---

https://crsh.github.io/papaja_man/index.html

title : "Thinking Longitudinal: A Framework for Scientific Inferences with Temporal Data" shorttitle : "LONGITUDINAL INFERENCES"

author: - name : "Christopher R. Dishop" affiliation : "1" corresponding : yes # Define only one corresponding author address : "316 Physics Road, Psychology Building, Room 348, East Lansing, MI, 48823" email : "dishopch@msu.edu" - name : "Michael T. Braun" affiliation : "2"

- name : "Goran Kuljanin" affiliation : "3" - name : "Richard P. DeShon" affiliation : "1"

affiliation: - id : "1" institution : "Michigan State University" - id : "2" institution : "University of South Florida" - id : "3" institution : "DePaul University"

author_note: |

Dishop, C. R., Braun, M. T., Kuljanin, G., & DeShon, R. P. (2020). Thinking Longitudinal: A Framework for Scientific Inferences with Temporal Data. In Griep, Y. Hansen, S. D., Vantilborgh, T., & Hofmans, J. (Ed.), Handbook on the Temporal Dynamics of Organizational Behavior. Edward Elgar Publishing.

abstract: | In this manuscript we explore how to think about patterns contained in longitudinal or panel data structures. Organizational scientists recognize that psychological phenomena and processes unfold over time and, to better understand them, organizational researchers increasingly work with longitudinal data and explore inferences within those data structures. Longitudinal inferences may focus on any number of fundamental patterns, including construct trajectories, relationships between constructs, or dynamics. Although the diversity of longitudinal inferences provides a wide foundation for garnering knowledge in any given area, it also makes it difficult for researchers to know the set of inferences they

may explore with longitudinal data, which statistical models to use given their questions, and how to situate their specific inquiries within the broader set of longitudinal inferences. Moreover, the diversity of statistical models that can be applied to longitudinal data requires that researchers understand how one inference category differs from another.

keywords : "longitudinal inferences, between-unit, growth, trends, dynamics, relationships over time, processes" #wordcount : "149"

bibliography : ["references.bib"] nocite: "(**???**); (**???**); (**???**); (**???**)"

figsintext : no figurelist : no tablelist : no footnotelist : no lineno : yes

lang : "english" class : "man" output : papaja::apa6_pdf —

Organizational scientists recognize that psychological phenomena and processes unfold over time (**???**; **???**). Individuals in the workplace, over time, strive to accomplish work goals, team members collaborate so the whole eventually becomes greater than the sum of its parts, and managers repeatedly promote values to build vibrant, innovative work cultures. To better understand psychological phenomena, such as motivation, teamwork, and organizational culture, researchers must attend not to static snapshots of behavior (**???**; **???**; **???**) but to actors, events, variables and relationships as they move through time. Obtaining longitudinal data allows researchers to capture the unfolding set of events, interactions, behaviors, cognitions, or affective reactions across a variety of psychological phenomena. If practitioners and scholars, however, are unfamiliar with the many questions that they can ask of longitudinal data – and how those questions do or do not pair with certain statistical models – then the inferences that they draw from it are at best restricted and, at worst, spurious.

Researchers have the opportunity to explore many inferences when they analyze longitudinal data. For example, researchers may examine the shape of trajectories on psychological constructs (e.g., Did job satisfaction generally increase or decrease during the past six months?), how two or more constructs relate to each other (e.g., Did team communication and cohesion positively correlate over time?), or whether changes in one variable relate to changes in another (e.g., Did changes in goal-setting relate to changes in employee performance? **???**; **???**; **???**; **???**; **???**; **???**; **???**). Given the variety of available inferences with longitudinal data, an organizing framework would elucidate their subtle differences and enhance theoretical insight.

We developed a framework that organizes some of the fundamental patterns that researchers may explore with longitudinal data. In this paper, we use it to describe how to think about patterns contained in longitudinal or panel data structures. Our manuscript is timely for two reasons. First, it consolidates disparate literature. The ways of thinking (i.e., inference categories) that we describe are not new, they are contained in the organizational literature already, but they are often discussed in isolation which limits a common understanding of how they fit together. To demonstrate this point, we conducted a brief review of articles published in the *Journal of Applied Psychology* and the *Journal of Business and Psychology* in the years 2017 and 2018 that contained three or more waves of data with every variable measured at each time point. Twenty-eight studies were identified and, using the study hypotheses, introductions, and discussions, classified according to the inference categories that we discuss in this paper. Table 1 reports the frequencies of each inference across the 28 studies. The specific inference categories (trend, relationships, and dynamics) will be fully developed later, what matters here is that a majority of the reviewed studies explored a single inference category, and no study focused on all three. We are not saying that researchers and practitioners must ask all questions possible of their longitudinal data, but we are pointing out that other inferences and ways of thinking about patterns exist that researchers may not be aware of. This paper presents each inference in a single location rather than forcing researchers to locate and parse disparate literature to understand what they can ask of longitudinal data. Second, a recent article noted that, despite a growing emphasis on dynamics in organizational science, certain statistical models are inappropriate for inferences about dynamics (**???**). The authors state that researchers should consider whether their interest is dynamics or other over time patterns and choose their statistical model accordingly. Researchers, therefore, need to know how dynamic inferences differ from other, related inferences. Here, we fully describe those differences.

Longitudinal Research in Applied Psychology

This paper is devoted to inferences with repeated measures, so we begin with a few labels and definitions. Authors typically identify a "longitudinal" study by contrasting 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 (**???**). We therefore emphasize differences on the number of observations when we distinguish longitudinal from other types of research. Longitudinal or panel *data* are repeated observations on several units (i.e., N or $i > 1$), whereas time-series data are observations of one unit over time – a distinction that focuses on the amount of people in the study (given repeated measures). Most organizational studies collect data on more than one unit, therefore our discussion below focuses on longitudinal research with panel data, or designs with $N > 1$, $t \geq 3$, and the same construct(s) measured on each i at each t. That is, we focus on designs with repeated measures across many people (units) where every variable is measured at each time point.

Longitudinal applies to both short and long-term research. An experiment with ten trials is longitudinal, as is a study spanning 10 years that assesses its measures once every year. Longitudinal is not reserved for "long-term" studies that last more than one year irrespective of the frequency of their observations. Rather, certain processes unfold over short time horizons (e.g., decision-making on simple tasks, swift action teams; **???**) whereas other psychological phenomena unfold over long time horizons (e.g., the development of a shared organizational culture; **???**), so the informativeness of a particular study depends on its rationale, research design, analytical work, and effective interpretation of results – as with any study. Short and long time horizons both offer valuable insights.

Insert Table 1 about here

Framework for Longitudinal Inferences

We use three inference categories to partition our discussion, including trends, relationships, and dynamics. Briefly, longitudinal inferences focusing on trends assess whether trajectories follow a predictable growth pattern, longitudinal inferences focusing on relationships assess the relationship between one or more constructs, and longitudinal inferences focusing on dynamics assess how one or more constructs move through time as functions of themselves and each other and emphasize how the past constrains the future. Each category comes with box-and-arrow model heuristics[1] that represent the broad inferences, research questions to orient the reader as to what the sub-inferences capture (i.e., inferences are the answers to the research questions that we present), and references for statistical models that are appropriate for a given inference.

