but both do not know which conditions the participants are in

(b) The experimenter knows the purpose of the study and knows which condition the participants are assigned to

(c) Both the participant and the experimenter do not know which condition they are in

(d) The participants do not know which condition they are in but the experimenter does know which condition they are in

99. In lecture we discussed a study by Adam and Galinksy (2012) who showed that wearing a white lab coat can prime participants to adopt a scientific focus and behave more carefully and attentively. One advantage to their findings was that:

(a) Although they showed that wearing a lab coat did influence measures of attention, they did so in two different tasks and this limited their ability to show consistent effects across tasks.

(b) They showed that wearing a lab coat influenced measures of attention in two different tasks, thereby replicating their own effect and showing that it generalizes across tasks.

(c) They had relatively large sample sizes but generally had more females than males in their study, thus limiting the generalization of their conclusions.

(d) They only used a lab coat as a clothing item and did not test how other items of clothing might influence behavior.

Sub # A	Level 1	Level 2	Diff. Score (L2-L1)
1	8	14	6
2	8	14	6
3	6	11	5
4	7	12	5
5	7	7	0
6	6	7	1
7	7	7	0
8	5	12	7
9	10	5	-5
10	8	9	1

Sub # B	Level 1	Level 2	Diff. Score (L2-L1)
1	9	13	4
2	7	9	2
3	8	14	6
4	9	12	3
5	10	9	-1
6	9	15	6
7	10	9	-1
8	7	5	-2
9	9	8	-1
10	10	8	-2

100. The above tables A & B show sample data for a paired samples t-test between 2 levels. The scores for each subject for each level are presented, and the difference scores (level 2 - level 1) are also presented. Based on the numbers in these tables alone, which data set would be more likely to produce a significant p-value?

(a) A

(b) Neither A or B

(c) A and B

(d) B

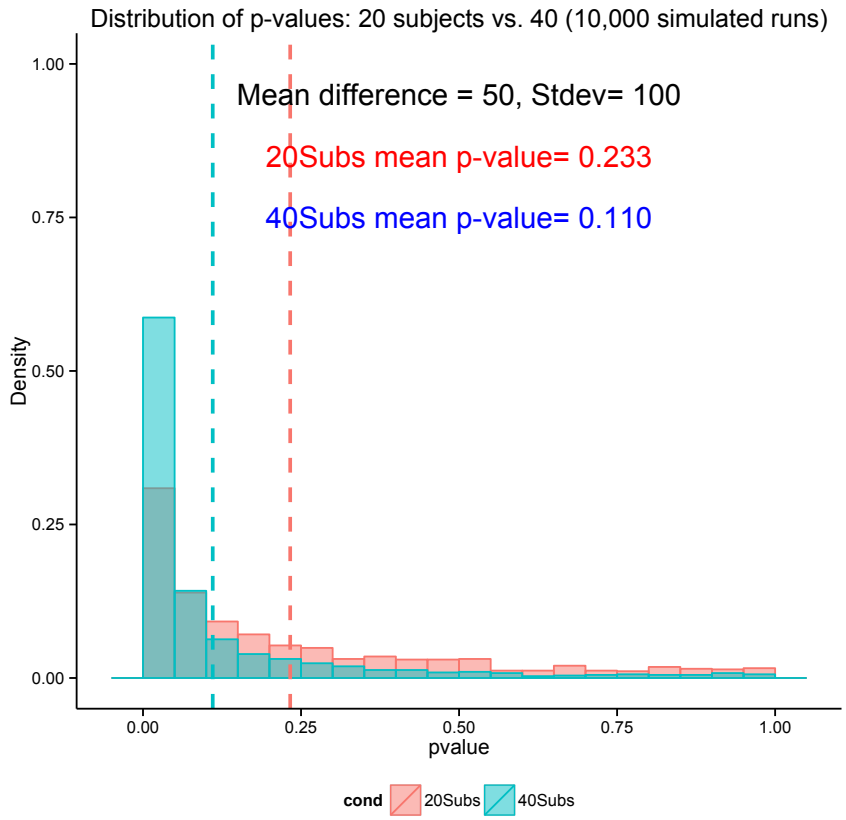
101. If the difference scores are thought of as data from a coin-toss to determine whether a coin is fair, which coin A or B appears to be fair?

