Quasi-Experimental Designs

Dewayne E Perry
ENS 623
perry@mail.utexas.edu
Non-random Assignments

- Random assignment not always possible
- Alternative approaches serve as approximations
- Quasi - same except for randomization
- Basic presumption: groups are non-equivalent
  - Result: internal validity threatened by a full range of threats
- Three classes of designs
  - Non-equivalent group designs
  - Interrupted time-series designs
  - Correlational designs
Non-equivalent Group Designs

- Most widely used in quasi experiments
- Pre/post measures on treatment/control
- Problem: expect subjects in different groups to differ because assignment not controlled
  - Must make assumptions about variables

Alternatives
- Randomization after assignment into treatment and control if sample mandated
- Match groups as closely as possible
- Non-volunteers as wait-list; compare against volunteers
- Compare different amount of treatments
Interrupted Time Series

- Effects of treatments are inferred
  - Compare outcome measures at different time intervals
    - A single data point for each point in time
  - Before and after treatment is introduced
    - A clear dividing line at the beginning of treatment
  - Four considerations
    - Need a sufficient number of data points
    - Same units throughout equally spaced
    - Sensitive to the particular effects being studied
    - Measurements should not fluctuate
Box-Jenkins Procedure

- **Auto-regressive integrated moving average**
- **Aim:** identify underlying model of serial effects
  - Abrupt change at point of treatment
  - Gradual constant changes in levels
  - Abrupt change but lasting only a short while – a pulse

- **Assumptions**
  - The series of observations must be stationary
    - Fluctuate around the mean rather than drift
    - Secular trend handled by differencing
      - $2 3 4 5 6 \rightarrow 1 1 1 1 1$
  - **Autocorrelation**
    - Dependence or independence of observations on each other
    - Regular: adjacent observations on one another
    - Seasonal: observations separated by a period
Single Case, Small N

- **N=1, single case**
  - Widely used to evaluate effects of behavioral control treatments
  - Widely use in SWE
  - Problematic to call these experimental – randomization not a consideration at all

- **Argument**
  - Subjects serve as own controls
    - Behavior monitored as treatment effects replicated over time
    - Changes in patterns of performance are basis for inferences about treatment
Single Case, Small N

- Start by establishing a behavioral baseline:
  - the continuous, continuing performance of a single individual
- Found niche for effects of clinical, counseling and educational interventions
- Prototype procedure: A B A (variant of AB)
  - A is pretreatment phase
  - B denotes introduction of independent variable
  - A treatment is withdrawn at the end and behavior measured
- Variants - non unambiguous wrt internal validity
  - A B BC B
    - to tease out effects of BC and B alone
  - A B A B
    - To emphasize positive effects of treatment variable
- Seldom report elaborate statistical analyses, but use good graphical representations
Cross-Lagged Panels

- Frequent in past, now employed with skeptical advocacy
- **Cross-lagged:**
  - a time series design
  - some data treated as temporarily lagged values of the outcome variable
- **Panel:** another name for longitudinal
  - Two motivations
    - Increase precision by measuring each subject in all conditions
    - Examine individuals change response over time
Cross-Lagged Panels

- Assumption: longitudinal measurements of same two variables (A, B) would provide information about causal relationship between them
- Hence: a method for choosing between competing causal hypotheses
Cross-Lagged Panels

- 3 sets of paired correlations
  - **Test-retest:** \( r_{A1A2}, r_{B1B2} \)
    - Indicates reliability of A and B over time
  - **Synchronous:** \( r_{A1B1}, r_{A2B2} \)
    - Reliability of relationship between A and B over time
  - **Cross-lagged:** \( r_{A1B2}, r_{B1A2} \)
    - Relationship between two sets of data points
    - Is A a stronger cause of B than B of A
    - Yes if \( r_{A1B2} \) is higher than \( r_{B1A2} \)
      - Eg, \( r_{A1B2} = 0.585 \) and \( r_{B1A2} = 0.405 \)
Cross-Lagged Panels

- Interpretability considered maximum when r values remain the same at each period
- However, seldom stationary
  - Temporal erosion
  - Attenuation leaves us with a residual effect
- Seldom reliable and clear cut
  - seldom a clear inference
Path Analysis

