Research design

- No one research design is suitable for all topics
- Depending on what one wants to study, different designs may be appropriate, inappropriate, or even impossible
- 3 basic types:
  1. Case
  2. Experimental
  3. Correlational

1. Case method

- The simplest, most obvious, and most widely used—just look at it
- Closely studying a particular event or person in order to find out as much as possible
  - I.e. Alaska airlines plane crash in CA
- At its best, the case method yields not only explanations of particular events, but also more general lessons and possibly scientific principles
- Used by all sciences
  - I.e. volcano eruption, extinct animal, medical difficult cases, business—companies who’ve succeeded or failed
- Freud, Jung, Adler, Horney and Allport

1. Case method—Adv & Disadv

Advantages
1. Above all other methods, it’s the one that feels like it does justice to the topic
   - Describes the whole phenomenon and not just isolated variables
2. A well-chosen case study can be the source of general ideas
   - May reveal general facts about the situations studied
3. Sometimes the method is necessary
   - I.e. when a plane goes down—we must figure out why and how to prevent it
   - I.e. complicated medical cases

Disadvantages
1. It’s not controllable—of all of the facts which are crucial and which are incidental?

Good for 2 things
1. Gets a researcher interested in a problem
   - Newton’s apple falling—interest in gravity
2. It can be a source of ideas

2. Experimental study

- Exam the relationship between 2 things
  - I.e. between anxiety and test performance
- Divide participants by random assignment into 2 groups
  1. Experimental group
  2. Control group
- Random assignment allows for the presumption that the 2 groups or relatively equal in ability, personality, etc.
- Do something (manipulate the independent variable) to one group (exp grp) and do nothing to the other group (control grp)
  - I.e. tell them “your life depends on your performance on this test”
- Then give both groups a 30 item math test
- If anxiety hurts perf, then you’d expect the participants in the “your life depends” group to do worse on the test than the participants in the control group who didn’t get this message
- To test whether you got the results you predicted (hypothesis), write them down and display them
- Ensure that any difference is more than what would be expected from chance alone—a t-test
- See p. 68
3. Correlational study

- Measure the amount of anxiety that your participants already have
- Give them all a questionnaire asking them to rate how anxious they feel right now on a scale of 1-7
- Then give them the math test
- Now the hypothesis would be that if anxiety hurts perf, then those who scored higher on the anx measure will score worse on the math tests than those who received lower anx scores
- See p. 69

Comparing experimental and correlational methods

- These methods are often discussed as if they’re completely different and opposed.
- Both methods attempt to assess the relationship between 2 variables—anxiety and performance.
- The stats used in both studies are interchangeable.
- The statistic from the experiment can be converted, using algebra, into a correlation coefficient (r), and vice versa.
- One major difference between the 2 designs is that in the exp method the assumed causal variable (anxiety) is manipulated and in the corre the same variable is measured as it already exists.
- Advantages of exp:
  1. Ability to assess causality
     - Performance could have caused the anx.
- Disadvantages of exps:
  1. How do you know the statement made them anx?
  2. It can create levels of a variable that are unlikely or even impossible in real life
  3. Often requires deception
  4. Sometimes exp are simply not possible.
- Advantages of corre:
  1. Ability to assess causality
     - Performance could have caused the anx
- Disadvantages of corre:
  1. Either of 2 variables might have caused the other—the “third variable” problem
  2. Both variables might be the result of some other, unmeasured factor
     - Anxiety → poor performance OR poor performance → anxiety
Comparing experimental and correlational methods

- Are the participants fairly representative of the population to which the results will be generalized?
- Participants are not the only factor to generalize
- Stimuli and responses
- Research should be designed to sample across all of the domains to which the results will be generalized
- Should try to affect different kinds of performance that reflects the range of effects that exist in real life
- Sampling is almost nonexistent
  - The typical exp uses one kind of exp manipulation and measures just one behavior