(a) A

(b) A and B

(c) B

(d) Neither A or B



102. The above graph shows a simulation for an experiment (1 IV, 2 levels) that was replicated 10,000 times assuming 20 subjects and 10,000 times assuming 40 subjects. Each time a significance test was conducted. The x-axis shows a p-values from 0 to 1, and the y-axis shows the density or proportion of the total number of experiments. Each bar shows the proportion of times that the experiment resulted in a particular p-value. Based on this graph, answer the following questions. The mean p-values across all 10,000 replications are .233 and .110, what conclusion is not true?

(a) Some proportion of the experiments will not produce significant p-values

(b) Running this kind of experiment will never produce a significant p-value.

(c) Some proportion of the experiments will produce significant p-values

(d) Increasing the number of subjects will probably lower the p-value

A	Test Land	Test Water
Training Land	10	10
Training Water	20	20

B	Test Land	Test Water
Training Land	20	10
Training Water	20	10

C	Test Land	Test Water
Training Land	20	10
Training Water	10	20

D	Test Land	Test Water
Training Land	10	20
Training Water	20	10

103. Godden and Baddeley (1975) investigated whether memory performance depends on the match between training and testing environments. They had groups of divers read a list of words either on land or under water. The memory test involved recalling as many words as possible. Some subjects recalled the words in the same environment, and some in a different environment. They hypothesized that recall performance would be better when the test environment matched than mismatched the training environment. Which table best reflects these predictions?

(a) D

(b) C

(c) B

(d) A

104. Using the four tables from above, which of the following statements accurately summarizes the patterns predicted by each of the tables. Main effect of Training means that the table shows such an effect, Main of Test means that the table shows such an effect, Interaction means that the table shows and interaction.

(a) A=Main effect of Test, B=Interaction, C=Interaction, D=Main effect of Training

(b) A=Interaction, B=Main effect of Test, C=Main effect of Training, D=Interaction

(c) A=Main effect of Training, B=Main effect of Test, C=Interaction, D=Interaction

(d) A=Interaction, B=Interaction, C=Main effect of Training, D=Main effect of Test

Column	A	B	C	D
	Caffeine		No Caffeine	
Subjects	Morning	Afternoon	Morning	Afternoon
1	12	8	7	2
2	13	6	9	1
3	13	6	10	3
4	14	8	8	1
5	14	9	7	5
6	13	6	8	3
7	13	7	9	1
8	15	10	10	3
9	13	9	10	2
10	12	10	7	4
Means	13.2	7.9	8.5	2.5

105. Using the table above answer the following questions. The letters ABCD refer to each of the columns in the table. Which of the following would be appropriate for computing the main effect of Caffeine?

(a) Average of BC vs Average of AD
(b) Average of AC vs Average of BD
(c) Average of AB vs Average of CD
(d) Average of A vs. Average of BCD

106. Which of the following would be appropriate for computing the main effect of Time of Day

(a) Average of AC vs Average of BD
(b) Average of AB vs Average of CD
(c) Average of BC vs Average of AD
(d) Average of A vs. Average of BCD

107. Which of the following would be appropriate for computing the interaction between Caffeine and Time of Day

(a) a

(b) (Average of AB - Average of CD) compared to (Average of BA - Average of DC) c. (Average of A - Average of D) compared to (Average of B - Average of C)

(c) (Average of A - Average of B) compared to (Average of C - Average of D)

(d) (Average of AD - Average of BC) compared to (Average of BC - Average of DA)

