Intensive Longitudinal Data: Methodological Challenges and Opportunities

Daniel J. Bauer & Patrick J. Curran

CenterStat

APA Science Training Sessions:
The Collection and Analysis of Intensive Longitudinal Data

centerstat.org

Intensive Longitudinal Data

Part 1: Methodological Challenges & Opportunities

Daniel J. Bauer & Patrick J. Curran
University of North Carolina at Chapel Hill

AMERICAN PSYCHOLOGICAL ASSOCIATION

CenterStat

Objectives

- Define "longitudinal data" and consider why it is so important
- Review different types of longitudinal designs
- Note limitations of panel data and focus on intensive longitudinal data
- Describe unique features of intensive longitudinal data
- Highlight ability to disaggregate within- and between-person effects
- Briefly note measurement challenges
- Describe where we will go next

1.2

What is Traditionally Considered "Longitudinal"?

- There is a surprising degree of disagreement on what exactly constitutes longitudinal design, data, and analysis
  - much controversy is historical leftover from back-in-the-day
- Baltes & Nesselroade (1979):
  - "*The one sine qua non of longitudinal research is that the entity under investigation is observed repeatedly as it exists and evolves over time*"
- We adopt this minimalist definition with one caveat:
  - at least a *subset* of entities under investigation are observed repeatedly
  - allows for missing data and single observation data

1.3

Different Types of Longitudinal Designs

- **Pre-post design**: Classic two time point design where an assessment is obtained before and after some treatment or manipulation
  - analysis focused on raw (e.g., ANOVA) or residualized change (e.g., ANCOVA)
  - David Rogosa: two time points are better than one, but not much better
- **Panel design**: often seen as "traditional" design where two or more assessments obtained at fixed and widely-spaced time points
  - e.g., three or more assessments taken at 6 or 12 or 24 month periods on a large sample of individuals
  - examples include National Longitudinal Survey of Youth, Early Childhood Longitudinal Study, Longitudinal Study of American Youth, Add Health Study, etc.

1.4

Different Types of Longitudinal Designs

- **Time series design**: often a high number of closely-spaced assessments taken on one or a small number of individuals
  - e.g., stock market data, wearable device measures, etc.
- **Intensive longitudinal design**: cross between time series and panel designs
  - larger number of assessments than panel designs, but fewer than time series
  - larger number of individuals than time series, but fewer than panel designs
- **Burst designs**: clever hybrid of a more traditional panel design that includes "bursts" of intensive measures between panel assessments
  - e.g., annual assessments on core constructs, but an ILD burst in interim time period
- Regardless of design, important to first consider *why* we want longitudinal data

1.5

Why Do We Even Want Longitudinal Data?

- Both hard and expensive to obtain, so why want in the first place?
  - many reasons, some of which are less obvious
- Most commonly stated is to establish *temporal precedence*
- John Stuart Mill conditions to infer causation
  1. the cause must precede the effect
  2. the cause must be related to the effect
  3. no plausible alternative explanations for the effect other than the cause
- Many other philosophers of causation, but nearly all require that a cause must temporally precede an effect

1.6

Temporal Precedence

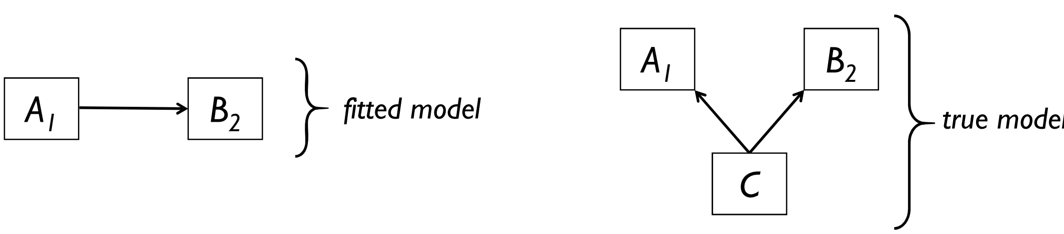
- Longitudinal designs can establish temporal precedence, but this *necessary* condition is rarely *sufficient* to infer cause
- Say we hypothesize that construct *A* at time 1 causes construct *B* at time 2

