Inferences With Longitudinal Data

. . .[1]

[1] ...

Author Note

. . . .

Correspondence concerning this article should be addressed to . . . , ... E-mail: . . .

Abstract

Begin here...

Keywords: ....

Word count: 95

Inferences With Longitudinal Data

There is now a common understanding that the phenomena organizational researchers study unfold over time. Beal (2015), for example, states that psychological phenomena are “sequences of events and event reactions that play out within each person’s stream of experience,” and “describing these within-person processes is fundamental to understanding” (p. 5). Similarly, Pitariu and Ployhart (2010) note that the processes we study “are not static but instead develop, change, and evolve over time” (p. 405). Moreover, organizations are systems with many connected parts, and systems are inherently dynamic. Studying these systems and processes, therefore, requires paying attention to how things happen over time; doing so puts us in a better position to capture the sequence, understand it, and can lead to new and interesting insights (Kozlowski & Bell, 2003).

This sentiment is reflected in our empirical literature, where repeated assessments are now common. For instance, Jones et al. (2016) observed the work attitudes of pregnant women in their second trimester every week until they gave birth. Meier and Spector (2013) examined counterproductive work behavior over five waves. Hardy, Day, and Steele (2018) investigated self-regulation over 20 lab trials. Finally, Johnson, Lanaj, and Barnes (2014) observed justice behavior and resource depletion across 10 consecutive workdays.

Armed with repeated observations, there are then different inferences and research questions that we can explore. Jones et al. (2016) ask about trend: they propose negative slopes for concealing behaviors and positive slopes for revealing behaviors among pregnant women. Johnson et al. (2014) about change: they argue that changes in justice behaviors will predict changes in depletion. Hardy et al. (2018) inquire about dynamic relationships, where prior self-efficacy relates to subsequent metacognition, and metacognition then influences self-efficacy at a later time point. Finally, Meier and Spector (2013) examine how effect sizes change when they vary the time lag between their independent and dependent variable.

None of these inferences perfectly discovers the data generating mechanism. Rather, each asks an interesting and important question about how DVs relate to IVs. Only with lots of asking about lots of different patterns of relationships across the variables could we piece together one (of many) possible representation(s) of the data generating process – hopefully having a good theory to guide the way.

The spine of an investigation is then the statistical model that researchers apply and ultimately connect to their inference or question. Meier and Spector (2013) present a sequence of path models that test increasingly longer time lags. Hardy et al. (2018) and Jones et al. (2016) employ bivariate cross-lagged latent growth curves, an approach similar to the latent change model used by (**???**). We also find complex hierarchical linear models in many event-sampling studies (e.g., **???**; Rosen, Koopman, Gabriel, & Johnson, 2016). All of these researchers evoke a particular model to support or reject a particular inference.

We want to link inferences to models in this paper so that researchers know which of the many models they can use when they are interested in one of the many possible inferences in a longitudinal investigation. As should be clear to anyone reading our literature, there is great excitement for the utility of longitudinal studies; they can pose interesting questions and discover patterns that would otherwise be impossible to capture in a static investigation. One of our goals is to bring attention to the span of questions available so that researchers can fully appreciate and take advantage of their data. Second, although the inferences concern trajectories or relationships over time, their small differences have large implications for what we take away from them – what we ultimately conclude. Here, we unpack the differences fully to ensure this point is emphasized. Third, there are many inferences, many models, and different models can be used to understand or explore the same inference. In this paper, we provide readers with a specific model for each inference so that they can be sure that the model they evoke is appropriate for the research question that they are interested in. In summary, this paper exposes researchers to the span of inferences they may investigate when they collect longitudinal data, links those inferences to models, and parses some of the modeling literature that may be difficult to consume for researchers with only graduate level training in statistics.

Below, we do these things.

A similar hook but with a separate parts for research questions and inferences. Paragraphs 3, 4, and 5 are different.