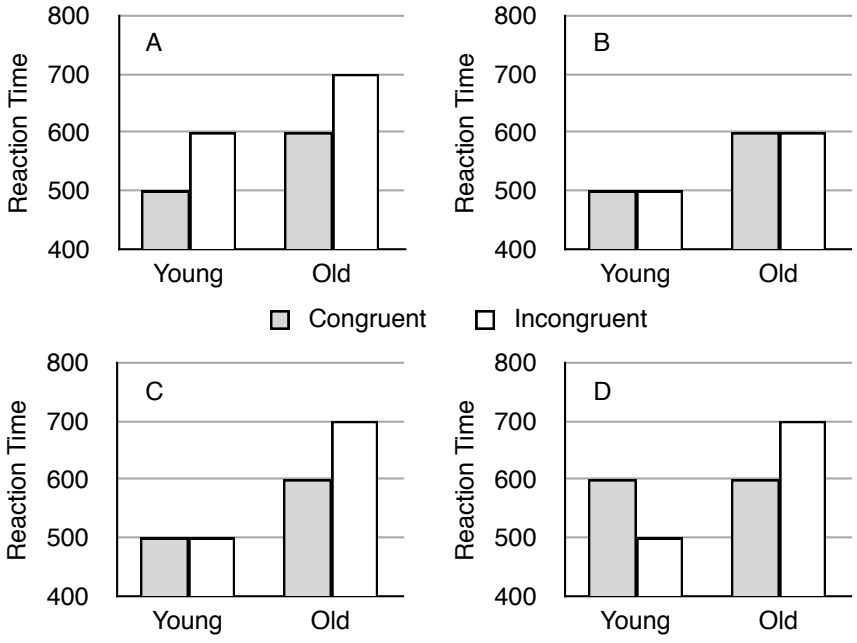
108. Assume that a main effect or interaction from the table needs to have a difference of at least 2 points to be significant. Which of the following accurately describes all of the significant effects?

(a) Main effect of time of day is significant, Interaction is significant

(b) Main effect of caffeine is significant, Main effect of time of day is significant

(c) Main effect of caffeine is significant, Interaction is significant

(d) Main effect of caffeine is significant, no other significant effects



109. Use the above graphs to answer the following questions. All graphs show hypothetical data from a Stroop experiment that was conducted on a population of young and old participants. Which graph(s) show a main effect of Age?

(a) ABC, but not D

(b) AB, but not CD

(c) A, but not BCD

(d) ABCD

110. Which graph(s) could potentially show a main effect of congruency

(a) AB

(b) BD

(c) AC

(d) ACD

111. Which graph(s) could potentially show an interaction between age and congruency?

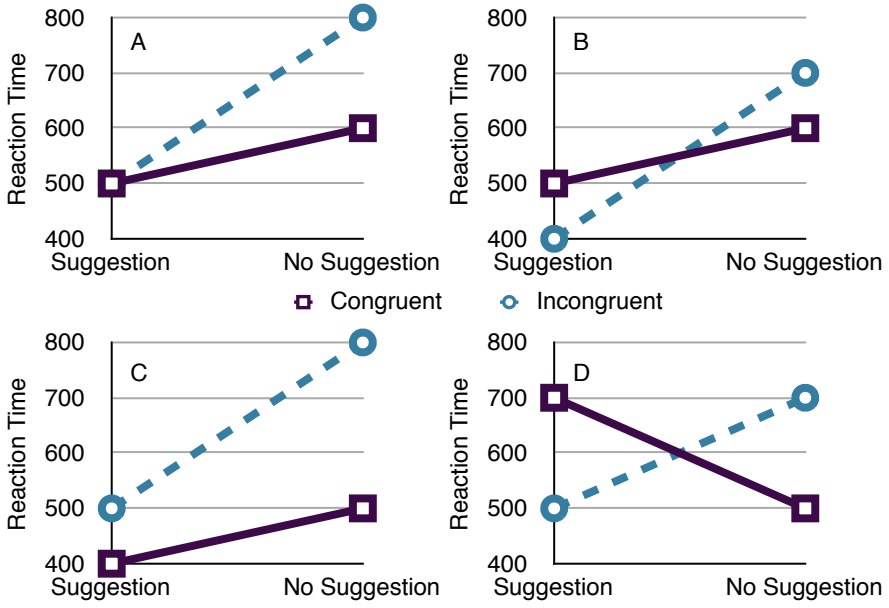
(a) AC
(b) CD
(c) BD
(d) AD

112. Which graph(s) could be used to describe to the following statement, “The pattern of Stroop effect was different for young vs. old subjects”.

