



# Longitudinal research

# Issues in Longitudinal Research

---

- Longitudinal research, like cross-sectional research, relies upon three fundamental methods of gathering data:
  - asking people questions
  - observing people's behavior
  - observing the physical traces or results of peoples behavior

# Issues in Longitudinal Research

---

- problems of internal and external validity
- measurement reliability
- sampling
- appropriateness of questions to the populations being studied
- adequacy of the randomization procedures in experimental designs
- effects of interaction between subjects or respondents and experimenters, interviews, or observers (in micro social data collection)
- relevance of the research (do we measure what is important or just what is easily measureable?)

# Issues in Longitudinal Research

---

## ■ **Genesis versus Prediction**

- Genesis focuses on stages or sequences of qualitative changes, with the goal of discovering laws of growth or developmental change.
- Prediction according is concerned with the stability of population characteristics over time and the extent to which external influences (changes in environment, therapeutic intervention) may modify those characteristics

To obtain a more complete account of patterns of development, then, we may study both the sequencing of behavior and the timing and correlates of progression from one stage to another.

# Issues in Longitudinal Research

---

## ■ **Changes in Measurement Over time**

Some researchers have suggested that the findings from participant observation studies may be differently “idiosyncratic and difficult to replicate” (Blalock & Blalock, 1982:97)

# Issues in Longitudinal Research

---

## ■ **Changes in Measurement Over time**

Two important questions:

1. Can one variable measured at one stage of a respondent's life be equated with a conceptually similar variable (e.g., stress in family of orientation and stress in family of procreation) at a later stage?
2. Is the transition abrupt or gradual? Do respondents undergo a period when involvement in both contexts is important (e.g., working and going to school at the same time), or is this change an abrupt one that involves completely leaving one context and immersing oneself in the other?

# Issues in Longitudinal Research

---

## ■ **Changes in Measurement Over time**

Shifting hypotheses, variables, and measurements part way would mean that the two parts of the research, before and after the shift, might not be comparable.

This potentially destroys the utility of the data, both before and after the shift, for the sort of longitudinal analysis that was originally intended. In addition, it runs the risk of having ones research dictated by what may later be recognized as a transient theoretical fashion.

# Issues in Longitudinal Research

---

## ■ Panel Attrition

respondents who are lost may differ from those who are retained in some systematic way (they may have had different average values on variables to begin with, or they may have changed in ways different from the rest of the sample)

# Issues in Longitudinal Research

---

## ■ Panel Attrition

This especially serious if losses come disproportionately from those with extreme values on the variables on which the research focuses.

Not only the magnitude of the attrition but also the pattern of attrition with respect to critical variables in the study that may be problematic.

# Issues in Longitudinal Research

---

## ■ Reducing Panel Attrition:

- obtaining, on the first and subsequent interviews names and addresses of parents, other relatives, friends, or other individuals with whom the respondent is likely to stay in touch
- repeated annual mailings (e.g., birthday cards or other special occasion greeting cards) with a request to the post office to supply a forwarding address if the respondent moved

# Issues in Longitudinal Research

---

## ■ Panel Attrition

To maintain low attrition rates regardless of the resources available, substantial effort and persistence must be devoted to tracking respondents.

The binomial test (Bulmer, 1979) may be used to test whether the proportion of individuals in different demographic categories (male or female, white or nonwhite, etc.) changes significantly over time.

# Issues in Longitudinal Research

---

## ■ Panel Attrition

Retrospective studies may have problems of selection. It may miss individuals who, for example, died or moved out of the area from which the sample was drawn and who did so during the period for which the data were collected. Those individuals may differ systematically from the rest of the population.

# Issues in Longitudinal Research

---

## ■ Treatment of Missing Data in Longitudinal Research

It is possible to have missing data for a single item on a multiple-item scale, for an entire scale, for a particular case at a particular wave but not for all waves (were nonresponsive), or for a particular case for all waves, for example, as a result of initial nonresponsive.

In this case it would be relatively simple to substitute a scale mean for the missing item.

# Issues in Longitudinal Research

---

## ■ Treatment of Missing Data in Longitudinal Research

For wave nonresponse or initial nonresponse, if there is a clear pattern of “missingness” by a variable that has been measured (e.g., ethnicity or socioeconomic status), weighing cases so the cases that are most like the missing cases “count” a little more is a frequently used option to reduce potential bias.

One may attempt to adjust the coefficients in a regression or similar model by modeling the nonresponse.

# Issues in Longitudinal Research

---

## ■ Treatment of Missing Data in Longitudinal Research

It appears that weighting may be preferable when nonresponse rates are low, while modeling nonresponse may be preferable when nonresponse rates are high.

To fill in some data by interpolation (using both prior and subsequent waves of data to estimate the data at the missing wave) or extrapolation (using only prior waves or only subsequent waves to estimate the missing data).