There is now a common understanding that the phenomena organizational researchers study unfold over time…

This sentiment is reflected in our empirical literature, where repeated assessments are now common…

Armed with repeated observations, there are then different research questions that we can explore. Jones et al. (2016) ask about trend: they want to determine if the trajectories among certain variables increase or decrease over time. Johnson et al. (2014) about change: they are interested in how changes in one variable relate to changes in another across time. Hardy et al. (2018) inquire about dynamic relationships, where prior values on one variable predict subsequent values on another, and this second variable then goes back to predict the first at a later point in time. Finally, Meier and Spector (2013) examine how effect sizes change when they vary the time lag between their independent and dependent variable.

Researchers then evoke statistical models that are determined by their research questions. Meier and Spector (2013) present a sequence of path models that test increasingly longer time lags. Hardy et al. (2018) and Jones et al. (2016) employ bivariate cross-lagged latent growth curves, an approach similar to the latent change model used by (**???**). We also find complex hierarchical linear models in many event-sampling studies (e.g., **???**; Rosen et al., 2016).

The spine of an investigation, finally, is to interpret the model to make an inference regarding the original question. Jones et al. (2016) infer negative slopes for concealing behaviors and positive slopes for revealing behaviors. Johnson et al. (2014) argue that "justice is dynamic: The frequency of actors' justice behaviors varies day to day" (p. 10), and these daily fluctuations predict daily changes in depletion. Hardy et al. (2018) find support for dynamic relationships between self-efficacy, metacognition, and exploratory behaviors. Finally, Meier and Spector (2013) suggest that the effects of work stressors on counterproductive work behaviors are not substantially different across different time lags.

None of these inferences perfectly discovers the data generating mechanism. Rather, each asks an interesting and important question about how DVs relate to IVs. Only with lots of asking about lots of different patterns of relationships across the variables could we piece together one (of many) possible representation(s) of the data generating process – hopefully having a good theory to guide the way.

We want to link inferences to models in this paper so that researchers…

Longitudinal Definitions

This paper is exclusively devoted to the inferences we make with repeated observations, so we begin by identifying a few labels and definitions. Authors typically identify a "longitudinal" study by making a contrast with respect to either a) research designs or b) data structures. Longitudinal *research* is different from cross-sectional research because longitudinal designs entail three or more repeated observations (Ployhart & Bliese, Singer & Willett). We therefore emphasize differences on the number of observations when we distinguish longitudinal from other types of research. Longitudinal *data* are repeated observations on several units (i.e., N or $i > 1$), whereas panel data are observations of one unit over time – a distinction that focuses on the amount of people in our study (given repeated measures). Most organizational studies collect data on more than one unit, therefore our discussion below focuses on longitudinal research with longitudinal data, or designs with $N > 1$, $t >= 3$, and the same construct(s) measured on (potentially) each i at (potentially) each t.

Framework

Relationships. Growth. Change. Dynamics. These are umbrella research foci, each has its own sub-inferences and models.

Relationships

General discussion.

Inference 1

A stable x relates to y.

Model.

.

```
'
perf.1 ~ b1*gender
perf.2 ~ b1*gender
perf.3 ~ b1*gender
perf.4 ~ b1*gender


'
```

Inference 2

A fluctuating x relates to y.

Model. .

```
'
perf.1 ~ b1*affect.1
perf.2 ~ b1*affect.2
perf.3 ~ b1*affect.3
perf.4 ~ b1*affect.4

'
```

Growth

General discussion.

Inference 1

There is growth (positive or negative) in a given variable. Other terms: trend, slope, some call this change; we won't.

Model. .

```
'
latent_perf_slope =~ 0*perf.1 + 1*perf.2 + 2*perf.3 + 3*perf.4

'
```

Inference 2

There is a relationship between growth (slope) and level in a given variable.

Model. .

```
'
latent_perf_level ~~ latent_perf_slope


'
```

Inference 3

There is a relationship between a stable x and growth in y.

Model. .

```
'
latent_perf_slope ~ b1*gender


'
```

Could also do this for level.

Inference 4

There is a relationship between a fluctuating x and y after partialling the growth in y; or there is growth in y after partialling the relationship between a fluctuating x and y.

