

# Survey Research and Design in Psychology

## Lecture 10 - Summary and conclusions

Dr Ananthan Ambikairajah

University of Canberra

# What we have covered

- ▶ Survey research and design (lectures 1 & 2)
- ▶ Univariate and bivariate descriptives and graphs (lectures 3 & 4)
- ▶ Psychometrics (lectures 5 & 6)
- ▶ Multiple linear regression (lectures 7 & 8)
- ▶ Power and effect sizes (lecture 9)

- ▶ What is a survey?
  - ▶ A standardised stimulus designed to convert fuzzy psychological phenomenon into hard data
- ▶ Why do we do research?
  - ▶ To help us gather information, or test and build theories

# Pros and cons of survey research

- ▶ Pros:
  - ▶ Ecological validity
  - ▶ Cost efficiency
  - ▶ Can obtain lots of data
- ▶ Cons:
  - ▶ Low compliance
  - ▶ Reliance on self-report

- ▶ Self-administered surveys are:
  - ▶ Cheaper
  - ▶ Suffer less from demand characteristics
  - ▶ Allow us to access a representative sample
  - ▶ Anonymous
  - ▶ ... but they suffer from non-response, and it is difficult to adjust for factors such as cultural differences and special needs
- ▶ Survey design is an art and a science

- ▶ Objective versus subjective
  - ▶ Objective - verifiable true answer
  - ▶ Subjective - perspective of respondent
- ▶ Open versus closed
  - ▶ Open - empty space for answer
  - ▶ Closed - pre-set response options

- ▶ Nominal: Arbitrary numerical labels, could be in any order
- ▶ Ordinal: Ordered numerical labels, intervals may not be equal
- ▶ Interval: Ordered numerical labels, equal intervals
- ▶ Ratio: Continuous, meaningful 0

- ▶ Sampling requires an understanding of your target population, sampling frame, and sample
- ▶ We can use probability sampling methods (random), including simple, systematic, or stratified
- ▶ Or we can use non-probability sampling, including convenience, purposive, and snowball sampling

- ▶ Sampling biases - Sample doesn't represent target population
- ▶ Non-sampling biases - Measurement tool issues (reliability and validity)
  - ▶ Response biases
  - ▶ Acquiescence, Order effects
  - ▶ Demand characteristics
  - ▶ Self-serving bias, Social desirability, Hawthorne effect

- ▶ What is the central tendency?
  - ▶ Frequencies, Percentages
  - ▶ Mode, Median, Mean
- ▶ What is the variability?
  - ▶ Min, Max, Range, Quartiles
  - ▶ Standard Deviation, Variance

- ▶ Nominal/Ordinal data
  - ▶ Bar graph or pie chart
- ▶ Interval/ratio data
  - ▶ Histogram, stem & leaf plot, box plot

## Correlation and covariation

- ▶ The world is made of covariations.
- ▶ Covariations are the building blocks of more complex multivariate relationships.
- ▶ Correlation is a standardised measure of the covariance (extent to which two phenomenon co-relate)
- ▶ Correlation does not prove causation

- ▶ What is the relationship/association/shared variance/co-relation between two variables?
- ▶ To what extent do two variables covary/depend on one another/explain one another?

- ▶ Nominal X nominal: Chi-Square with Phi ( $\phi$ ) or Cramer's V
- ▶ Ordinal X ordinal: Spearman's rank or Kendall's Tau ( $\tau$ ) *b*
- ▶ Dichotomous X interval/ratio: Point bi-serial  $r_{pb}$
- ▶ Interval/ratio X interval/ratio: Product-moment/Pearson's  $r$

# The steps we take

- ▶ Choose the type of correlation and graph (based on level of measurement)
- ▶ Check assumptions (e.g. linearity, outliers)
- ▶ Report effect size, direction, significance, and percent of shared variation
- ▶ Be wary of limitations (correlation  $\neq$  causation)

- ▶ Factor analysis:
  - ▶ Summarizes correlations amongst items
  - ▶ Common clusters form factors

- ▶ Check the assumptions
- ▶ Choose extraction method and rotation
- ▶ Determine the number of factors (based on an examination of Eigen Values, Scree plot, percentage of variance explained)
- ▶ Select items (based on factor loadings)
- ▶ Name and describe factors
- ▶ Examine correlations amongst factors
- ▶ Analyse internal reliability
- ▶ Compute composite scores to use factors as subscales

## The steps we take

- ▶ Ask yourself: What is the concept that I am trying to measure?
- ▶ Brainstorm possible indicators of the concept, define it, draft measurement items
- ▶ Pre-test and pilot-test them, examine the psychometric properties, and redraft/re-test