Consistent with the typical inferences in organizational science, the inferences described here are between-unit, meaning that they reflect the average relationship across the unit of focus for a particular study (e.g., the relationship across people, or the relationship across firms). The typical statistical model applied to panel data in organizational science reflects the

[1] Note that statistical models differ from the term, "model heuristic." A model heuristic is a visual representation only, whereas a statistical model is characterized by a formula explaining the data and assumptions on the errors, and the parameters of statistical models are estimated using an estimation technique. In this paper, we never use the term, "model" without pairing it either with "statistical" or "heuristic" – the two differ substantially.

average relationship of interest across units in a given sample, providing a population-level estimate of a given relationship. Alternatively, time-series designs focus on single units and all inferences are interpreted at the within-unit level, generating a unit-level estimate of a given relationship. Time-series designs, however, are not popular in our literature so they are not described further.

Although we use box-and-arrow diagrams throughout to represent the broad inferences, we also graph a few of the more challenging inferences with mock data – some of the inferences in the trend and relationships sections are difficult to grasp without seeing them in data form. Keep in mind, however, that data are always messy. It is rare to find data in which the inferences present themselves simply by plotting – althought it is certainly possible. We use these "data plots" to clearly convey what the inferences mean, but be aware that field data are often noisy.

Finally, despite pointing researchers to statistical models, our paper puts a majority of its emphasis on inferences, therefore researchers need to be sure that they appreciate all of the nuance before applying a recommended statistical model. Numerous issues arise when modeling longitudinal data and the statistical models differ in how they handle these issues, the assumptions they make, and the data format they require. We do not speak directly to those issues here, but we refer readers to a number of informative references for each statistical model.

Trend

Made popular in the organizational literature by (**???**) and (**???**), trend inferences represent a class of thinking in which researchers create an index of time and relate it to their response variable to understand the trajectory of the dependent variable. The first panel of Figure 1 shows a box-and-arrow model heuristic in which time is related to an outcome, y, and ultimately the analyst is interested in a variety of questions about trend and its correlates.

Trend inferences have two components: trend itself and level. For clarity, we discuss them separately.

Component 1 - Trend. Does affect, in general, increase or decrease across time, or is its trajectory relatively flat? Does trainee skill generally increase over the training session? These are questions about trend, and these first two are specifically about linear trend. It is also possible to explore how variables bend or curve across time. Do newcomer perceptions of climate increase and then plateau over time? Does the response time of a medical team decrease with each successive case but then remain stable once the team can no longer improve their coordination? These latter questions concern curvilinear trajectories.

Trend has to do with the systematic direction or global shape of a trajectory across time. If we put a variable on the y-axis and plot its values against time on the x-axis, do the values display a stable temporal pattern? It can be thought of as the coarse-grained direction of a trajectory. A positive trend indicates that, on average, we expect the variable to increase over time and a negative trend indicates that we expect the variable to decrease over time. Our first trend research question, therefore, concerns the shape of the trajectory.

> **Research Question 1:** On average in the population, is there a positive/negative/curvilinear trend?

Many research questions and inferences begin with the average pattern (or relationship) and then move to variability, the same applies here. After learning about the average trend across the sample, researchers then focus on trend variability. How much consistency is there in the trend pattern? Do all trainees develop greater skill across time? Is there variability in the trend of helping behaviors, or counterproductive work behaviors over time?

> **Research Question 2:** Does trend differ across units?

Research questions one and two concern one variable, but they can also be iterated across all observed variables. For example, we might discover that – on average – affect and performance trends both decrease, but there is greater variability across people in the affect trend. Or we might learn that affect has a negative trend while performance has a positive trend.

Correlating these trends is the next inference. Correlating indicates co-occuring patterns, where a large, positive, correlation between affect and performance trends indicates that people with a positive affect trend (usually) have a positive performance trend and people with a negative affect trend (usually) have a negative performance trend.

Figure 2 shows the inuition behind this inference with a set of graphs. In Panel A, we plot affect and performance across time for three individuals. Affect goes up while performance goes down for person one, this pattern is reversed for person two, and person three reports trendless affect and performance (i.e., zero trend).

Panel B then maps those pairings onto a scatterplot that demonstrates the (between-person) relationship among affect and performance trends. For example, person one has a positive affect trend and a negative performance trend, so their value in Panel B goes on the bottom right, whereas person two has the opposite pattern and therefore is placed on the top left (where the performance trend is positive and the affect trend is negative). Producing this bottom scatter plot tells us that the between-person association among affect and performance trends is negative. That is, people with a positive affect trend are expected to have a negative performance trend, people with a negative affect trend are expected to have a positive performance trend, and people with an affect trend of zero are expected to have a performance trend of zero.

Insert Figure 2 about here

Research Question 3: On average in the population, what is the relationship among two trends?

The final trend inference is about identifying covariates or predictors of trend. Does gender predict depletion trends? Does the trend in unit climate covary with differences in leader quality?

Figure 3 demonstrates the inference in a plot. We graph affect across time for six employees that differ by job type. The first three individuals work in research and development, whereas the final three work in sales. Affect trajectories tend to decrease over time for employees in research and development, whereas affect trajectories tend to increase for those in sales. An individual's job type, then, gives us a clue to their likely affect trend – said formally, job type covaries with affect trend, such that we expect individuals in sales to have positive affect trends and individuals in research and development to have negative affect trends.

Insert Figure 3 about here

Research Question 4: On average in the population, what is the relationship between a trend and a covariate?

Note the difference between research questions three and four. Both are between unit, but three is about co-occuring trend patterns whereas four is about the relationship between trend and a covariate/predictor. With respect to our examples, inference three (i.e., the answer to research question three) says, on average, if an individual has a positive affect trend then we expect her to have a negative performance trend. Inference four says, on average, if an individual is in research and development then we expect him to have a negative affect trend.

Component 2 - Level. Researchers that explore trend also assess its predicted value at a given time t, and this second component is called level. Level is almost always evaluated at the first or last observed time point – e.g., What is the predicted level of the trainee skill trend, on average across trainees, at the beginning of a training session? On average across departments, what is the expected level of the department climate trend at the end of a two-week socialization process? – although researchers are free to asssess level at any t.

> **Research Question 5:** On average in the population, what is the expected level of the y trend at time t?

After exploring the average trend level, we then move to its variability. Trend lines have a beginning (or end) point, how consistent do we expect that point to be across the sample? Is there variability in affect trend starting level? Are there large differences in the expected level of the performance trend at the last time point?

> **Research Question 6:** Is there variability across units in the expected level of the y trend at time t?

It is also possible to assess correlations among level and (a) trend in the same variable or (b) level or (c) trend in a different variable. First, consider a relationship among level and trend in the same variable. On average, do people with low initial skill show positive skill trends whereas people with high initial skill show negative skill trends? Do organizations with high initial CWBs, on average across organizations, tend to have negative CWB trends?

> **Research Question 7:** On average in the population, what is the relationship between trend and level in y?