Effect sizes

- Problems with significant testing
  - A significant result isn’t necessarily large, important, or dramatic, but probably
  - Significant at the “5%” level is different from zero to a degree that, by chance alone, would be expected about 5% of the time.
  - at the “1%” level by chance about 1% of the time and so is a stronger result.
  - P-level—probability level—of the difference in means in exp or some
    - Given the probability that the difference is really zero and so didn’t occur by chance, but its
      “real value” is probably not zero
    - Null hypothesis criterion
      - The criteria for a significant result is arbitrary
      - Why is the result of \( p < .05 \) significant, when a result of \( p < .06 \) isn’t?
    - The p-level only tells about the probability of one kind of error—the Type I error
      - Type II error
        - Deciding that one variable doesn’t have an effect on or a relationship with another variable, when really it does.
        - Unfortunately there isn’t any way to estimate the probability of a Type II error w/o making additional assumptions.
      - Psych research seems to be slowly moving away from it
        - What you really want to know from your data is, are your results important?
        - Sign testing was designed to answer this question, but isn’t really doing the job.
      - Correlations
        - Effect size
          - Rather than just stopping with significance, we also calculate a number that will reflect the size, as opposed to the likelihood, of that mean.
          - More meaningful than a sign level.

Calculating correlations

- Correlation coefficient
  - Can be used to describe the strength of the effect in a corre or exp study
  - Start with 2 variables
    - Arrange all of the scores on the two variables into 2 columns, with each row containing the scores for one participant
    - Label these columns X (variable you think is the cause) and Y (variable you think is the effect
    - \( r \) or \( r \) is the corre coeff
      - If calc’s are right: -1.0 to +1.0
      - If 2 variables are unrelated, the corre between them will be near zero
      - If the variables are positively associated—as one goes up, the other tends to go up too, like height and weight—then the corre coeff will be \( r > 0 \).
      - If the variables are negatively associated—as one goes up, the other goes down, like anx and perf—then the corre coeff will be \( r < 0 \).
      - If 2 variables are corre \( r = 0 \) or \( r \) this means that one of them can be predicted from the other
    - I.e. If I known how anx you are, I can predict (to a degree) how well you’ll do on a maths test
    - You can also be a corre coeff from exps
      - \( X \) and \( Y \) are the corre coeff for the same technique.

Representative design

- Are the participants fairly representative of the population to which the results will be generalized?
- Participants are not the only factor to generalize
- Stimuli and responses
- Research should be designed to sample across all of the domains to which the results will be generalized
- Should try to affect different kinds of performance that reflects the range of effects that exist in real life
- Sampling is almost nonexistent
  - The typical exp uses one kind of exp manipulation and measures just one behavior

Effect sizes

- Problems with significant testing
  - A significant result isn’t necessarily large, important, or dramatic, but probably
  - Significant at the “5%” level is different from zero to a degree that, by chance alone, would be expected about 5% of the time.
  - at the “1%” level by chance about 1% of the time and so is a stronger result.
  - P-level—probability level—of the difference in means in exp or some
    - Given the probability that the difference is really zero and so didn’t occur by chance, but its
      “real value” is probably not zero
    - Null hypothesis criterion
      - The criteria for a significant result is arbitrary
      - Why is the result of \( p < .05 \) significant, when a result of \( p < .06 \) isn’t?
    - The p-level only tells about the probability of one kind of error—the Type I error
      - Type II error
        - Deciding that one variable doesn’t have an effect on or a relationship with another variable, when really it does.
        - Unfortunately there isn’t any way to estimate the probability of a Type II error w/o making additional assumptions.
      - Psych research seems to be slowly moving away from it
        - What you really want to know from your data is, are your results important?
        - Sign testing was designed to answer this question, but isn’t really doing the job.
      - Correlations
        - Effect size
          - Rather than just stopping with significance, we also calculate a number that will reflect the size, as opposed to the likelihood, of that mean.
          - More meaningful than a sign level.

Calculating correlations

- Correlation coefficient
  - Can be used to describe the strength of the effect in a corre or exp study
  - Start with 2 variables
    - Arrange all of the scores on the two variables into 2 columns, with each row containing the scores for one participant
    - Label these columns X (variable you think is the cause) and Y (variable you think is the effect
    - \( r \) or \( r \) is the corre coeff
      - If calc’s are right: -1.0 to +1.0
      - If 2 variables are unrelated, the corre between them will be near zero
      - If the variables are positively associated—as one goes up, the other tends to go up too, like height and weight—then the corre coeff will be \( r > 0 \).
      - If the variables are negatively associated—as one goes up, the other goes down, like anx and perf—then the corre coeff will be \( r < 0 \).
      - If 2 variables are corre \( r = 0 \) or \( r \) this means that one of them can be predicted from the other
    - I.e. If I known how anx you are, I can predict (to a degree) how well you’ll do on a maths test
    - You can also be a corre coeff from exps
      - \( X \) and \( Y \) are the corre coeff for the same technique.