- Observing *A* earlier in time than *B* unambiguously establishes temporal precedence, so what could possibly go wrong??
- There remain **many** alternative models that account for the observed data

1.7

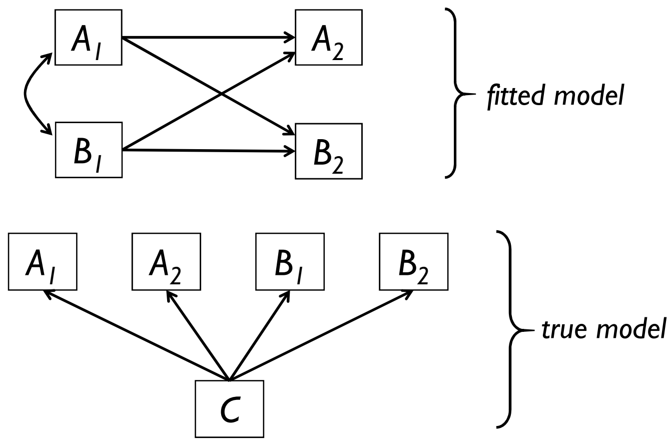
Alternative Models

- For example, *C* might be the true underlying cause of both *A* and *B*, but because *C* is omitted, *A* appears to be causally related to *B* in the absence of *C*



1.8

And As We Approach Halloween: Even Scarier



- Temporal precedence is critical, but must rule out other explanations to infer cause

1.9

Isolating Intra-Individual Variability & Change

- Another key advantage of longitudinal data is that we can simultaneously consider and model two components of variability:
  1. pattern of variability and change ***within*** each individual
  2. differences in these patterns ***across*** individuals
- With longitudinal data, can explicitly examine *between*-person differences in *within*-person **change**
  - ***inter-individual differences*** in ***intra-individual change***
- Can also consider *within*-person **processes**
  - within individuals, variability in x over time predicts variability within y over time
  - can examine ***inter-individual differences*** in ***intra-individual processes***
- This will be our focus for much of the week

1.10

Traditional Panel Data

- Traditional longitudinal data is typically what is referred to as "panel data"
  - Typically 3 to 7 repeated measures obtained on sample sizes over 100
  - Repeated measures often spaced far apart (e.g., by one or more years)
- When analyzing panel data, a typical goal to describe and predict individual differences in within-person **change** over time via growth models
  - Can be fit in either structural equation modeling framework (latent curve models) or multilevel modeling framework
- But traditional panel data is often less well suited to study of within-person **processes**

1.11

Key Limitation of Panel Data

- Why is panel data poorly suited to study of within-person processes?
- Consider *negative reinforcement hypothesis*: individuals consume alcohol to mitigate negative feelings (i.e., self-medication)
- Traditional panel data might use negative affect at age 15 to predict alcohol use at age 16
  - but a *between-person* relation: people with more negative affect tend to drink more
- Theory predicts more proximal effects, i.e., that negative affect *during the day* predicts alcohol use *that night*
  - this is a *within-person* relation: on average, when a person experiences more negative affect *than they usually do,* they consume more alcohol *than they usually drink*
- Need higher numbers of more closely spaced repeated assessments

1.12

Traditional Time Series Data

- A classic approach characterized by a high number of closely spaced repeated assessments is *time series analyses*
- Number of time points often > 50 assessed on <u>one</u> or <u>a few</u> units
- Repeated measures often taken close together in time (e.g., day-by-day)
- Analytic goal is typically prediction (e.g., forecasting) or inferring causation (e.g., interrupted time series)
- Powerful methodology, but not always optimal for behavioral sciences

1.13

Key Limitation of Time Series Data

- A core goal in the behavioral sciences is to generalize findings from a sample of individuals to a broader population
  - **external validity**: the generalization of findings across person, place, and time
- Time series data provides intense understanding of an individual unit
  - e.g., *idiographic:* the study of particular facts rather than general laws
- But time series data often extremely limited for drawing generalizations
  - e.g., *nomothetic*: the study general laws and theory
- If the scientific goal is focused on one or small number of individuals, time series may be ideal
- But if the scientific goal is to draw general conclusions based on a sample of individuals, time series can be quite limited