Second, consider a correlation between level in one variable and level in another. On average across people, do people with low initial performance also have low initial depletion (based on the initial levels predicted by the performance and depletion trends)? Are organizations with high initial turnover also expected, on average across organizations, to have high burnout (based on the initial levels predicted by the turnover and burnout trends)?

> **Research Question 8:** On average in the population, what is the relationship between level of the x trend and level of the y trend at t?

Finally, researchers are free to mix the inferences above and assess whether levels in one variable covary with trend in another. Are people with high initial voice (predicted by the voice trend) expected to have negative satisfaction trends?

> **Research Question 9:** On average in the population, what is the relationship between the level of the x trend at time t and the trend in y?

A note on phrasing. The inferences we explored in this section have to do with correlating levels and trends, where a statement like, "affect and performance trends covary between-units, such that people with a negative affect trend have a positive performance trend" is appropriate. There is a less precise phrase that is easy to fall into – and we have seen it used in our literature. Sometimes, researchers will correlate trends and then state, "when affect decreases performance goes up." We urge researchers to avoid this second statement because it is not clear if it refers to a static relationship about trends or a statement about how trajectories move from one moment to the next. That is, the phrase "when affect decreases performance goes up" could refer to between-unit correlated trends, but it could also mean something like, "when affect decreases performance immediately or subsequently goes up." This second statement is far different and it should not be used when an analysis only correlates trends or evokes predictors of trend. Again, we urge researchers to phrase their inferences as we show here.

Literature on Statistical Models for Trend and Level

Currently, the dominant method for analyzing longitudinal data with respect to trend and level inferences in the organizational sciences is growth curve modeling (GCM; **???**; **???**). Broad theoretical discussions of growth are in (**???**) and (**???**) (keep in mind that they call growth "change"), whereas (**???**) describe actual growth curve analysis. Growth curves are a core topic in developmental psychology, so there are many articles and textbooks to read from their field. See (**???**) and (**???**) for two great textbooks on growth curve modeling and (**???**) for an empirical discussion. Two straight-forward empirical examples from our own field include (**???**) and (**???**).

GCM is able to model the dependent variable – performance, for example – as a result of predictors at multiple levels of analysis. Level one predictors vary at the same level as the depenent variable, meaning that if individual performance is the outcome of interest then level one predictors might include individual goal-striving and individual cognitive ability. If, on the other hand, the outcome is organizational performance then level one predictors might include organizational climate or culture. Level two predictors occur at higher units of analysis – team cohesion if the dependent variable is individual performance or national culture if the dependent variable is organizational performance. Variables at any level can enter into the statistical model either as fixed or random. Fixed predictors estimate only the average relationship across all units, whereas random predictors estimate not only the average IV-DV relationship across units but also estimate the degree of between-unit variability in the relationship.

The most basic growth model is the unconditional means model (UMM). Using notation from (**???**), this statistical model is specified as

$$Y_{ij} = \pi_{0i} + \varepsilon_{ij} \tag{1}$$

$$\pi_{0i} = \gamma_{00} + \zeta_{0i} \tag{2}$$

where $\varepsilon \sim N(0, \sigma^2_\varepsilon)$ and $\zeta_{0i} \sim N(0, \sigma^2_0)$, Y_{ij} is the dependent variable measured for person i at time j, π_{0i} is the mean of Y for individual i, γ_{00} is the mean of Y across everyone in the population, ε_{ij} is the residual for individual i on occasion j, σ^2_ε is the pooled within-person variance of each individual's data around his or her mean, ζ_{0i} is the random effect for individual i (i.e., deviation of the person-specific mean from the grand mean), and σ^2_0 is the random effect variance.

In words with individual performance as an example, this UMM says that performance for Rachel at any time is a function of her across time individual performance mean and error (equation one). Moreover, Rachel's individual performance mean is a function of the population performance mean (i.e., the mean of everyone's individual performance) and error (equation two). What this statistical model embodies is that (1) a reasonable prediction for Rachel's performance given no other information, such as other predictors like goal-striving or cognitive ability, is the mean of individual performance and (2) there are between-person differences in individual performance.

The initial UMM model is typically used to calculate the intraclass correlation coefficient (ICC(1)) which estimates the proportion of total variance attributed to between-unit differences. Researchers, for example, would conduct an initial UMM on individual performance and then state, perhaps, that 57% of the total variance resides between individuals. It is also possible to conduct a χ^2 test to assess whether the estimated between-person variance differs from zero. Both results are used to argue that it is reasonable to move forward with more complicated statistical analyses, to include additional predictors that may explain (in a statistical sense) the observed between-person variability.

Assuming differences across units exist, it is then recommended to conduct the unconditional linear growth model (ULGM). The ULGM regresses the dependent variable on a fixed linear time variable while allowing variability across units in intercepts. The regression weight on the time variable (ie., slope) models the average trend across units and is used to answer RQ1. It is then common to allow the Time-DV relationship to vary across units by entering the time variable as a level one random predictor, as shown in the equation below.

$$Y_{ij} = \pi_{0i} + \pi_{i1} Time_{ij} + \varepsilon_{ij} \tag{3}$$

$$\pi_{0i} = \gamma_{00} + \zeta_{0i} \tag{4}$$

$$\pi_{i1} = \gamma_{10} + \zeta_{1i} \tag{5}$$

where $\varepsilon \sim N(0, \sigma^2_\varepsilon)$ and $\begin{bmatrix} \zeta_{0i} \\ \zeta_{1i} \end{bmatrix} \sim N\Big(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2_0 & \sigma_{10} \\ \sigma_{10} & \sigma^2_1 \end{bmatrix} \Big)$, π_{0i} is now the initial status (i.e., intercept) of Y for individual i, γ_{00} is the average initial status of Y across everyone in the population, π_{1i} is the rate of change (i.e., slope) of Y for individual i, γ_{10} is the average rate of change of Y across everyone in the population, σ^2_ε is the pooled variance of each individuals' data around his or her linear change trajectory, ζ_{0i} is the intercept random effect for individual i, σ^2_0 is the variance of intercept random effects, ζ_{1i} is the slope random effect for individual i, σ^2_1 is the variance of the slope random effects, σ_{10} is the population covariance between intercepts and slopes, and all other terms are defined above.

In words with individual performance as example, this ULGM says that Rachel's performance is a function of her initial level of performance (which is a function of the population initial performance level and error) and time (equation 3). Time, therefore, can be thought of as a predictor in the case of the ULGM which makes it an inherently static, as opposed to dynamic, statistical model (**???**) and emphasizes description rather than explanation because time cannot be a true underlying cause (**???**). We moved beyond the unconditional means model to “explain” more variation in Rachel’s performance by including additional predictors. In the case of the ULGM, our additional predictor is time and the estimate of the coefficient relating it to the outcome describes both the expected trend and whether there are between-unit differences in trend in the sample (research questions one and two). Time is a level one predictor because it varies on the same level as the outcome (the individual level) and it is incorporated in the statistical model as a random effect (equation 5).