Model. .

```
'
latent_per_slope =~ 0*perf.1 + 1*perf.2 + 2*perf.3 + 3*perf.4


perf.1 ~ b1*affect.1
perf.2 ~ b1*affect.2
perf.3 ~ b1*affect.3
perf.4 ~ b1*affect.4


'
```

Growth 2.0

Above, we examined growth in y and how it related to correlates or predictors – but those predictors/correlates were assumed to have no growth. There is also a class of models for examining relationships between two variables where both are assumed to grow.

Inference 1

There are correlated slopes among two growth curves.

Model. .

```
'


latent_perf_slope ~~ latent_affect_slope


'
```

Change

Partial technique in MLM

Dual change in SEM

Relationships

Sub Inference 1.

.

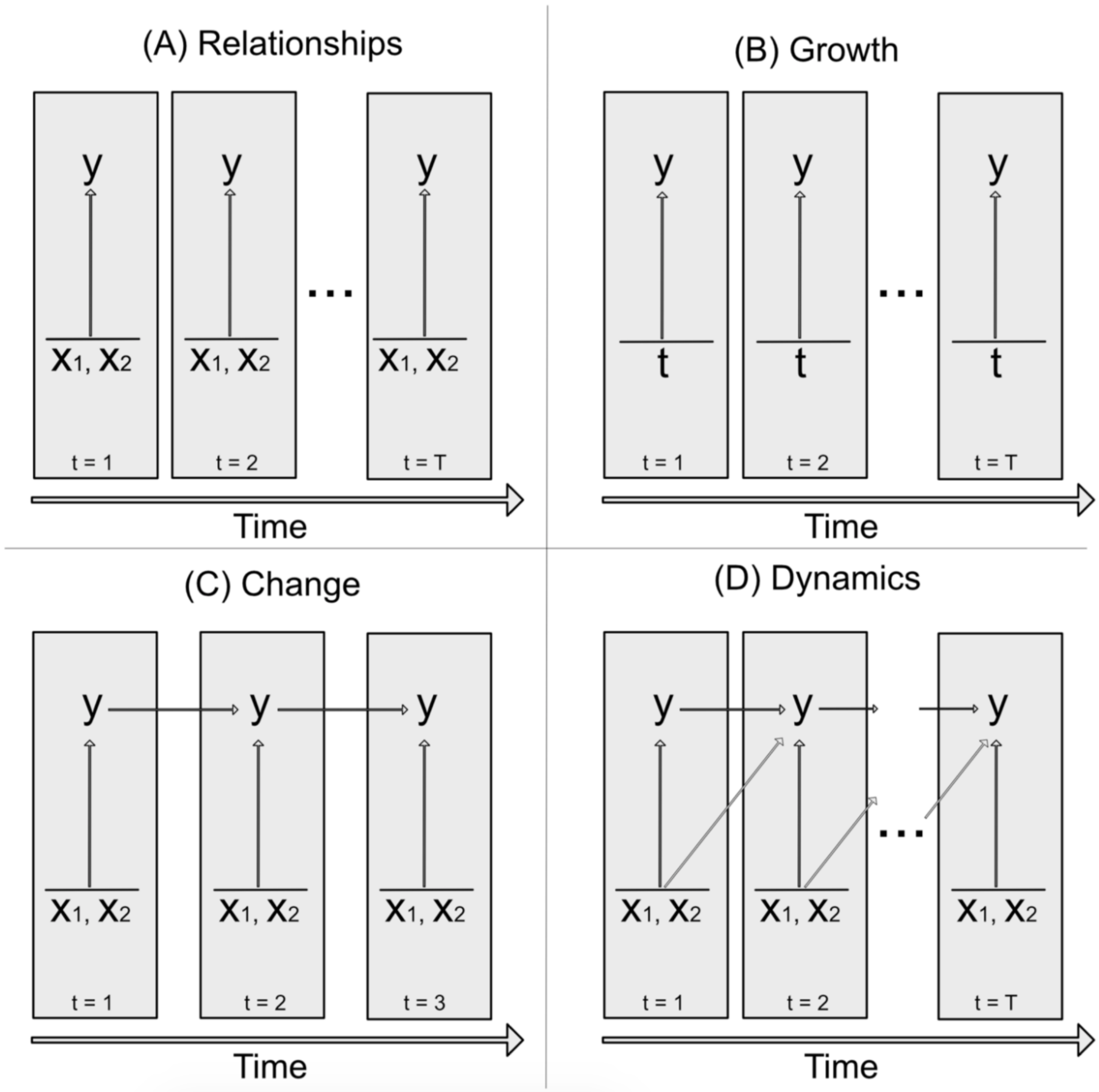
Sub Model 1.