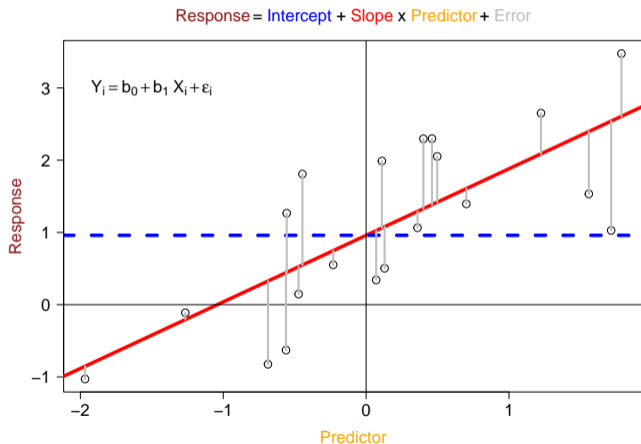
- ▶ Measurement error = Deviation of measure from true score
- ▶ Sources = Non-sampling (e.g., paradigm, respondent bias, researcher bias) or Sampling (e.g., non-representativeness)
- ▶ Minimise measurement error by using:
  - ▶ Well-designed measures
  - ▶ Representative samples and maximising response rate
  - ▶ Reducing demand effects
  - ▶ Ensuring administrative accuracy – entering data

- ▶ Reliability is about consistency of:
  - ▶ Items within the measure
  - ▶ The measure over time
- ▶ Validity is about whether the measure actually measures what it is intended to measure

## The steps we take

- ▶ Develop model and hypotheses
- ▶ Check assumptions
- ▶ Choose type
- ▶ Interpret output
- ▶ Develop a regression equation (if needed)

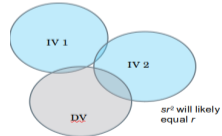
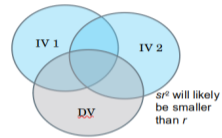
- ▶ Best-fitting straight line for a scatterplot of two variables



- ▶ Using multiple independent variables to predict a single dependent variable
- ▶ Overall fit and coefficients
- ▶ Three types: standard, hierarchical, stepwise

# Semi-partial correlations in multiple linear regression

- ▶ Semi-partial correlations ( $sr$ ) can be used to get  $sr^2$
- ▶  $sr^2$  indicates the percentage of variance in the dependent variable which is uniquely explained by an independent variable
- ▶ We can compare each  $sr^2$  with the  $r^2$  (or  $sr$  with the  $r$ ), and if they differ, think about why this might be the case. If it stays the same, there may be no overlap between predictors. If it drops substantially, then there is possibly a strong relationship between predictors.



- ▶ In multiple linear regression independent variables may interact to:
  - ▶ Have no effect
  - ▶ Increase the independent variable's effect on the dependent variable
  - ▶ Decrease the independent variable's effect on the dependent variable
- ▶ Model interactions using hierarchical multiple linear regression:
  - ▶ Step 1: Enter independent variables
  - ▶ Step 2: Enter cross-product of independent variables
  - ▶ Examine change in  $R^2$

- ▶ Analysis of changes over time can be assessed either by:
  - ▶ Standard regression:
    - ▶ Calculate difference scores (post-score minus pre-score) and use as a dependent variable
    - ▶ Doesn't tell you how much additional variance is accounted for by independent variables (over and above the baseline dependent variable)
  - ▶ Hierarchical multiple linear regression:
    - ▶ Step 1: "partial out" baseline scores
    - ▶ Step 2: enter other independent variables to help predict variance in changes over time

## Significance testing

- ▶ Logic – At what point do you reject the null hypothesis?
- ▶ History – Started in 1920s & became very popular through 2nd half of 20th century
- ▶ Criticisms – Binary, dependent on sample size, effect size, and critical  $\alpha$
- ▶ Practical significance
  - ▶ Is an effect noticeable? Is it valued? How does it compare with benchmarks?

- ▶ Effect size - standardised difference or strength of relationship
- ▶ Inferential tests should be accompanied by effect sizes and confidence intervals
- ▶ Common bivariate effect sizes include:
  - ▶ Cohen's  $d$
  - ▶ Correlation  $r$

# What is going wrong in psychology?

- ▶ Publication bias: Tendency for studies with statistically significant findings to be published more often
- ▶ Replication crisis
- ▶ Academic integrity issues (e.g. Retraction Watch)

# Tying it all together

- ▶ What kind of researcher do you want to be?

- ▶ Think about what you have learnt in this class:
  - ▶ When you're reading (e.g. articles in journals, news sites, shared on social media)
  - ▶ When you're working (preparing summaries/reports, evaluating a program, making graphs or infographics)
  - ▶ When you need to prove a point (literature search)
  - ▶ When you're filling in a survey or poll, or interpreting the results of them
  - ▶ When making important decisions (what does the evidence say I should do?)

Thank you

▶ Thank you and good luck!

# Contributions to this course

Dr James Neill

Dr Samantha Stanley

Dr Jeroen van Boxtel