To review, we first modeled individual performance as a function of across time individual performance means (which, themselves, were functions of the population individual performance). That basic statistical model was a UMM and we turned it into an ULGM by incorporating time as a predictor. Once time enters the equation, we update our view of performance and it becomes a function of initial performance level and time, meaning that one-unit increases in time are seen as relating to increases or decreases in performance. Those increases and decreases across time, in aggregate, form trend. In practice, the statistical model returns one number for the estimate of the coefficient relating time to the outcome and it describes the expected between-person performance trend.

Understanding the basic ULGM with time as a random level one predictor allows researchers to explore, with simple extensions such as incorporating additional predictors or modeling two or more variables as outcomes (multivariate systems), any number of further inferences. Researchers can enter additional multiples of time as predictors – e.g., include $Time^2$ and/or $Time^3$ in equation 3 – to determine whether trajectories are curvilinear or follow other temporal patterns. Researchers can also enter additional level one or two substantive predictors to determine whether there are covariates of trend. Consider, again, the example in Figure 3 which plots affect trends that differ by job type. Affect is the outcome that is regressed on time, forming the underlying (descriptive) ULGM. Entering job type as a random, level two predictor returns a coefficient that describes whether the expected affect trend differs according to this additional predictor. Statistically, the model estimates whether higher values on the level two predictor relate to stronger IV-DV relationships. In the case of growth models, time is the IV so "stronger IV-DV" relationships means different trend patterns. The level two predictor therefore estimates whether higher values – or in the case of job type, different types of jobs – demonstrate different trend (RQ4). Beyond incorporating more predictors of a single outcome, researchers can also model multiple outcomes with simultaneous ULGMs. Consider two independent ULGMs, one with individual performance regressed on time and another with individual OCBs regressed on time. All inferences, research questions, and statitical models described above can be explored independently with these two outcomes. Typically, however, when random effects are incorporated covariances among all random effect variables are estimated, meaning that the two outcomes in the multivariate system are no longer viewed as independent. The covariance estimate between the slope term for performance and the slope term for OCBs, in this example, are used to answer research question three.

Once time is included in the statistical model (e.g., the ULGM), the intercept value represents the level of the DV at the time point coded 0 (typically the first or last time point). The intercept value is almost always modeled as random whereby the analysis will return a mean estimate which tells you the average level across units (answering RQ5), and it will also return a variance estimate that indicates variability in level between units. As such, the variance component on the intercept term determines whether there is significant between-unit variability in the level of the DV when time equals zero, answering RQ6. As previously stated, it is common to estimate covariance among random predictors, therefore the covariance between the intercept and slope random effects is used to determine whether units with higher (lower) initial values exhibit stronger (weaker) growth, answering RQ7. Finally, it is also possible to estimate covariances among the intercept and slope among different variables in multivariate systems, answering RQ8 and RQ9.

Relationships

A relationships inference explores between-unit relationships over time. The second panel of Figure 1 shows a model heuristic, where a predictor is concurrently related to a response variable at each time point and the relationship is typically constrained to equality or is averaged over time. Essentially, the inference compiles single-moment correlations. For example, we assess the correlation between, say, OCBs and depletion at time one, again and times two and three, and then ultimately take the average of each individual, between-unit correlation.

Questions about static relationships over time take the following forms. What is the relationship between helping behaviors and incivility? Does burnout predict turnover intention? Is unethical behavior related to self-control?

Figure 4 shows the inuition of the inference with data. Panel A plots affect and performance trajectories for three people. The triangles in Panel A highlight each individual's affect and performance values at time point six. Given that we have three people at time point six, we can calculate a correlation between affect and performance at that moment (granted, it is a small sample). The calculated coefficient is then graphed in Panel B with another triangle. At time point six, the (across person) correlation among affect and performance is large and positive.

Insert Figure 4 about here

Panel B also shows across-person correlation coefficients for the rest of the time points. Often these correlations are either averaged over time or constrained to be equal. Note that when a researcher uses a time-varying covariates, hierarchical linear, random-coefficient, or multi-level model on longitudinal data to explore concurrent relationships among one or more variables (and they are not analyzing trend) they are making this inference.

> **Research Question 1:** On average in the population, what is the relationship between x and y? (Typically constrained to be equal or averaged over time).

The first relationships inference emphasizes the expected average. As with the trend inferences, the next question is to examine variability in that estimated relationship across the sample. How consistent across the sample is the relationship between distractions and fatigue? Is there variability in the relationship between emotions and volunteering behaviors?

> **Research Question 2:** What is the variability across units in the relationship among x and y?

Literature on Statistical Models for Relationships

Time-varying covariates (TVC) analysis is the workhorse behind relationship inferences. TVC models are simply growth curve models that include level 1 predictors (either fixed or random). The equation below shows an ULGM ammended to include an additional level one predictor, X.

$$Y_{ij} = \pi_{0i} + \pi_{i1} Time_{ij} + \pi_{i2} X_{ij} + \varepsilon_{ij} \tag{6}$$

$$\pi_{0i} = \gamma_{00} + \zeta_{0i} \tag{7}$$

$$\pi_{i1} = \gamma_{10} + \zeta_{1i} \tag{8}$$

$$\pi_{i2} = \gamma_{20} + \zeta_{2i} \tag{9}$$

where X is the additional random predictor and it is related to the outcome, Y, through π_{i2}. The average relationship between X and Y across units is used to answer RQ1 whereas the variance component estimating the between-unit variaiblity in the X-Y relationship is used to answer RQ2. It is important to note that TVC analyses can either be conducted by building upon the ULGM, as is presented here, or can be done by building directly upon the UMM (e.g., **???**). The difference is whether the predictor time is included to control for growth in the DV. Typically, if it is anticipated or observed that the DV exhibits a consistent trajectory over time, then time is included and TVC models build from the ULGM. Alternatively, if the DV is not expected or observed to exhibit linear (or curvilinear) growth, then level one predictors are added directly to the UMM. Complete discussions of TVC models are in (**???**) and (**???**) and two relatively straight-forward empirical examples are in (**???**) and (**???**).

Dynamics

The word "dynamics" takes on a variety of meanings throughout our literature. Informally, it is used to mean "change," "fluctuating," "longitudinal," or "over time" (among others), but the fundemantal concept to identify with dynamics is that the past constrains what happens next, variables have memory as they move through time. For example, (**???**) notes that in dynamic analysis, "it is essential to know how variables depend upon their own past history" (p. 409), (**???**) state that dynamic variables "behave as if they have memory; that is, their value at any one time depends somewhat on their previous value" (p. 604), and Wang, Zhou, and Zhang (**???**) define a dynamic model as a "representation of a system that evolves over time. In particular it describes how the system evolves from a given state at time t to another state at time $t + 1$ as governed by the transition rules and potential external inputs" (p. 242). In this section we discuss a number of inferences couched in the idea that the past constrains future behavior.

Does performance relate to itself over time? Do current helping behaviors depend on prior helping behaviors? Does unit climate demonstrate self-similarity across time? Does income now predict income in the future? These are questions about the relationship of a single variable with itself over time – does it predict itself at each subsequent moment? Is it constrained by where it was in the past?

Panel A of Figure 5 shows the concept with a box-and-arrow model heuristic. y predicts itself across every moment – it has self-similarity and its value now is constrained by where it was a moment ago. In our diagram, we show that y at time t is related to y at time $t+1$. In other words, we posit that y shows a lag-one relationship, where y is related to its future value one time-step away. Researchers are of course free to suggest any lag amount that they believe captures the actual relationship. Note that the statistical term to capture self-similarity or memory is called autoregression.