# Issues in Longitudinal Research

---

## ■ Treatment of Missing Data in Longitudinal Research

In practical terms, the researcher may be faced with three choices:

1. Be a proficient statistician and use maximum likelihood or multiple imputation models plus selection models for nonresponse;
2. Hire a proficient statistician and use maximum likelihood or multiple imputation methods plus selection models for nonresponse;
3. Use listwise deletion.

# Issues in Longitudinal Research

---

## ■ **Repeated Measurement and Panel Conditioning**

Effects of repeated testing may damage internal validity in experimental and quasi-experimental designs, but the use of control groups may allow the researcher to measure this effect and if it is present, to determine whether there are treatment effects in addition to the changes resulting from repeat measurement.

- Willingness of respondents to answer questions in a way that will evoke a known response.
- Unwillingness to participate in the study.
- Respondents will change as a result of participation in survey.

# Issues in Longitudinal Research

---

## ■ Respondent Recall

Increasing memory decay for longer intervals.

The phenomena of *telescoping* and *reverse telescoping*: reporting an event that happened in one period for more or less recent period than one in which it occurred.

Other studies attempt to use *memory-bounding techniques* (e.g. reference to particular events such as birthdays or holidays) to reduce the impact of telescoping and reverse telescoping.

# Issues in Longitudinal Research

---

## ■ Respondent Recall

Use of introductory material to prepare the respondent for the question; asking more detailed questions.

Providing clear reference periods bounded by easily remembered dates.

Attempting to link events about which the question is being asked with salient events in the respondent's life.

“Life history calendars”

# Issues in Longitudinal Research

---

## ■ **The costs of Longitudinal Research**

If the purpose is to measure historical or developmental change, a longitudinal design is essential, especially to separate age, period, and cohort effects.

If the purpose of the research is to identify or estimate the strength of causal relationships, longitudinal research may again be preferable to cross-sectional research, especially if the true causal and temporal order of changes in variables is unknown.

# Issues in Longitudinal Research

---

## ■ **The costs of Longitudinal Research**

If the measurement of change is not a concern, if causal and temporal order are known, or if there is no concern with causal relationships, then cross-sectional data and analysis may be sufficient.

# Longitudinal Analysis

---

Describing change, it is important to make distinctions between:

1. Qualitative and quantitative change
2. Short- term vs. long-term change
3. Whether a change in values of one or more variables or a change in relationships between two or more variables
4. Whether we are interested in describing, predicting, or explaining change.

# Longitudinal Analysis

---

In causal analysis we also need to make distinctions between:

1. Qualitative and quantitative outcomes.
2. Short-term and long-term analysis.
3. Whether we are interested in the values of a particular outcome or in the relationship of two or more outcomes to each other.

# Longitudinal Analysis

---

In most *experimental and quasi-experimental research*, the concern is with short-term change.

Although *developmental research* sometimes considers short-term patterns of change, it is more often concerned with longer term change.

*Causal analysis of differences* (implicitly, differences between individuals, not differences within individuals over time) to distinguish it from the (causal or other) analysis of change.

# Longitudinal Analysis

---

## ■ Types of Longitudinal causal Models

There are four “pure” types of causal models:

(A)  $X \rightarrow Y$ , the value of the dependent variable is expressed as a function of the independent variable;

(B)  $\Delta X \rightarrow Y$ , where  $\Delta X$  represents a change in  $X$ , and the value of the dependent variable is expressed as a function of the change in the independent variable:

(C)  $X \rightarrow \Delta Y$ , where  $\Delta Y$  represents a change in  $Y$  and the change in the dependent variable is expressed as a function of the value of the independent variable:

(D)  $\Delta X \rightarrow \Delta Y$  the change in the dependent variable is expressed as a function of the change in the independent variable.

# Longitudinal Analysis

---

## ■ **Measuring Change**

Measuring or describing change, at its most basic level, involves a bivariate relationship between two variables, one measuring chronological time or age and another measuring some outcome of interest.

For a continuous measurement scales, two measure of change are commonly considered.

# Longitudinal Analysis

---

## ■ Measuring Change

- One is the difference between the later score and the earlier score on a variable as defined by subtraction:  $X_2 - X_1$ , where the subscripts refer to time periods. This may be called a difference, a change score, a raw change, or a raw gain.
- A second measure that has been used in research on change is the residual gain.

# Longitudinal Analysis

---

## ■ Measuring Change

For ratio scales a third commonly used measure is:

1. The percent change
2. A rate change

Graphical descriptions of change generally involve plotting values of a variable for different periods on a graph on which time is the horizontal axis and the variable is represented on the vertical axis.

# Longitudinal Analysis

---

## ■ Measuring Change

Whenever we attempt to measure change, we need to consider whether apparent differences from one time to another really indicate change or whether they may indicate unreliability of measurement instead.

The dilemma of separating unreliability of measurement from real change may best be addressed by replication and the use of multiple, independent measures of reliability and change.