1.14

A Middle Ground: Intensive Longitudinal Data

- ILD is a cross between panel and time series
- Number of time points ≥ 10 and sample sizes ≥ 25
- Repeated measures taken close together in time (hourly, daily)
- Motivating analytic goals are
  1. To capture within-person **processes** (e.g., negative reinforcement motive for drinking) and individual differences in these processes
  2. Understand individual differences in within-person **change** over time over a specific period (e.g., changes in cognition during treatment for cancer)
- Given density of observations, ILD is exceptionally well-suited to test a variety of research hypotheses

1.15

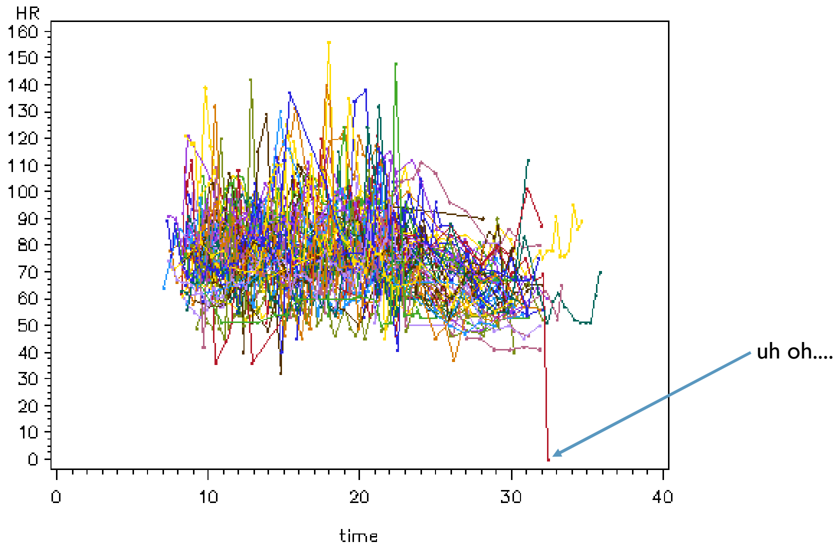
But Be Careful For What You Ask…

- Arguably the biggest challenge in ILD is *data management*
  - data to right: n=61 individuals assessed six to 60 times over 24 hours on heart rate
- Both MLM and DSEM structure data in "long" format
  - e.g., one line of data per person, per assessment
- For typical applications, this data file might be thousands of lines long
- Graphics are an excellent starting point

Obs	ID	time	HR	posture	awake	discrim	MALE	AGE
1	1.00	8.4833	79	sitting	1	1.9	0	27
2	1.00	8.7333	99	standing	1	1.9	0	27
3	1.00	9.1333	89	standing	1	1.9	0	27
4	1.00	9.4000	90	standing	1	1.9	0	27
5	1.00	9.8000	86	standing	1	1.9	0	27
6	1.00	10.0500	86	standing	1	1.9	0	27
7	1.00	10.3500	83	standing	1	1.9	0	27
8	1.00	10.6000	80	standing	1	1.9	0	27
9	1.00	10.9333	84	standing	1	1.9	0	27
10	1.00	11.3000	80	standing	1	1.9	0	27
11	1.00	11.7000	76	sitting	1	1.9	0	27
12	1.00	12.0667	81	standing	1	1.9	0	27
13	1.00	12.6500	85	standing	1	1.9	0	27
14	1.00	13.0333	90	standing	1	1.9	0	27
15	1.00	13.4000	79	sitting	1	1.9	0	27
16	1.00	13.8167	79	sitting	1	1.9	0	27
17	1.00	14.6000	87	standing	1	1.9	0	27
18	1.00	14.9000	81	sitting	1	1.9	0	27
19	1.00	15.6000	60	lying	1	1.9	0	27
20	1.00	16.0167	57	lying	1	1.9	0	27