> **Research Question 1:** On average in the population, what is the relationship of y to itself over time? (Autoregression)

Insert Figure 5 about here

As before, after exploring the expected average we turn to variability. How consistent are the self-similarity relationships? Are there between-unit differences in autoregression in, for example, employee voice? Do we expect a large variance in the autoregression of helping behaviors?

> **Research Question 2:** What is the variability across units in the expected autoregression of y?

The next inference is about relating a predictor to our response variable while it still retains memory. Panel B of Figure 5 shows a box-and-arrow diagram: y is predicted by concurrent values of x but it also retains self-similarity. This model heuristic is therefore said to partial prior y: it examines the concurrent relationship between x and y while statistically partialling values of y at $t-1$, or statistically accounting for y at the prior moment.

Our literature has converged on calling this kind of relationship "change" because it emphasizes the difference between y now and where it was in the past (e.g., **???**; **???**). The association asks how current x relates to the difference between y now and its immediately prior value. How does affect relate to change in performance? Does depletion covary with change in OCBs? Note that change can be construed from any prior time point (baseline, $t-1$, $t-3$); our literature commonly emphasizes lag-one change.

> **Research Question 3:** On average in the population, what is the relationship betweeen concurrent x and change in y?

The analyst is also free to assess variability in the expected change relationship.

> **Research Question 4:** What is the variability across units in the expected change relationship between concurrent x and y?

Change relationships do not have to be concurrent. Panel C of Figure 5 shows concurrent relationships as we saw above but it also includes lags from the predictor to the outcome. y retains memory, but it is predicted by both concurrent and prior values of x. Typically, we call these cross-lag relationships.

Questions about lag-one change relationships take the following forms. Does affect predict subsequent performance change? Do prior counterproductive work behaviors relate to current incivility change? Does metacognition predict subsequent exploratory behavior change? Of course, researchers can also explore longer lags by relating predictors to more distal outcomes.

> **Research Question 5:** On average in the population, what is the cross-lag relationship between x and change in y at a different point in

time?

Again, typically researchers explore variability after assessing the average estimate.

Research Question 6: What is the variability over units in the expected cross-lag relationship of change?

Extensions

We described a simple set of inferences above, but the ideas generalize to more complex dynamics as well. Often researchers are interested in reciprocal relationships, where x influences subsequent y, which then goes back to influence x at the next time point. Said formally, x_t influences y_{t+1}, which then influences x_{t+2}. Said informally, current performance influences subsequent self-efficacy, which then influences performance on the next trial. These inferences are no different than what we saw above – they are cross-lag predictions – all we did was add more of them. Panel D of Figure 5 shows reciprocal dynamics, in which both x and y show self-similarity and cross-lag relationships with one another.

Researchers typically posit a sequence of single cross-lag predictions when they are interested in reciprocal dynamics. For example, (**???**) explored reciprocal relationships among performance and motivation (self-efficacy, metacognition, and exploratory behavior). Their hypotheses include, (1) prior self-efficacy negatively relates to subsequent exploratory behavior and (2) prior exploratory behavior positively relates to subsequent self-efficacy (among others). These single inferences are used in aggregate to make conclusions about reciprocal influence.

The dynamic inferences shown here also generalize to systems of variables where a researcher posits self-similarity and cross-lag predictions across many variables. There could be reciprocal dynamics between a set of variables like performance, self-efficacy, and affect, or a sequence of influence between dyadic exchanges, performance, and team perceptions:

perhaps initial dyadic exchanges influence subsequent team perceptions, which later influence performance. Following the performance change, the structure of the task updates and this initiates new dyadic exchanges. Once a researcher grasps the foundational ideas presented here he or she is free to explore any number of complex relationships.

Literature on Statistical Models for Dynamics

(**???**) review a variety of dynamic models and, although their paper does not provide readers with specific code, it is an excellent resource to become familiar with potential dynamic models. (**???**) describe why multi-level models are innapropriate for inferences about dynamics and instead recommend a general panel model described in (**???**). See (**???**) for a similar discussion. Other statistical models that are appropriate for dynamic inferences are discussed in (**???**), (**???**); (**???**), (**???**), (**???**), and (**???**). Finally, (**???**) discuss a number of dynamic statistical models.

The principle addition for dynamic models is the inclusion of a lagged version of the DV as a predictor (Y_{t-1}). The inclusion of Y_{t-1} controls for prior observations of the DV when predicting current values, essentially modeling the change in the DV from one time point to another without relying on difference scores (e.g., **???**). As such, the first research question is answered by evaluating the average relationship between the DV and a prior version of itself as a level one predictor. Similarly, once the autoregressive term is modeled as random, evaluating the variance component answers RQ2 regarding whether the autoregressive relationship differs across units. To answer the subsequent research questions, the inclusion of an additional substantive predictor, X_t, is required. When X_t is modeled at only the concurrent time point with the DV, then the $X_t \rightarrow Y_t$ relationship determines whether values of X at a given time point relate to the change in Y, addressing RQ3. The variance component on X_t when it is modeled as a random level one preditor determines whether the relationship varies across units, answering RQ4. Finally, if the researcher is interested in determining whether changes in the predictor, X, relate to changes in the DV, Y, an additional level one predictor in included in the model that represnts prior realizations of X, X_{t-1}. With the inclusion of X_{t-1}, the paramater on the predictor X_t now determines whether changes in X relate to changes in Y, answering RQ5 whereas the variance component on X_t determines whether significant variability in the relationship exists (RQ6). There are many additional dynamic models that can be estimated within the GCM framework. (**???**) review a variety of dynamic models and, although their paper does not provide readers with specific code, it is an excellent resource to become familiar with potential dynamic models.

Discussion

There are many different patterns to explore with longitudinal data structures. This manuscript, by unpacking between-unit patterns, mirrors the common questions and inferences currently emphasized by organizational scientists. What is the relationship among a set of constructs (averaged over time)? What is the expected trend? Are there differences in trend (also phrased as, "between-unit differences in within-unit change")? We organized these questions and inferences into a fundamental set, discussed what they mean, and consolidated disparate literature so that researchers have a single source to better understand how to think about over time patterns. Our discussion and figures were designed to reduce ambiguity about patterns in longitudinal data and help researchers ask questions effectively. Ultimately, researchers should now be able to understand the spectrum of inferences that they can explore with rich, longitudinal data.

Between-unit questions are common and useful and they are the sibling to an alternative lens to asking questions and making inferences with repeated measures: within-units. Within-unit inferences emphasize fluctuations over time rather than across units. For example, Beal (**???**) notes that many of the psychological phenomenon in which we are interested are "sequences of events and event reactions that happen within each person's stream of experience" (p. 5). This is a within-unit statement: it emphasizes how a construct moves through time within a single individual. Organizational scientists have become increasingly interested in within-unit perspectives over the past decade. (???) review theory and research on within-person job performance, (???) review emotional labor and differentiate a variety of within-person perspectives, (???) present a team motivation model describing within-individual resource allocation and within-team feedback, (???) present a within-person model of multiple-goal pursuit, (???) describes recent within-person approaches to sleep and organizational behavior, and (???) present a within-person perspective of organizational citizenship behaviors. There are many within-person inferences accumulating in our literature, but they are occasionally accompanied by between-person models or are dispersed and unconnected among different content areas. An immediate next step for future research is to create a framework for the fundamental within-unit inferences.