# Longitudinal Analysis

---

## ■ **Measuring Change**

Its also possible, however, to examine changes in the relationships among various variables over time.

Testing for factorial invariance, whether the same indicators have the same relationship to underlying concepts at one time as at another.

# Longitudinal Analysis

---

## ■ **Deterministic Versus Probabilistic Models**

In the deterministic descriptive model of change, all cases having a specified set of characteristics are expected to change in exactly the same way.

In probabilistic causal models, knowing the values of some relevant, finite set of predictor variables, or whether or how much they change for a specific case permits us to know with (in principle) certainty the value of a predicted or dependent variable, whether it will change, how it will change (increase or decrease), and by now how much it will change for that case.

# Longitudinal Analysis

---

## ■ **Deterministic Versus Probabilistic Models**

Descriptions of change in a variable should include only that variable and time in the mathematical formula;

Explanations of change involve the introduction of other variables into the equation;

Predictions of change may or may not involve predictors other than time.

# Longitudinal Analysis

---

## ■ **Deterministic Versus Probabilistic Models**

In probabilistic descriptive model of change, what we expect is not that all cases will change in the same way, even if they have the same characteristics, but that a certain proportion of cases will change in a certain way.

If we know the values of some relevant, finite set of predictor variables or how they change for the total population, the sample, or a group of sufficient size (males and females, more and less developed countries), we can predict with some accuracy the proportion or percentage of cases with specific outcomes or whether they will change, the proportion or percentage that will change in a certain way, and the average (mean, median, modal) amount by which they will change.

# Longitudinal Analysis

---

## ■ **Deterministic Versus Probabilistic Models**

In social sciences, deterministic models are relatively rare, and when they are used, they are usually used to model quantitative variables.

Probabilistic models, particularly statistical models, are much more common in the social sciences, and they are used extensively for both qualitative and quantitative variables.

# Longitudinal Analysis

## ■ Deterministic Versus Probabilistic Models

TABLE 5.2  
Analytical Models and Data Structures

<i>FEW CASES</i> <i>MANY PERIODS</i>	$(n < 20)$ $(t > 20)$	<i>MANY CASES</i> <i>MANY PERIODS</i>	$(n > 100)$ $(t > 10)$
ARIMA models: covariates, transfer function models, interrupted time series models		Continuous time event history analysis: Cox proportional hazards and parametric hazard models	
Autoregressive (AR) time series models			
Lagged endogenous variable (LEV) models		Multilevel growth curve models	
Multivariate dynamic analysis of categorical data with optimal scaling			

# Longitudinal Analysis

## ■ Deterministic Versus Probabilistic Models

---

*FEW CASES*      ( $n < 20$ )  
*FEW PERIODS*    ( $t < 10$ )

*MANY CASES*    ( $n > 100$ )  
*FEW PERIODS*    ( $t < 10$ )

---

Pooled cross-sectional/time-series analysis

Linear panel analysis conditional change model (lagged endogenous variable)

Linear panel analysis unconditional change model (change score)

Latent growth curve analysis

Discrete time event history analysis

Multilevel growth curve models

---

# Longitudinal Analysis

---

## ■ Pooling Cross-Sectional and Time series Data

Pooled cross-sectional and time series data need to be understood in two distinct ways: as a data structure and as an approach to analyzing longitudinal data.

Periods are “Stacked” as though they have part of the same cross-section.

The repeated cross-sectional data structure is useful for analyzing aggregate historical change, at the level of the unit (country, city, etc.) from which cases were sampled and can be analyzed with familiar techniques.

# Longitudinal Analysis

---

## ■ Time Series Analysis

Technically, whenever we have data on the same variable for the same case for two or more periods, we have a time series, but the term “time series analysis” is usually reserved to describe a family of methods for analyzing relatively long time series for a single case at a time.

# Longitudinal Analysis

---

## ■ Time Series Analysis

Long series of time ordered data in terms of some combination of four processes is attempted to be described by ARIMA.

A white noise process is a series of random shocks or changes; this is the probabilistic component that is present in all stochastic time series models.

An autoregressive (AR) process is one in which the present values of a variable depend on past values of that same variable at some specified lag(s) or interval(s).

# Longitudinal Analysis

---

## ■ Time Series Analysis

A moving average (MA) process is one in which past values of the white noise process continue to influence present values of the modeled variable for some finite, specified lag(s) or interval(s).

An integrated (I) process is one in which there is a detectable trend or drift over time in the values of the modeled variable but in which there is no trend or drift in the series that results from subtracting values of the variable from values of the variable at some later time.

# Longitudinal Analysis

---

## ■ **Methods for Short Time series With Many Cases**

For quantitative data, but increasingly expanding to include qualitative data as well, structural equation modeling (SEM) techniques are used for the casual analysis of differences and change in linear panel analysis and for the description and explanation of change in latent growth curve models.