Is random assignment used?

Counterbalancing

- Complete counterbalancing and partial counterbalancing (enough groups that each condition is occurs in each ordinal position)
- Possible orders:
  - 2 groups – 2 orders
  - 3 groups – 6 orders
  - 4 groups – 24 orders
  - 5 groups – 120 orders
- Latin Square: Each condition is in each order but always proceeds the same condition, i.e. 1234, 2341, 3412,4123

Internal Validity – Are the findings valid?

- Threats:
  - Third variables - confounding variables
  - Extraneous variables
  - History (events that happen during study)
  - Maturation (participant changes i.e. height, weight)
  - Instrumentation (technical issues, i.e. changing testing materials part way through)
  - Testing effects (practice, fatigue)
  - Regression towards the mean

External Validity – Generalisation

- Threats:
  - Volunteer bias
  - Representation of the population
  - Using different measures (i.e. repeated measures)
  - Novelty effect (acting differently in a new situation)
  - Reactivity (acting differently when being observed)
  - Multiple treatment interference (effects of previous treatments)
  - Experimenter characteristics
  - Timing of measurements

Types of Research

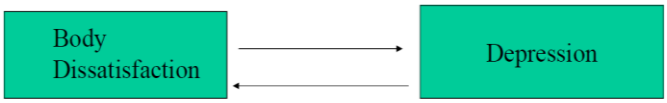
Subject Design
Between – Independent
Within – Repeated-measures
Time related effects:

- Long term: history, maturation
- Short term: order effects
  - Carryover – long lasting effects, resolved by using between subjects
  - Progressive error – fatigue, practice, resolved with counterbalancing

  - High external, low internal
- Correlational
  - Looks at relationships
  - High external, low internal
- Descriptive
  - Observational, case studies
- Qualitative

Directionality Problem

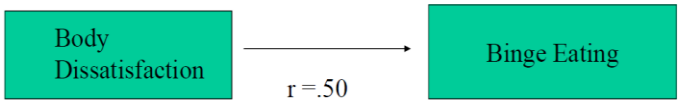
Which causes which?



r = .60

Third Variable Problem

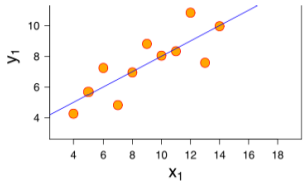
Depression can cause both of these and can be their connection

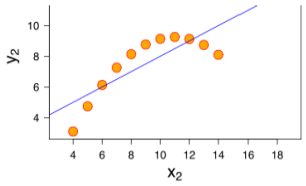


r = .50

Correlation

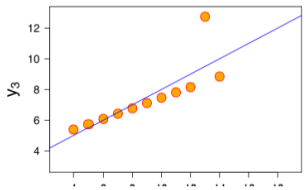
- Degree of linear association between two variables
- Measures:
  - Direction (positive, negative)
  - Form (linear, nonlinear)
  - Degree/strength (weak, moderate, strong – measured with r-value)
- Relationships:
  - Positive – score high in A, score high in B
  - Negative – score high in A, score low in B

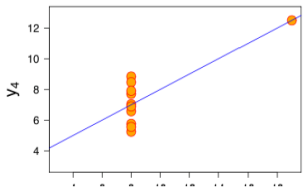




Should be plotted as a quadratic/curved relationship

Outlier skews line of best fit - removing outlier or reducing influence will make it a better prediction model





Not a good model or measure for the data; It's a violation of homogeneity of variance of arrays

All four graphs below have Mean of X = 9, variance of X = 11, Mean of Y = 7.5, variance of Y = 4.122 or 4.127, correlation between x and y = 0.816, linear regression line: y = 3 + 0.5X

Bivariate Regression

- Used for prediction when variables correlate
- The stronger the correlation, the more reliable the prediction
- Predictor: IV, x-axis
- Predicted (criterion): DV, y-axis
- Regression lines – line of best fit
  - Finds 'centre' of relationship, makes it easier to see but not always reliable
  - Equation: $\hat{Y} = bX + a$
  - *Predicted y value = slope* x *X value + y intercept*

R Values

- Correlation Coefficient (r)
- Coefficient of determination (r^2)
  - i.e. r = 0.7, r^2 = 0.49, this means 49% of variance in one variable is explained by the other; 51% is *not* explained by the other variable
- Significance tests
  - Transform r so it takes on a normal distribution
  - Can use z scores or SPSS t-test

Correlations

		Anxiety	Negative Mood
Anxiety	Pearson Correlation	1	.751*
	Sig. (2-tailed)		.012
	N	10	10
Negative Mood	Pearson Correlation	.751*	1
	Sig. (2-tailed)	.012	
	N	10	10

*. Correlation is significant at the 0.05 level (2-tailed).

Factors that Affect Correlation

- Range restriction: only collecting a small amount of data can prevent us from seeing the actual trend
- Heterogenous sample: overall correlation of two groups may differ to their separate correlations
  - i.e. Males might have a low rate of X and females a high rate of X but when combined, the average turns into a moderate rate of X
- Outliers: out of range values
  - Univariate: scores three SD away from the mean
    - Regression analyses can check the influence of an outlier, outliers can be recoded into range, bootstrapping reduces outliers influence
  - Multivariate: scores out of range on numerous variables
    - i.e. age (15 - 60) vs. income ($5,000 - $500,000)
    - Age = 16 and income = $200,000 are both normal but combined are an outlier
    - Regression and bootstrapping can be used but the case can also be deleted