Our focus was on between-unit patterns because these inferences are the backbone of longitudinal modeling and inference in organizational science. Moreover, there can be a tendency for researchers to believe that they are making within-unit inferences simply because they collect longitudinal data, our goal was to build consensus and clarity on the fundamental between-unit ideas in longitudinal data structures.

We close with four key takeaways. First, knowing the different patterns that you can explore with longitudinal data is crucial to operating as an organizational scholar. These data sets are becomming more and more common (???) so graduate students, practitioners, and academics are bound to come across such data in their career. Without knowing the fundamental patterns that are possible to explore, researchers will have a hard time asking good questions to guide their research designs, have a limited understanding of how to evaluate other longitudinal research, and miss potential insights. The manuscript has utility for both teachers and learners.

Second, questions come before statistical models. We presented ways of asking questions and thinking about longitudinal data, and only after discussing those concepts were possible citations for statistical models provided. The insight here is that questions should drive research design, data collection, and statistical modeling. A researcher first asks, "what questions do I have?" "What inferences do I want to make?" And only after thinking hard about those concepts moves to questions about reserach design, data collection, and possible statistical models that can be applied to the data that align with the initial question. A researcher should never couch him or herself within a single inference category simply because he or she is only familiar with one statistical model. What is your question? What inference do you want to make? Then choose a research approach (i.e., research design, data collection, and statistical model) that is consistent with those interests.

Third, and building off point two, certain questions and inferences require specific statistical models. (**???**) demonstrated that multi-level (or hierarchical linear) models are innapropriate for inferences about dynamics. (**???**), (**???**), (**???**) state that the assumptions (e.g., stationarity) of multi-level models may not best suit data following dynamic patterns. Knowing how inferences about dynamics differ from inferences about trends or relationships, therefore, is critical for choosing an appropriate statistical model.

Fourth, all of the ideas presented here can be used in harmony to learn about the temporal nature of a phenomenon. A researcher does not have to limit him or herself to a single inference category. Each is a unique way of asking questions about patterns contained in longitudinal data structures and, after potentially asking many questions about the same assessed constructs, a researcher can learn about multiple aspects of the phenomenon.

References

Table 1

Number of times a recent article emphasized one or more inference category.

Type	Occurence
Trend	4
Relationships	13
Dynamics	10
Any 2	1
All 3	0

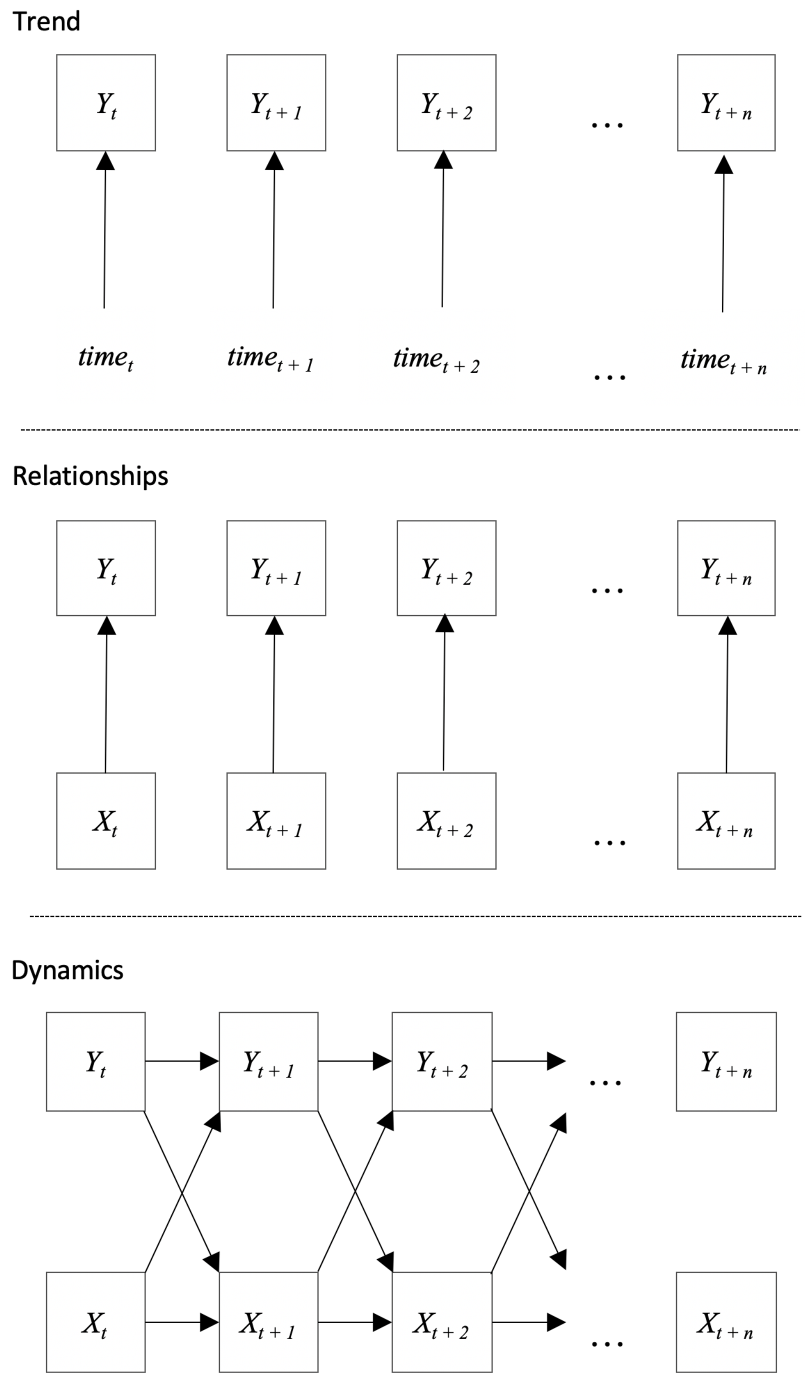

Figure 1. Common inference categories with models applied to longitudinal data.