Time-invarying covariates analysis.

Introduce framework

Presenting the entire inference and modeling literature that uses longitudinal data would be impossible. Instead, we focus on four related streams that we feel can be organized using a framework proposed by Xu and DeShon. Figure one shows each inference we will discuss in this paper: relationships, growth, change, and dynamics.

In each panel in figure one, time is on the x-axis to portray that we are investigating these inferences over time. Each slice contains an observation of y, such that at time t we observe y_t and at $t+1$ we observe y_{t+1}. What differentiates the panels – the inferences – is the pattern of relationships we investigate – and we add complexity as we move from A to D. For example, researchers do not include lag effects when they are interested in relationships over time (panel A), but they do include a lag effect when they study change or dynamics (panels C and D). We will devote a section to each of these inferences below, but we first describe some preliminary pieces about code and data that we refer to repeatedly in this paper.



Introduce data sets and generic variables

We will use two generic variables – affect (x) and performance (y) – throughout this paper. These variables will hopefully provide continuity across the inferences and also provide an illuminating backdrop after we refer to x or y generically. Moreover, we use code examples in each section that refer to data sets that contain measures of affect and performance on 50 simulated subjects across five time points.

These data are contained in two data sets: `long_df` and `wide_df`. The raw data are the same, but their formatting differs because different estimation techniques require different data structures. Structural equations modeling (SEM) requires wide data, whereas HLM requires long data. The first data set, `long_df` contain the data in long form,

affect	performance	id	time
10.269606	17.23310	1	1
9.370015	22.81949	2	1
10.868660	17.64258	3	1
11.727196	20.56741	4	1
10.024188	12.08548	5	1
10.368025	13.17514	6	1
8.690796	12.76671	7	1
10.738622	13.62823	8	1

where `id` refers to a person identifier, `time` refers to the observation period, and scores on `affect` and `performance` are listed in their respective columns.

The second data set, `wide_df` contain the data in wide form,

id	affect.1	performance.1	affect.2	performance.2	affect.3	performance.3
1	10.269606	17.23310	9.945866	13.653924	9.059233	20.944278
2	9.370015	22.81949	10.461612	10.886965	10.093910	16.578068
3	10.868660	17.64258	9.403230	9.386411	10.066253	10.834513
4	11.727196	20.56741	11.263258	20.491540	9.009344	8.547953
5	10.024188	12.08548	8.854671	16.052982	9.927885	16.759775
6	10.368025	13.17514	11.084624	13.839404	9.776743	26.595386
7	8.690796	12.76671	8.471005	13.229533	9.857223	10.831559
8	10.738622	13.62823	8.426258	16.364850	10.535407	8.794605

where `id` is defined above and now `affect` and `performance` are given new columns at each measurement period – such that, as an example, person one reports 10.3 for affect at time one and 9.9 for affect at time two (the data continue for five time points). Again, we will refer to these variables and data sets as we unpack each section, but remember that the values within each data set are the same – the only difference is their format.

Introduce Code and Its Use

Finally, we will also present code snippets that estimate models related to each inference. Code is typically published in methods literature to demonstrate how to estimate models, but it can also be used as a tool – a language – to build a greater understanding of the phenomenon (Shiffman, 2012). We would like to use it for both reasons. Our goal is to provide readers with code for estimating models, but also allow them to see the inferences represented in various ways – words and code – to help clarify what they represent.