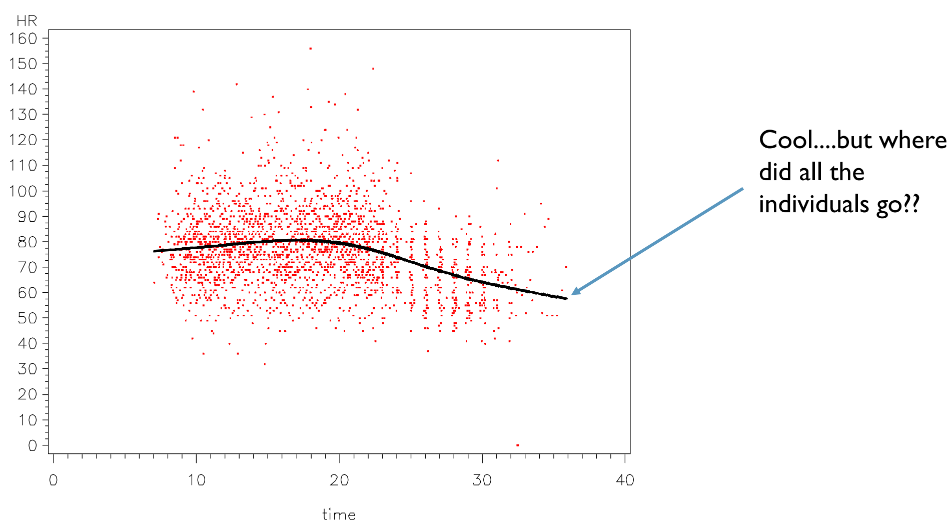
1.16

All 61 Subjects Considered at Once: Yikes



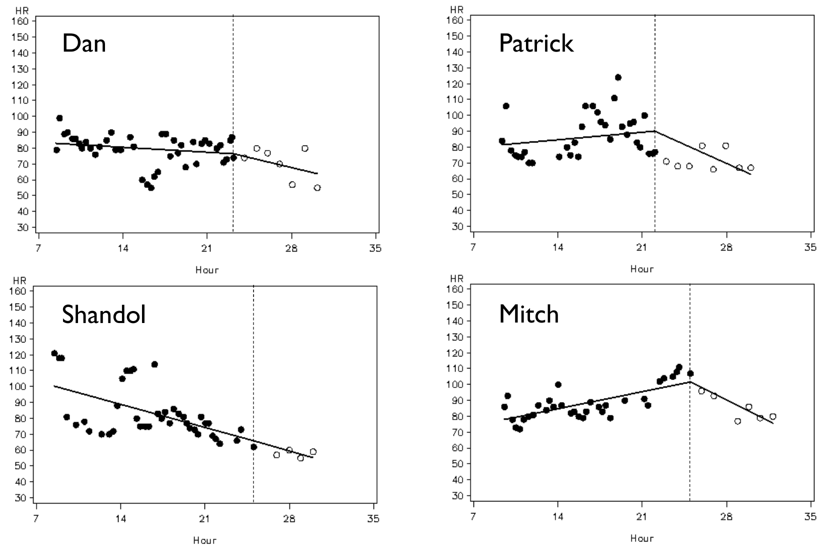
1.17

Aggregate Time Trend Pooling Over Subjects



1.18

Individual Plots For Four Subjects



1.19

Unique Features Posed by ILD

- Now that we have managed the data and examined graphics, we need to think about how to statistically model the data to test our hypotheses
  - But several unique features of ILD pose challenges when using more traditional statistical modeling such as latent curve modeling
- Examples include
  1. potential for *transition points* and *time trends*
  2. the likely occurrence of *serially correlated residuals*
  3. challenges and opportunities for isolating ***intra**-individual* versus ***inter**-individual* effects
- We will explore each of these in turn here in general terms, then revisit in next sessions within specific contexts of MLM (Day 2) and DSEM (Day 3)

