Correlational Research

Collecting data to determine to what degree a relationship exists between two or more quantifiable variables

This degree is expressed as a *correlation* coefficient

Correlational Research

- The purpose of the research is:
  - To determine relationships between variables
  - To use these relationships to make predictions

Correlational Research

- Used to study variables that may be related to complex variable
- Variables with a low degree of relationships are discarded
- Variables that are highly related may be further studied in causal – comparative or experimental studies.

Selection of Problem

- Relationships of variables should be based on theory or experience
  - Avoid treasure hunts and shotgun or fishing approaches.

Selection of Participants

- Use acceptable methods
- Rule of thumb – 30 participants minimum
- Low validity and reliability requires an increase in sample size

Instrument Selection

- Must be valid and reliable
- Must represent intended variables

Correlation Design

- Two (or more) scores are obtained for each sample member (one for each variable)
- Paired scores are correlated
- Result is expressed as a correlation coefficient

Data Analysis

- Correlation coefficient
  - Indicates size and direction of relationships
  - Decimal number 1.00 to 0 to –1.00
  - Near +1.00 - high size and positive direction (person with high score on variable will have high score on second variable)
  - Near 0.0 – variables are not related
  - Near –1.00 - high size and negative direction

Correlation Coefficient

- DOES NOT indicate the percentage of the relationship between variables
- A correlation coefficient of .70 does not mean that the variables are 70% related

Correlation Coefficient Squared

- Indicates the amount of common shared variance between variables
- If the correlation coefficient is .76 = 58% (approx) of common variance between variables which also means that 42% (approx) of the variance in unexplained

Interpretation

- Prediction study
  - Correlation coefficient must be high
- Exploration or testing of hypothesis
  - Must have statistical significance

Statistical Significance

- Depends on sample size
- Small samples require high Correlation coefficients

Statistical Significance

- Tests to determine if correlation reflects a true relationship or chance
- Indicates the probability of a true significant relationship
- Use Table A.2 and Tables 2 & 3 in the article

Level of Confidence

- As level of confidence increases the value of the correlation coefficient needed for significance increases
- .05 = 95% .01 = 99%

Relationship vs Causality

- Correlation coefficients reflect relationships between or among variables NOT causality

Cause – Effect Relationships

- Can be established by experimental research

Importance of Relationship Studies

- Suggest relationships between variables for further study / causal – comparative and experimental
- Allows for removal of influence of variables on dependent variables in other types of studies
- Allows identification of variables that can be excluded from further studies

Data Collection

- 1st identify the variables
  - Logically
  - Based on literature review
- 2nd identify appropriate population
  - Must be able to collect data on all identified variables

Data Analysis

- Identify variable or primary interest (VPI)
- Correlate each of the other variables to the VPI
- Most common computing

Pearson *r*

- Used on continuous data
  - Ratio data
  - Interval data
- Most precise estimate of correlation

Spearman *rho*

- Used for rank difference correlation
- Rank data
  - Participants ranked in order of score
  - Two participants with same score are averaged (2 with highest score are ranked 1 and 2 then averaged 1.5)
  - Variables to be correlated must be expressed in terms of rank
  - Easier to compute with small numbers of participants (less than 30) than Pearson *r*

Other types of analysis

- Categorical dichotomies (329)
- Artificial dichotomies (3290
- Turn to page 330 Table 9.2

Linear vs Curvilinear

- Linear
  - –Linear relationships can be plotted in straight line
  - –Perfect relationship = straight line
  - –No relationship = scattered
- Curvilinear
  - –Increase in one variable results in increase in second variable
  - –Then increase in one variable results in decrease in second variable

Subcategories in Correlational Studies

- Can be valuable to generalize to populations
- Can only be done with larger sample sizes
- Use stratified samples to ensure similar numbers in groups

Factors that Contribute to Inaccuracies

- Attenuation (unreliability)
- Corrects for un-reliable measures
- Should not be used in prediction studies
- Restricted range of scores (spread)
- More variability in set of scores the higher the correlation coefficient

Prediction Studies

- Predictor – variable used to predict another variable's scores
- Criterion – variable that is predicted
- Studies are used:
  - To facilitate decisions
  - To test predictors
  - To determine predictive validity of measures
  - By counselors, admission directors, employers, and researchers

Prediction Studies

- Several variables correlated with each other can provide better prediction

Data Collection

- Use valid measures esp. for criterion variable
- Define terms quantifiably ways

Data Collection

- Predictor variables are usually obtained before criterion variables
- Data may span over time
- Shrinkage – tendency of prediction equation to become less accurate when used with subsequent groups
- Cross validation – validation of prediction equation with at least one other group (after original) and discarding variables that do not show relationship

Data Analysis

- Use statistical processes

Data Interpretation

- Single predictive studies
  - Uses single predictor
  - Interpretation formula (page 333)
    $y = a + bx$

Multiple Regression Equation

- Contains more predictors
- Formula page 334

Accuracy of Predictions

- Predictors and criterion variables must be reliable
- Longer length of time between predictor and criterion variables the lower the accuracy
- General criterion variables have lower prediction accuracy than narrow criterion variables

Standard Error

- Refers to the predicted range of sample score

Other Types of Analysis

- Discriminate function analysis
  - Criterion variable is categorical, not continuous
  - Predictor variables are continuous
- Path analysis
  - Shows relationships and patterns among variables
  - Shown as diagram

Other Types of Analysis

- Canonical correlation
  - Produces correlation between a group of predictor variables and criterion variables
- Factor analysis
  - Way to take large number of variables and break them into smaller parts
  - Computes correlation among all variables then groups them together