The code snippets will take consistent forms throughout. HLM models will be expressed as follows:

```
hlm_model <- lme(

  code
  more code

  data = long_df

)
```

where `hlm_model` is an object that stores the results of our model, `lme` is a function call to a linear mixed-effects regression (HLM), and within that we specify our effects and reference our data. In the inference sections, `code` and `more code` will be replaced by the effects we estimate, and `data` will always reference `long_df` because HLM requires long formatted data. Finally, after storing our results in the object `hlm_model` we can actually view them with:

```
summary(hlm_model)
```

Similarly, SEM models will be expressed as:

```
sem_string <- '

    code
    more code

'

sem_model <- sem(sem_string,
                 data = wide_df)
```

where `sem_model` is an object that stores the results of our SEM model and we refer to the `wide_df` data set because SEM requres wide data. Notice that in the SEM code snippet we first create a string object, `sem_string` to specify our effects. Again, the `code` and `more code` will be replaced by actual effects when we get to our inferences. Just like the HLM code, we then view our SEM results with:

```
summary(sem_model)
```

In summary, the style of code snippets just presented will be used throughout each inference section. What will change are the `code` and `more code` pieces, and in those areas we will impute effects specific to each inference. If you wish to run these models on your own computer, you will also need to load their packages in your `r` script with `library(nlme)` for HLM and `library(lavaan)` for SEM (and install them with `install.packages()` if you have not done so already).

We now turn to the inferences.

Relationships

Inference 1

What is the relationship between two or more variables across time?

Explanation

A common inference in our longitudinal literature is the relationship between an outcome (y) and one or more predictors (x_p) over time. As shown in figure one, researchers observe y and one or more predictors at each time point, and are then interested in the immediate effect of the predictors on y. Although there are multiple slices in our figure, typically the effect of x on y is treated as stable over time and therefore the analysis returns an estimate of a single parameter (e.g., a single beta weight). This parameter is essentially a summary statement of the immediate effect of x on y at any point in time. In other words, researchers observe x and y at every t from time t to $t + 5$, as an example, and then report a statement of the effect of x_t on y_t, or the expected immediate effect of x on y at any possible moment.

Although the effect (i.e., the parameter relating x to y) is treated as stable over time, the values on x and y – the raw data – are typically allowed to vary. When the values of x (potentially) change at each observation, the analysis is referred to as a time-varying covariates analysis – whereas a time-invariant analysis would be one where the raw data on x does not change at each observation (e.g., gender).

To clarify, consider this inference with respect to our generic variables: affect x and performance y. Researchers would collect performance and affect data on multiple people at each observation over several time points. Affect may vary at each observation – it may be 5 at time t and 9 at time $t + 1$ – so this analysis is referred to as a time-varying covariates analysis. Researchers are then interested in the stable, immediate influence of affect on performance. That is, they are interested in the effect of affect$_t$ on performance$_t$ at any potential t, where, on average, they expect the effect of affect on performance to be close to their parameter value.

There is a distinction with how we use the term "stable" that merits more explanation. In a relationships inference, "stable" means that we expect the parameter value relating affect to performance to be the same at each moment. In other words, the relationship between affect and performance will be the same at each t – high values of affect will result in high values of performance (if the parameter is positive) and low values of affect will result in low values of performance. "Stable" in this inference context does not mean that affect has a lasting impression on performance, or that affect at time t influences performance at some later time. This distinction is a difficult one, but it represents a major difference betweeen making a relationships inference versus some of the others we have yet to explore. A relationships inference, therefore, is concerned with the average influence over time, or what we expect the immediate effect of x on y to be at any given moment.

Example Hypotheses

Many studies in our literature explore relationships over time. When a researcher is interested in making a relationships inference, they propose a hypothesis of the form: x predicts y (given some direction, positive or negative). Using our generic variables, we would propose:

Affect positively (or negatively) predicts performance over time.

For example, Barnes, Schaubroeck, Huth, and Ghumman (2011) predict a negative relationship between poor sleep and cognitive self control. Similarly, Chi, Chang, and Huang (2015) hypothesize that daily negative mood negatively relates to daily task performance.