1.20

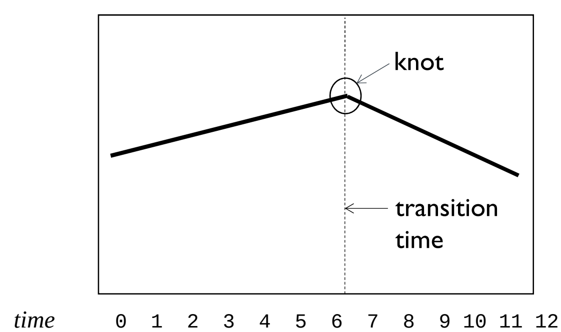
Transition Points

- With intensive longitudinal data, often of interest to observe change before and after an event, similar to interrupted time series analysis
  - Daily measures of stress and social support in couples prior to and after one member takes the bar exam
  - Daily measures of positive affect before and after taking part in a mindfulness meditation intervention
- The individual's trajectory may be 'deflected' by the event, or there may be a discontinuity in the trajectory
  - If change is roughly linear before and after the event, transitions can be modeled using a linear spline or piecewise linear model
  - If change is nonlinear before and/or after event, more complex splines used
  - If discontinuities (e.g., sudden jumps or falls at event), then still more complex

1.21

Modeling Transitions

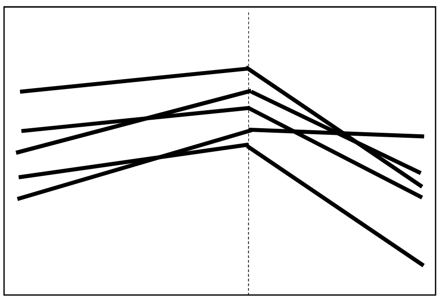
- Piecewise linear models are simple spline models in which there are two or more periods of linear growth.
- Where the lines are joined is sometimes called a knot.



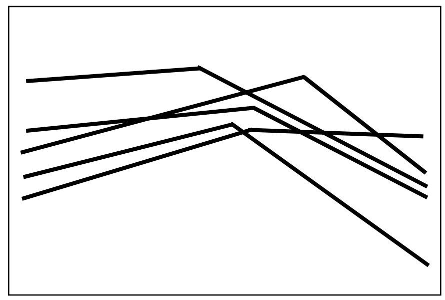
1.22

Constant vs. Varying Transition Point

- Can have the same transition point for everyone
  - e.g., spousal social support before and after bar exam



- Or transition points can vary across cases
  - e.g., momentary assessments of stress at work, stress after leaving work



1.23

Modeling Correlated Repeated Measures: Panel Data

- In addition to any overall time trends, also need to account for over-time correlations among repeated measures (dependence of observations)
- With long-term panel data, preference is to model over-time correlations among repeated measures with *random effects*
  - Random effects represent individual differences in level and rate of change over time, and these imply correlations among repeated measures
- Once we account for these random effects, residuals of repeated measures are assumed to be independent (within and across persons)
  - e.g., residuals for the same person at different points in time are unrelated
  - makes sense if perturbations from an individual's trajectory are short lived, vanishing over the long time intervals between assessments

1.24

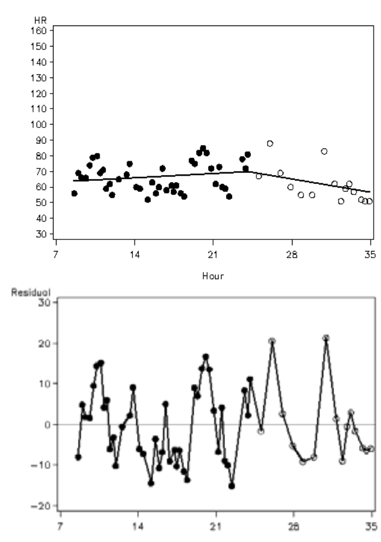
Modeling Correlated Repeated Measures: ILD

- With intensive longitudinal data, assumption that the residuals are independent within persons is often unrealistic given short intervals between measurements
  - Correlations among repeated measures often reflect *serially correlated residuals* in addition to random effects
- Short-lived effects of momentary perturbations ("shocks") remain in evidence at the next time point(s) given brief time intervals
  - Example: HR may elevate immediately with a shot of espresso or upon being yelled at by the boss and not return completely to normal levels until after several HR readings are taken
- Can borrow from time-series analysis to incorporate serial correlation structure for residuals (e.g., AR, MA, ARMA, etc.)