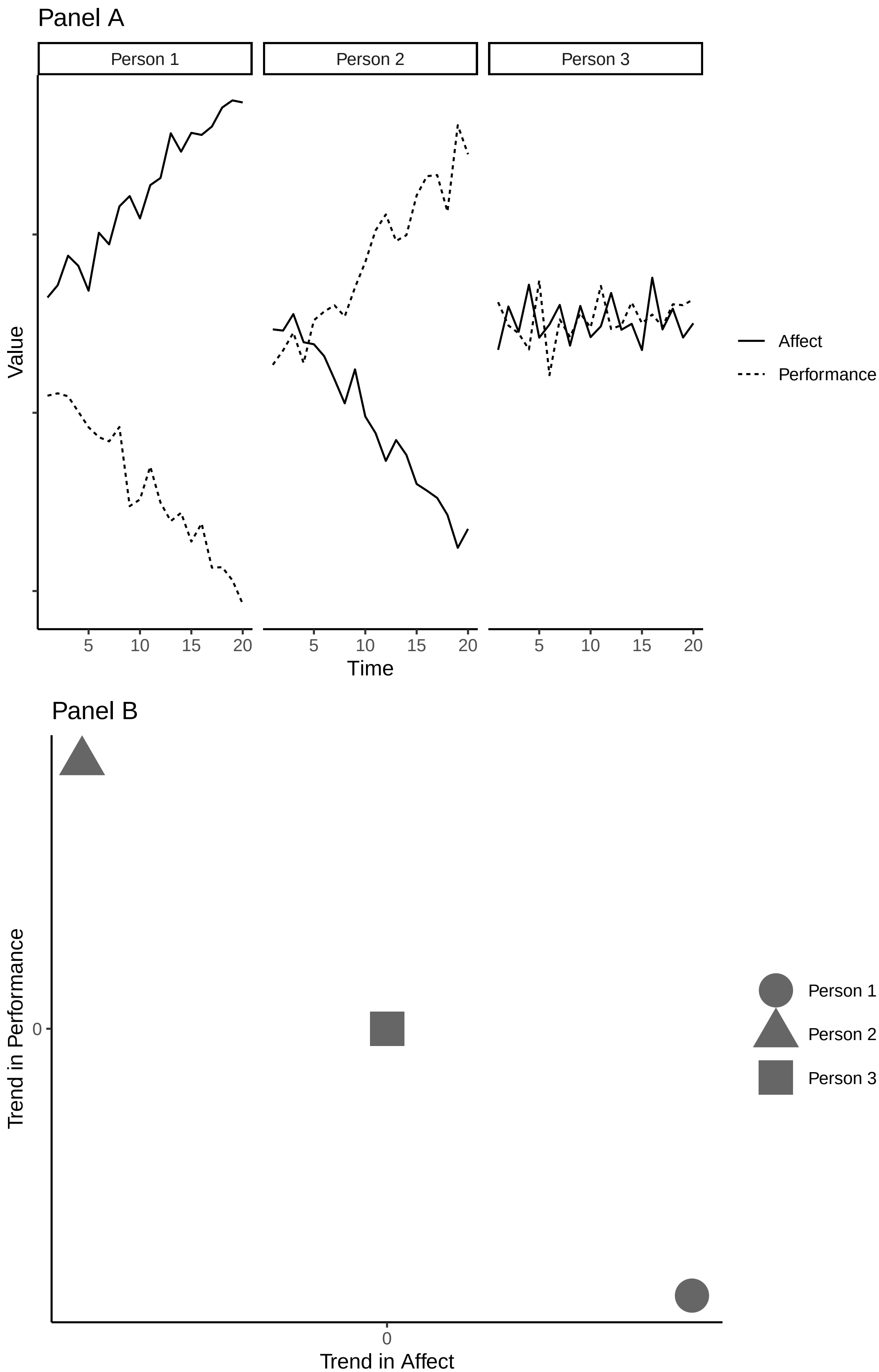

Figure 2. Between-person correlation of trend in affect and performance.

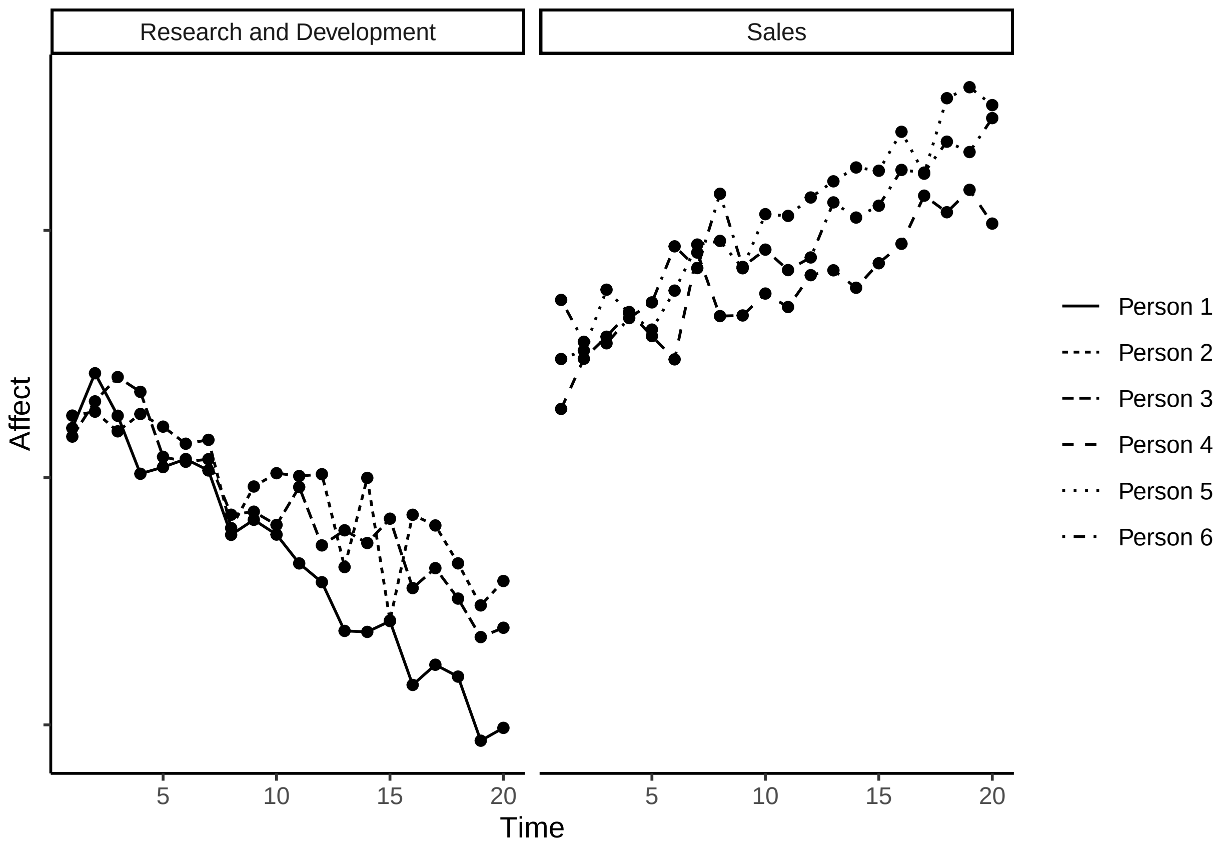

Figure 3. Job type as a covariate of affect trend.