Code

We gain even more clarity on this inference when we consider the code used to estimate models. A typical SEM model would be estimated with something similar to the following code.

```
sem_string <- '

      performance.1 ~ b1*affect.1
      performance.2 ~ b1*affect.2
      performance.3 ~ b1*affect.3
```

```
      performance.4 ~ b1*affect.4
      performance.5 ~ b1*affect.5

'
sem(sem_string, data = wide_df)
```

where this code snippet is identical to our introductory SEM code but we have replaced `code` and `more code` with actual effects we wish to estimate. In this case, we regress performance at time t on affect at time t and estimate `b1` – the parameter relating affect to performance over time. The analysis will return one number for `b1` because, again, we are treating this as a stable estimate over time.

The HLM equivalent would be:

```
lme(

  fixed = performance ~ affect,
  random = ~1|id,
  data = long_df

)
```

where we again replace `code` and `more code` from the introductory HLM snippet with a regression of performance on affect. Notice that we do not need to type the performance on affect regressions at each time point as we did with the SEM code. Typically, in SEM software you type out the entire model, whereas HLM packages use condensed code. There are two other new pieces of code: `fixed` and `random`. These commands specify how we want HLM to estimate the parameters we are interested in, but at this point we do not unpack them further – we will return to what they mean in the full modeling section at the end of the paper.

We will devote an entire section to the statistical properties of HLM and SEM below, what is important here is our regression of performance on affect. In both HLM and SEM we estimate a stable parameter relating affect at time t to performance at time t – in SEM software we explicitly identify a parameter (`b1`) whereas HLM software typically does so automatically.

Relationships Inference 1 Summary

- Hypothesis:
  - x relates to y over time.
- Parameters From SEM or HLM:
  - beta coefficient relating x to y at each time point.
- Inference:
  - x immediately predicts y at any moment in time.

Growth

We now move to our second inference panel, B and explore a variety of inferences related to growth.

Inference 1

What is the level of a construct at a given point in time?

References

```r
r_refs(file = "references.bib")
```

Barnes, C. M., Schaubroeck, J., Huth, M., & Ghumman, S. (2011). Lack of sleep and unethical conduct. *Organizational Behavior and Human Decision Processes*, *115*(2), 169–180.

Beal, D. J. (2015). ESM 2.0: State of the art and future potential of experience sampling methods in organizational research. *Annu. Rev. Organ. Psychol. Organ. Behav.*, *2*(1), 383–407.

Chi, N.-W., Chang, H.-T., & Huang, H.-L. (2015). Can personality traits and daily positive mood buffer the harmful effects of daily negative mood on task performance and service sabotage? A self-control perspective. *Organizational Behavior and Human Decision Processes*, *131*, 1–15.

Hardy, J. H., Day, E. A., & Steele, L. M. (2018). Interrelationships among self-regulated learning processes: Toward a dynamic process-based model of self-regulated learning. *Journal of Management*, 0149206318780440.

Johnson, R. E., Lanaj, K., & Barnes, C. M. (2014). The good and bad of being fair: Effects of procedural and interpersonal justice behaviors on regulatory resources. *Journal of Applied Psychology*, *99*(4), 635.

Jones, K. P., King, E. B., Gilrane, V. L., McCausland, T. C., Cortina, J. M., & Grimm, K. J. (2016). The baby bump: Managing a dynamic stigma over time. *Journal of Management*, *42*(6), 1530–1556.

Kozlowski, S. W., & Bell, B. S. (2003). Work groups and teams in organizations. *Handbook* *of Psychology*, 333–375.

Meier, L. L., & Spector, P. E. (2013). Reciprocal effects of work stressors and counterproductive work behavior: A five-wave longitudinal study. *Journal of Applied Psychology*, *98*(3), 529.

Pitariu, A. H., & Ployhart, R. E. (2010). Explaining change: Theorizing and testing dynamic mediated longitudinal relationships. *Journal of Management*, *36*(2), 405–429.

Rosen, C. C., Koopman, J., Gabriel, A. S., & Johnson, R. E. (2016). Who strikes back? A daily investigation of when and why incivility begets incivility. *Journal of Applied Psychology*, *101*(11), 1620.

Shiffman, D. (2012). *The nature of code: Simulating natural systems with processing.* Daniel Shiffman.