1.25

Example of Serially Correlated Residuals

- If there is minimal measurement error, can directly observe serially correlated residuals in individual time plots
- The apparent cycling around the line in HR values for this participant is a tell-tale sign
  - When one residual is positive, the next is more likely to also be positive
  - Must account for this serial correlation in our model specification



1.26

Within- and Between-Person Effects

- A common goal in collecting ILD is to be able to identify within-person processes
  - In part to evaluate the negative-reinforcement hypothesis, Hussong et al. (2001) used daily reports over three weeks to examine whether within-person increases in negative affect predicted within-person increases in alcohol use
- Evaluating within-person effects may be the primary goal of the study
  - Unlike long-term longitudinal studies, which tend to focus on between-person differences in time trends
- Indeed, the ability to disaggregate within-person and between-person effects is one of the key advantages of ILD

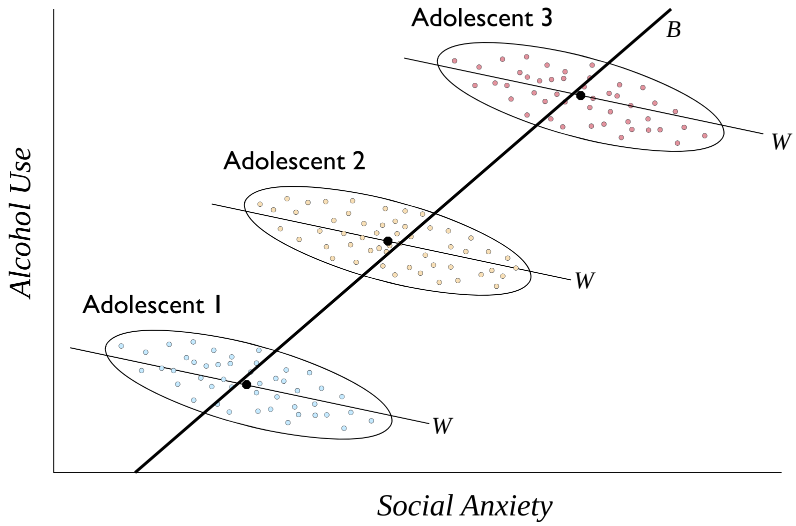
1.27

Within- and Between-Person Effects

- Historically, psychologists have often used *between-person* individual difference data to make inferences about *within-person* causal processes
- Carries the potential for errors of inference (akin to *Ecological Fallacy*)
  - Example: People are more likely to have a heart attack while exercising (within-person effect), but people who exercise more often on average are less likely to have a heart attack (between-person effect)
- With ILD, have the potential to separate within- and between-person effects of Time Varying Covariates (TVCs)
  - Large number of observations per person enables more precise estimation of within-person effects than is typically possible in long-term longitudinal studies
  - Smaller time intervals between observations allow us to capture these processes as they unfold

1.28

Opposite Within- and Between-Person Effects



Between: adolescents with higher average anxiety have higher average drinking (maladaptive coping via self-medication?)

Within: On days that an adolescent is more socially anxious *than usual*, they tend to drink less *than usual* (elevated anxiety keeps them out of social settings that provide alcohol?)

1.29

ILD and Decomposition of Effects

- Of all advantages of ILD, decomposition of effects may be most important
- Frequent assessments close together in time can be used to obtain reliable estimates of within-person effects
  - fear of panic attacks in the morning predicting the actual number of panic attacks experienced that day
- High-density of observations also allows for reliable estimates of traditional between-person differences
  - Can still incorporate person-level predictors of importance, e.g., treatment condition, racial group membership, self-identified gender
- ILD thus offers a powerful foundation for theory testing
  - but in the spirit of be careful for what you ask, now we need to *analyze* the data

1.30

Next Steps: Fitting Models to ILD

- There are many approaches to the analysis of ILD, but we focus on two over the next two sessions:

1. The multilevel model, or MLM (October 6)
   - MLM treats repeated assessments as nested within individual and is able to nicely parse within-person and between-person effects
   - well suited to studying both within-person change over time (like growth model) and within-person processes
2. The dynamic structural equation model, or DSEM (October 11)
   - DSEM combines elements of time series modeling, multilevel modeling, and SEM to leverage strengths of each for analyzing ILD
   - especially well suited to studying stable within-person processes

- So you're stuck with us for a bit

1.31

The Importance of Measurement

- Of course, these models are only as good as the measurements that we put into them
- A critical aspect of collecting ILD is to ensure the measures reflect valid and reliable measures of our theoretical constructs of interest
  - Can't assume traditional scales developed with strictly between-person data are also valid and reliable for measuring within-person variation
  - Often must also reduce items: If ping a subject five times a day on a cell phone, can't realistically give a multi-item scale at every assessment
- Lots of interesting work ongoing in this area
  - Lots more still to be done (Anyone need a dissertation project? Anyone?)
  - Not enough time for us to explore in detail, but see Laurenceau & Bolger, 2013, *Intensive Longitudinal Methods*, for a nice orientation to this topic

1.32

Summary

- A critical component of any scientific inquiry is to validly infer *cause*, and a critical component of causal inference is to establish *temporal precedence*
- There are many ways to design studies of stability and change over time
  - two time point, panel, time series, burst, etc.
- When focus is on within-person change / processes, ILD is ideal
  - balances moment-to-moment "life as lived" with basis for drawing generalizations
- ILD comes with unique challenges, but also offers unique opportunities
  - On Day 2 (Oct 6) we explore fitting models to ILD using the *multilevel model*
  - On Day 3 (Oct 11) we explicate the *dynamic structural equation model* for ILD
- We hope to see you then!

1.33

A Semi-Random Sampling of Resources

Asparouhov, T., & Muthén, B. (2020). Comparison of models for the analysis of intensive longitudinal data. *Structural Equation Modeling: A Multidisciplinary Journal*, *27*, 275-297.

Bolger, N., & Laurenceau, J. P. (2013). *Intensive longitudinal methods: An introduction to diary and experience sampling research*. Guilford Press.

Bolger, N., Stadler, G., & Laurenceau, J.-P. (2012). Power analysis for intensive longitudinal studies. In M. R. Mehl & T. S. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 285–301). Guilford Press.

Collins, L. M. (2006). Analysis of longitudinal data. *Annual Review of Psychology*, *57*, 505-528.

Curran, P.J., & Bauer, D.J. (2011). The disaggregation of within-person and between-person effects in longitudinal models of change. *Annual Review of Psychology, 62*, 583-619.

Hamaker, E. L., & Wichers, M. (2017). No time like the present: Discovering the hidden dynamics in intensive longitudinal data. *Current Directions in Psychological Science*, *26*, 10-15.

Ji, L., Chow, S. M., Schermerhorn, A. C., Jacobson, N. C., & Cummings, E. M. (2018). Handling missing data in the modeling of intensive longitudinal data. *Structural Equation Modeling: A Multidisciplinary Journal*, *25*, 715-736.

Laurenceau, J. P., & Bolger, N. (2012). Analyzing diary and intensive longitudinal data from dyads. *Handbook of research methods for studying daily life*, 407-422.

McNeish, D., Mackinnon, D. P., Marsch, L. A., & Poldrack, R. A. (2021). Measurement in intensive longitudinal data. *Structural Equation Modeling: A Multidisciplinary Journal*, *28*, 807-822.

Schwartz, J. E., & Stone, A. A. (1998). Strategies for analyzing ecological momentary assessment data. *Health Psychology, 17*, 6 –16.

Walls, T. A. (2013). Intensive longitudinal data. *The Oxford handbook of quantitative methods: Statistical analysis*, *2*, 432-440.

Walls, T. A., & Schafer, J. L. (Eds.). (2006). *Models for intensive longitudinal data*. Oxford University Press.

1.34