(a) CD
(b) AD
(c) BC
(d) AB

113. Which graph(s) could be used to describe to the following statement, “The pattern of Stroop effect was not different for young vs. old subjects”.

(a) CD
(b) AB
(c) BC
(d) AD



114. Using the above graphs to answer the following questions. Raz et al (2009) conducted a Stroop experiment and tested whether a suggestion to “view the words as nonsense characters” would diminish the size of the Stroop effect. One group received the suggestion, and the other did not. Which graph (s) best reflect the prediction that suggestion would diminish the size of the Stroop effect

(a) BC
(b) AC
(c) AD
(d) AB

115. Which graph best supports the conclusion that the Stroop effect was entirely eliminated for subjects in the Suggestion Group.

(a) C
(b) D
(c) B
(d) A

116. Which graph does not show any main effects?

(a) A
(b) D
(c) B
(d) C

117. Which graph (s) could show 2 main effects?

(a) AB
(b) AC
(c) AD
(d) BD

118. All factorial designs

(a) have at least one manipulated independent variable and one nonmanipulated
(b) have the potential for producing at least two interaction effects
(c) have at least two independent variables
(d) have the potential for producing at least three main effects

119. A 2x3 factorial design has six

(a) different conditions
(b) independent variables
(c) levels of the independent variable
(d) subjects per condition

120. A 2x4 factorial design has

(a) all of the other options
(b) eight different conditions
(c) an independent variable with four levels
(d) two different independent variables

121. A 2x2x4 factorial design

(a) has sixteen different conditions
(b) has three dependent variables
(c) includes an independent variable with three levels
(d) includes eight different conditions

122. A 3x3x3 factorial design has three

(a) levels
(b) conditions
(c) independent variables

(d) dependent variables

123. A 3x3x3 factorial design has

(a) the potential for three different main effects
(b) three independent variables, each with three levels
(c) all of the options
(d) 27 different conditions

124. In a factorial study, what is a main effect?

(a) it occurs when the effect of one variable depends on the level of the other variable
(b) it refers to any statistically significant difference between the levels of a single independent variable
(c) it refers to any statistically significant finding in the study
(d) it is any result that is significant at the .01 rather than the .05 level

125. All of the following are true about factorial designs except

(a) they have at least two independent variables
(b) they are best analyzed with a 1-factor ANOVA
(c) if both main effects and interactions occur, interactions should be interpreted first
(d) the independent variables can be either between-subjects or within-subjects

126. In a 2 (sex) x 2 (type of therapy) factorial design, males were helped by therapy A, but females were helped by therapy B. This describes

(a) an interaction between gender and therapy type
(b) none of the above
(c) a main effect for therapy type
(d) a main effect for gender

127. In a 2x2 design, the following results occur: upper left cell in the factorial matrix = 20; upper right cell = 40; lower left cell = 40; lower right cell = 80. Which of the following is true?

(a) there is one main effect and an interaction
(b) there is an interaction but no main effects
(c) there are two main effects but no interaction
(d) there are two main effects and an interaction

128. In a 2x2 design, the following results occur: upper left cell in the factorial matrix = 20; upper right cell = 20; lower left cell = 30; lower right cell = 30. Which of the following is true?

(a) there is one main effect and no interaction
(b) there are two main effects and an interaction
(c) there is an interaction but no main effects
(d) there are two main effects but no interaction

129. In a 2x2 design, the following results occur: upper left cell in the factorial matrix = 20; upper right cell = 30; lower left cell = 30; lower right cell = 40. Which of the following is true?

(a) there is an interaction but no main effects
(b) there is one main effect and no interaction

(c) there are two main effects and an interaction

(d) there are two main effects but no interaction

130. In a 2x2 design, the following results occur: upper left cell in the factorial matrix = 20; upper right cell = 30; lower left cell = 30; lower right cell = 20. Which of the following is true?