- Eg, violence in TV and aggression
  - Boys
  - Time periods 1960 and 1970
  - A1 and A2 - preference for violent TV
  - B1 and B2 - peer-rated aggression
  - $r_{A1A2} = 0.05$, $r_{B1B2} = 0.38$
  - $r_{A1B1} = 0.21$, $r_{A2B2} = -0.05$
  - $r_{A1B2} = 0.31$, $r_{B1A2} = 0.01$
  - Measures of aggression: who starts fights, takes others' things
  - Predictors of aggression: three favorite TV shows
  - Data indicates some not very reliable relationships
    - AB positive in 1960, negative in 1970
    - Test-retest only .05 for TV, .38 for aggression
    - Statistically a significant relationship between violent TV in 1960 and aggressive behavior in 1970 (.31)
    - Alternative causal pattern quite negligible (.01)
Cross-Lagged Panels

- **Plausible inferences**
  - Not possible to demonstrate a particular hypothesis is true
  - Possible to reject untenable hypotheses and narrow down rival explanations

- **A1 -> B2 – 5 plausible hypotheses**
  - **1**: A1 -> B1, A1 -> B2
    - Preferring to watch violent TV is a direct cause of aggressive behavior
    - \( r_{A1B1} = .21, r_{A1B2} = .31 \) is consistent with this
    - Low test-retest might be explained by different overtly violent activities in teens
  - **2**: A1 -> B1, B1 -> B2
    - Preference for violent TV stimulates children to be aggressive and carries over into teen years
    - Ruled out: correlation between A1B2 much higher than \( r_{A1B1} \times r_{B1B2} \)
Cross-Lagged Panels

★ 3: \( B1 \rightarrow A1, A1 \rightarrow B2 \)
   - Aggressive children prefer violent TV
   - Ruled out for reasons similar to above
   - \( r_{B1B2} \) much higher than product

★ 4: \( B1 \rightarrow A1, B1 \rightarrow B2 \)
   - Aggressive children are more likely to watch violent TV and to become aggressive teenagers
   - Not so easily rejected
   - Did a partial correlation
     ✓ Removed other influences:
       • \( A1 \) and \( B2 \) controlling for \( B1 \)
     ✓ Very close to original - .25 vs .31
   - Hence, implausible as complete causal explanation
Cross-Lagged Panels

★ 5: \( B1 \rightarrow A2, B1 \rightarrow B2 \)

- Early aggression causes both a weaker preference for violent TV as a teenager and a penchant to continue to be aggressive
- Rejected: needed cross correlation for this basis of rejection
  - \( r_{B1A2} = .01 \) was very close to comparison base

♦ Thus ruled out 2-5, leaving 1
  - Watching violent TV was a direct causal link to aggressive behavior in some viewers
Cohort Designs - Utility

- A wider set of longitudinal
  - Pure: one cohort followed over time
  - Mixed - several cohorts followed
- Age, time and cohort effects
  - Eg, believed that IQ increase to a maximum at age 30 and then declined
  - Confounded age and cohort effects
    - Cohort: different life experiences etc
  - Diachronic designs: changes in successive periods of time
  - Useful in uncovering relationships that remain shrouded in synchronic designs
Cohort Designs - Limitations

- Example of age and no religious affiliation of women in The Netherlands
  - Clearly cross-sectional conclusions cannot be correct
  - With full cohort data can do other analyses
    - Avoid fallacy of period centrism
      - One time period generalizable to another
    - Age effect: due to natural aging process
    - Time of measurement effect: impact of events on time that occur at points of measurement
    - Cohort effect: represents past history
Cohort Designs - Limitations

- Comparison where age, time and cohort effects are the major variables
  - Simple cross-sectional
    - Limitation: confounds age of subject with age of cohort
  - Simple longitudinal
    - Limitation: does not control for effects of history
      - Different results might be obtained using a different period of time
  - Cohort sequential
    - Takes into account age and cohort. But not the time of measurement fully
  - Time sequential
    - Does not take into account cohort
  - Cross-sectional
    - Does not take age fully into account

- Each has limitations
  - Hence best to employ a variety of methods