Longitudinal Studies

A. Alexander Beaujean, Ph.D.

Baylor University

Longitudinal studies provide information about phenomena constancy and change. Their core feature is that the phenomena under investigation are observed repeatedly in the same participants (e.g., persons, animals, organizations), which requires collecting data at multiple time points (called *waves*). The participants, phenomena observed, duration of study, and number of waves can greatly differ between studies. For example, longitudinal studies with animals often gather data across many waves, but the entire study may only last a few days or weeks. By contrast, longitudinal studies with humans typically gather data at just a few waves but may last years or decades.

Longitudinal studies cover a wide variety of designs. They can be prospective (e.g., studying pre-term and full-term children for 10 years after birth) or retrospective (e.g., finding a group of elderly adults whose cognitive ability was documented when they were children and then collecting information about various life events). Data can be collected on all phenomena from all participants at all waves, or the design can include *planned missingness* (i.e., participants differ in the phenomena collected or waves in which they provide information). Such designs can both decrease the resources needed to conduct the study as well as lessen the time burden on participants, which could decrease the likelihood of attrition or unplanned missing data.

Traditional longitudinal studies collect data from a single *cohort*, which is a group of participants that share a defining characteristic (i.e., time of birth). The major drawback of this design is that they confound age effects (i.e., influences due to developmental change) with both cohort effects (i.e., influences due to shared historical experiences) or time-of-data-collection effects (i.e., events that co-occur with the data collection occasions). To provide stronger evidence about developmental change, some developmental researchers collect data from multiple cohorts, in which case they are called sequential studies.

Data from longitudinal studies are dependent because the information is collected from the same participants at multiple waves. This dependency violates an assumption of many traditional statistical techniques. Consequently, analysis of longitudinal data requires techniques that allow for data dependency. The simplest technique involves calculating and analyzing *difference scores,* which is the difference in a variable's values between two waves—typically the first and last. Analyses of difference scores are often inadequate because they cannot determine if the observed change is due to: (a) the participants changing in the target phenomena (sometimes called α change); (b) both the participants changing in the target phenomena as well as changes in the instruments capturing the phenomena (β change); or (c) the target phenomena changing (γ change). Say researchers designed a longitudinal study to examine students' perceptions of school safety with information gathered at two wave: the beginning and end of the school year. If a traumatic event occurred in the middle of the year (e.g., tornado, shooting), it would be difficult to know if any differences found in the school safety variable values occurred because of change in students' actual perceptions or because their very understanding of safety changed after the traumatic event.

More robust analysis of longitudinal studies involves collecting data from more than two waves and using analysis techniques that address the substantive issues surrounding investigating change in the particular target phenomena. Table 1 provides some common techniques used in developmental research.

Further Reading

Baltes, P. B., & Nesselroade, J. R. (1979). History and rationale of longitudinal research. In J. R. Nesselroade & P. B. Baltes (Eds.), *Longitudinal research in the study of behavior and development* (pp. 1--39). New York, NY: Academic Press.

Card, N. A., & Little, T. D. (2007). Longitudinal modeling of developmental processes. *International Journal of Behavioral Development, 31*, 297-302.

Chan, D. (1998). The conceptualization and analysis of change over time: An integrative approach incorporating longitudinal mean and covariance structures analysis (LMACS) and multiple indicator latent growth modeling (MLGM). *Organizational Research Methods, 1*, 421-483.

Graham, J. W., Taylor, B. J., & Cumsille, P. E. (2001). Planned missing data designs in the analysis of change. In L. M. Collins & A.G. Sayer (Eds.), *New methods for the analysis of change* (pp. 335 353). Washington, DC: American Psychological Association.

Table 1. *Common Data Analysis Techniques for Longitudinal Studies.*

Technique	Description
Latent curve models	Structural equation model parameterized to measure intra- and inter-participant change.
Latent growth mixture models	Latent curve model extension used to identify population subgroups that have distinctive change trajectories.
Latent state-trait models	Latent curve model extension used to decompose outcome covariance into components representing trait variance across time points, state variance at specific time points, and error variance.
Latent transition analysis	Latent curve model extension for categorical data that examinees changes in class/stage membership over time
Multilevel models	Regression model extension for hierarchically structured data that accounts for observations at a given wave being nested within participants. Can be equivalent to latent curve models under certain parameterizations.
Event history/survival models	Useful for analyzing data on the occurrence and timing of discrete outcomes (e.g., school dropout, marriage).