Representative design

- Are the participants fairly representative of the population to which the results will be generalized?
- Participants are not the only factor to generalize
- Stimuli and responses
- Research should be designed to sample across all of the domains to which the results will be generalized
- Should try to affect different kinds of performance that reflects the range of effects that exist in real life
- Sampling is almost nonexistent
  - The typical exp uses one kind of exp manipulation and measures just one behavior

Effect sizes

- Problems with significant testing
  - A significant result isn’t necessarily large, important, or dramatic, but probably
  - Significant at the “5%” level is different from zero to a degree that, by chance alone, would be expected about 5% of the time.
  - at the “1%” level by chance about 1% of the time and so is a stronger result.
  - P-level—probability level—of the difference in means in exp or some
    - Given the probability that the difference is really zero and so didn’t occur by chance, but its
      “real value” is probably not zero
    - Null hypothesis criterion
      - The criteria for a significant result is arbitrary
      - Why is the result of \( p < .05 \) significant, when a result of \( p < .06 \) isn’t?
    - The p-level only tells about the probability of one kind of error—the Type I error
      - Type II error
        - Deciding that one variable doesn’t have an effect on or a relationship with another variable, when really it does.
        - Unfortunately there isn’t any way to estimate the probability of a Type II error w/o making additional assumptions.
      - Psych research seems to be slowly moving away from it
        - What you really want to know from your data is, are your results important?
        - Sign testing was designed to answer this question, but isn’t really doing the job.
      - Correlations
        - Effect size
          - Rather than just stopping with significance, we also calculate a number that will reflect the size, as opposed to the likelihood, of that mean.
          - More meaningful than a sign level.

Calculating correlations

- Correlation coefficient
  - Can be used to describe the strength of the effect in a corre or exp study
  - Start with 2 variables
    - Arrange all of the scores on the two variables into 2 columns, with each row containing the scores for one participant
    - Label these columns X (variable you think is the cause) and Y (variable you think is the effect
    - \( r \) or \( r \) is the corre coeff
      - If calc’s are right: -1.0 to +1.0
      - If 2 variables are unrelated, the corre between them will be near zero
      - If the variables are positively associated—as one goes up, the other tends to go up too, like height and weight—then the corre coeff will be \( r > 0 \).
      - If the variables are negatively associated—as one goes up, the other goes down, like anx and perf—then the corre coeff will be \( r < 0 \).
      - If 2 variables are corre \( r = 0 \) or \( r \) this means that one of them can be predicted from the other
    - I.e. If I known how anx you are, I can predict (to a degree) how well you’ll do on a maths test
    - You can also be a corre coeff from exps
      - \( X \) and \( Y \) are the corre coeff for the same technique.
Interpreting correlations

- To interpret a corre coeff isn't just using stats sign
- A corre becomes sign in a stats sense merely by probably not being zero, which depends as much on how strong the effect really is
- Commonly eval effect sizes by squaring them
  - A corre of .30 means that “only” 9% of the variance is explained by the corre
  - The effect sizes expressed in corre rarely exceed .40

The binomial effect size display (BESD)

- Demonstrates in a concrete manner how big hose effect-size corr es really are
  - i.e. 200 participants—all are sick—exp drug is given to 100 and nothing to 100—results of 100 alive and 100 dead: how much difference did the drug make?
  - This method begins by assuming a corre of zero, giving the 4 cells in the table an entry of 50—if there’s no effect, 50 pers rec rx will live and 50% before
  - Then take the actual corre (.40 here), remove the decimal (.40-.40), divide by 2 (.20) and add it to the 50 in the upper left (yielding 20). Then adjust the other 3 cells by subtraction
- BESD—corre effects need to be interpreted more carefully than usually are
- Shows vividly how much of the effect of an exp is likely to have, and how well one can predict an outcome from an individ measurement of difference

Ethics

- The uses of psychological research
  - Results may be used for harmful purposes
  - Truthfulness
  - Researchers plagiarizing the work of others
  - Science is based on truth and trust
  - All scientists must trust each other for the process to work
- Deception
  - The purpose of such research is to make the research “realistic”
  - Milgram’s study
  - APA ethical guidelines
  - Should be allowed
    1. Participants have given their “informed consent” to be deceived
    2. Little white lies told to participants usually do no harm
    3. Certain topics can’t be investigated w/o the use of some deception