(a) there is one main effect and no interaction

(b) there are two main effects but no interaction

(c) there are two main effects and an interaction

(d) there is an interaction but no main effects

131. A mixed factorial design has

(a) at least one between-subjects factor and at least one within-subjects factor

(b) both males and females serving as subjects

(c) at least one manipulated factor and at least one subject variable

(d) both two-level and multilevel factors operating

132. Suppose you have a balanced 2x2 design in which one of the variables is a between-subjects factor and one of the variables is a within-subjects factor. Further suppose that there will be 30 subjects in the upper left-hand cell of the 2x2 array. How many subjects are needed to complete the study?

(a) 60

(b) 120

(c) cannot be determined with the available information

(d) 30

133. Which of the following is true about a P x E factorial design?

(a) it includes both a subject variable and a manipulated variable

(b) it is a single-factor design

(c) both variables must be between-subjects variables

(d) both variables are manipulated variables

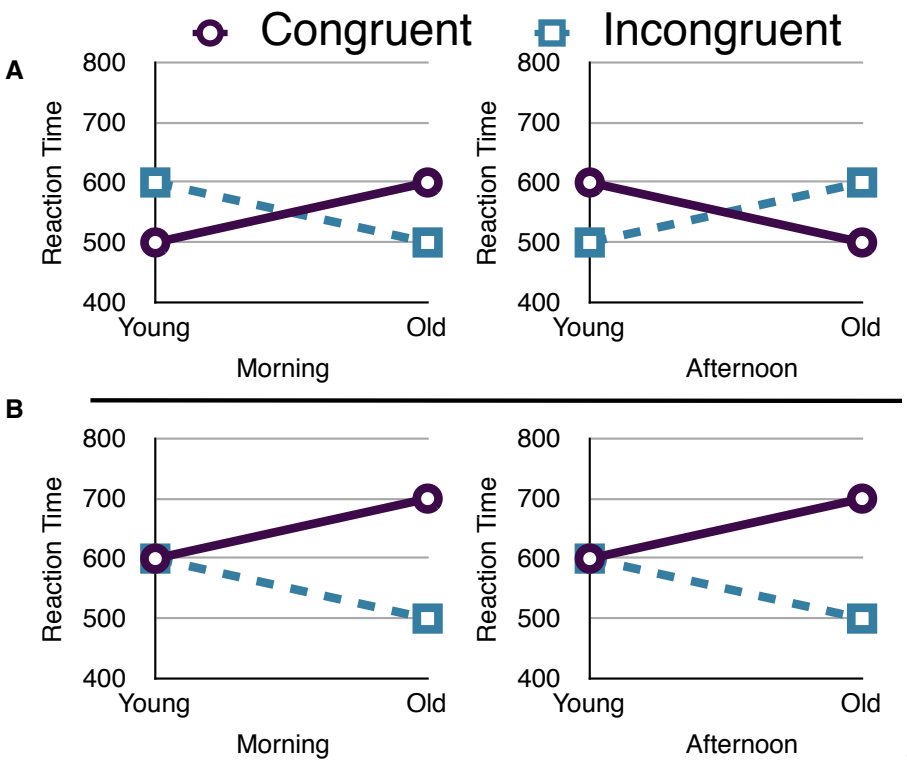
134. In a memory study using a 2x2 factorial, one of the factors is the presentation rate of the words, the two levels being 2 and 4 seconds per item. The researcher finds that recall is 98% accurate at 2 seconds per item and 99% accurate at 4 seconds per item (not a statistically significant difference). The most reasonable conclusion is that

(a) presentation rate has no significant effect on memory

(b) there is a main effect for presentation rate

(c) a ceiling effect has occurred

(d) an interaction has occurred



135. The above graphs show two sets of data for a 2x2x2 design, including congruency (congruent vs. incongruent), Age (old vs. young), and Time of Day (Morning vs. Afternoon).Interactions are said to exist when some effect changes or depends on the levels of another independent variable. All of these graphs show 2x2 interactions. However, which of these graphs (A or B) show a 3-way interaction, indicating that the 2x2 interaction changes across the levels of the 3rd independent variable?

(a) B

(b) A and B

(c) neither A nor B

(d) A