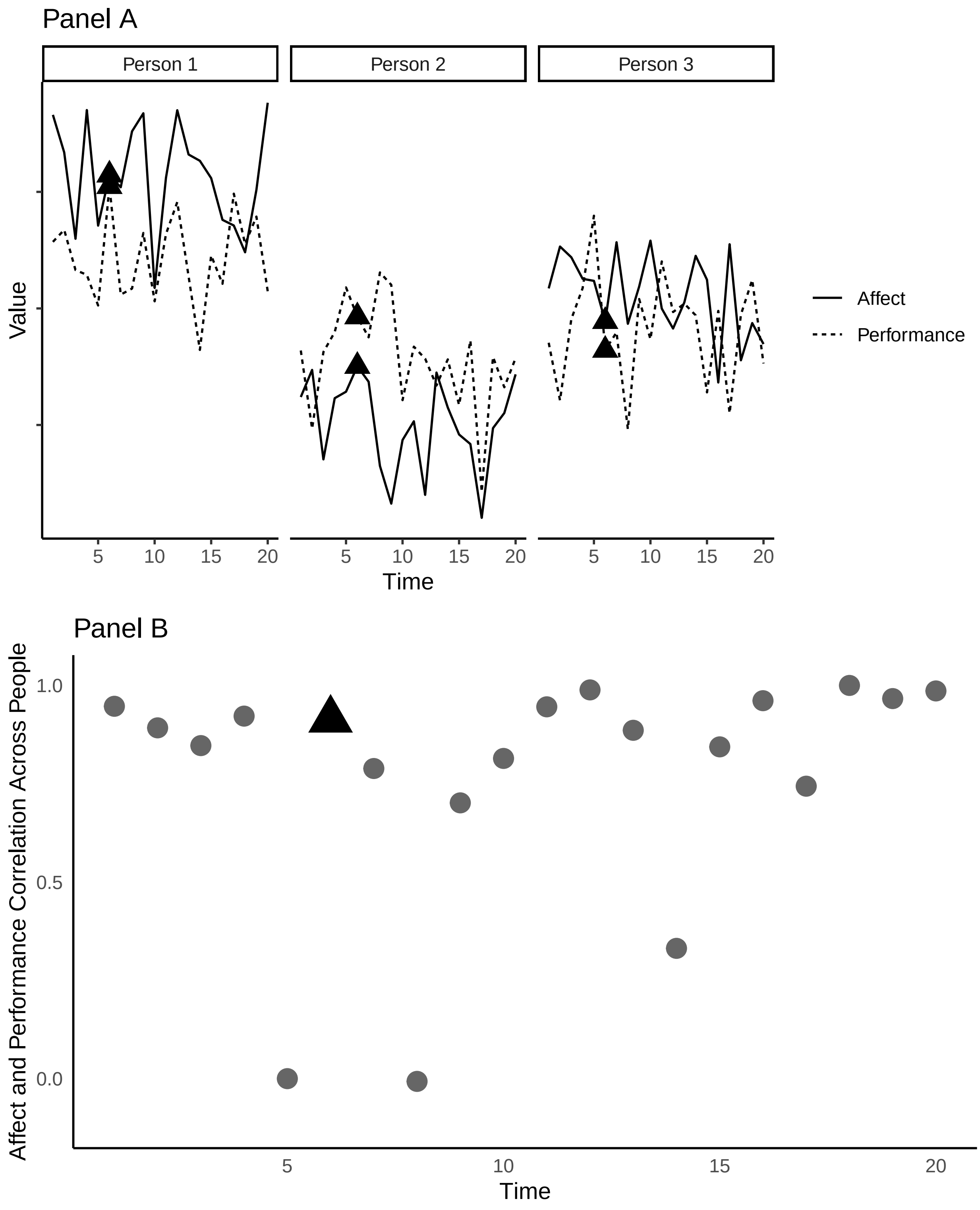

Figure 4. Relating affect to performance across people over time. The triangles demonstrate the between person correlation at time point six. A typical time-varying covariates model constrains the correlation to be equivalent across time. Here, the relationship is unconstrained at each time point.

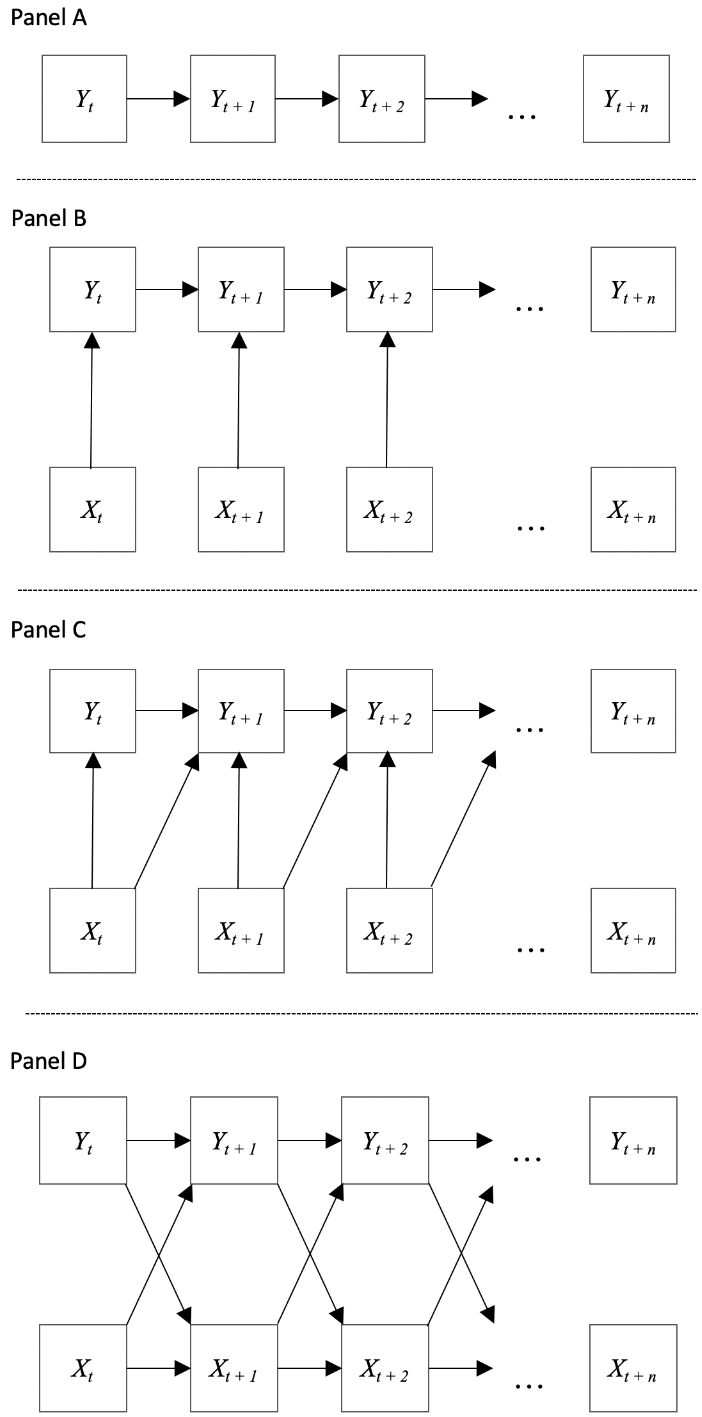

Figure 5. Univariate and bivariate dynamics adapted from DeShon (2012). Panel A shows self-similarity or autoregression in Y across time. Panel B shows concurrent X predicting change in Y. Panel C shows lagged change relationships. Panel D shows reciprocal dynamics between X and Y.