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

Organizational phenomena unfold over time. They are processes that develop, change, and evolve (Pitariu & Ployhart, 2010) that create a sequence of events within a person's stream of experience (Beal, 2015). Moreover, organizations are systems with many connected parts, and systems are inherently dynamic. Studying these systems and processes, therefore, requires paying attention not to static snapshots of behavior (Ilgen & Hulin, 2000), but variables and relationships as they move through 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 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 Ritter, Matthews, Ford, and Henderson (2016) We also find complex hierarchical linear models in many event-sampling studies (e.g., Koopman, Lanaj, & Scott, 2016; Rosen, Koopman, Gabriel, & Johnson, 2016).

The spine of an investigation, finally, is to interpret the model and 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) state that justice behaviors fluctuate day to day and predict 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 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. We bring attention to the span of questions available so that researchers can fully appreciate and take advantage of their data. 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. Moreoever, 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.

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 & Vandenberg, 2010). 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

Level. Trend. Dynamics. These are umbrella research foci, each has its own sub-inferences and models. Each section will basically have two notions: trying to understand the thing itself (one variable). Looking at trend or variability in trend. Then looking at correlates or predictors of the thing: correlates of trend, etc.

Each section will also point to models. But there is nuance. The models have different names, some require stationary, some don't. You need to appreciate that and make sure you are attending to all of its nuance. All we are doing here is pointing you in the direction.

Level

Level inferences concern the value that a variable takes (high or low) across time. If we put a variable on the y-axis and plot its values against time on the x-axis, what value does it take at time t, or what value does it take on average, across time, or what value does it take within certain windows of time?

Level has to do with the condition of a state with respect to its state space. It may have a state space of zero to ten, and its current condition is seven. This is level. It has also been called magnitude.

Is affect high, on average, across time? Is trainee skill low at the beginning of a training session? Are newcomer perceptions of unit climate high at the end of a two-week socialization process?

Level, with respect to a certain point or window in time, is called the intercept in the statistical literature, and we can make an inference about the intercept at any observation point. Sometimes we are interested in the level at the beginning of the study, other times we are interested in the level at the last observation. We can of course also be interested in the average level across time.

When we retain one variable but add multiple units – people or organizations, for example – then we can look at the variability in level. Does everyone have high affect across time? Is there variability in the level of skill among trainees at the beginning of a training session? Now there are multiple points on our graph at each time.

When we measure two or more variables we can examine all of their levels, but we can also correlate them, or regress some variables onto others. Is the level of affect related to the level of helping behaviors across time? That is, when affect is high, are helping behaviors also high or are they low? When team cohesion is low, is team performance low as well, or is it high?

figure 1

Inference List

The level of x at time t is... Picture 1.

There is variability in the level of x at time t. Picture 2.

There is a relationship between the levels of multiple variables. x relates to y across time. Stable vs fluctuating x, time-invariant vs time-varying covariates.

Pictures

One Variable, one unit.

One variable, multiple units

Muliple Variables

When affect is low performance is low. When affect is high performance is high.

Models

Intercept only models: these can be done in HLM or SEM. Time-varying or invariant covariates analysis, these can be done in HLM or SEM. Point to references.

Trend

Trend has to do with the global shape of the 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 tend to go up or down over time? Trend can be thought of as the coarse-grained direction of a trajectory. Also referred to as growth.

Does affect go up or down across our measurement period, or is it relatively stable? Does trainee skill increase over the training session?

These are all linear examples of trend. We can also look at curvilinear things. Do newcomer perceptions of the positive attitude in the workplace increase and then plateua over time? Does the response time of a medical team decrease with each successive case but then remain stable once every crew member has reached their maximum potential to coordinate their actions?

The statistical term for trend is slope, and linear slopes are straight lines whereas curvilinear slopes bend across time.

When we examine more than one unit on a variable we can also look at the variability in trend. Do all trainees develop greater skill across time? Is there variability in the trend of helping behaviors, or counterproductive work behaviors over time?

When we examine two or more variables, we can of course look at trend or slope in each of them. The next step is to then correlate slopes or regress one slope on another. Does the trend in affect relate to the trend in helping behaviors?

Inference List

There is growth in a variable across time. Its slope/trend is positive or negative.

There is variability in the growth of a variable across time. Slopes/trends differ across units.

There are correlates/predictors of trend.

There is a relationship between two growth curves. Two trends are correlated.

Pictures

Models

Growth curves in SEM or HLM. Bivariate growth curves.

Dynamics

Mediation

Discussion

Points to include. 1) Econometrics modeling levels vs. modeling differences. 2) Keep in mind you might see weird stuff in the literature. X at time 1 relates to Z at time 2, which relates to Y at time 3, but none are measured repeatedly across time. This is no good.

References

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

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.

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. doi:10.1177/0149206318780440

Ilgen, D. R., & Hulin, C. L. (2000). *Computational modeling of behavior in organizations: The third scientific discipline.* American Psychological Association.

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.

Koopman, J., Lanaj, K., & Scott, B. A. (2016). Integrating the Bright and Dark Sides of OCB: A Daily Investigation of the Benefits and Costs of Helping Others. *Academy of Management Journal*, *59*(2), 414–435. doi:10.5465/amj.2014.0262

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.

Ployhart, R. E., & Vandenberg, R. J. (2010). Longitudinal research: The theory, design, and analysis of change. *Journal of Management*, *36*(1), 94–120.

Ritter, K.-J., Matthews, R. A., Ford, M. T., & Henderson, A. A. (2016). Understanding role stressors and job satisfaction over time using adaptation theory. *Journal of Applied Psychology*, *101*(12), 1655.

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.

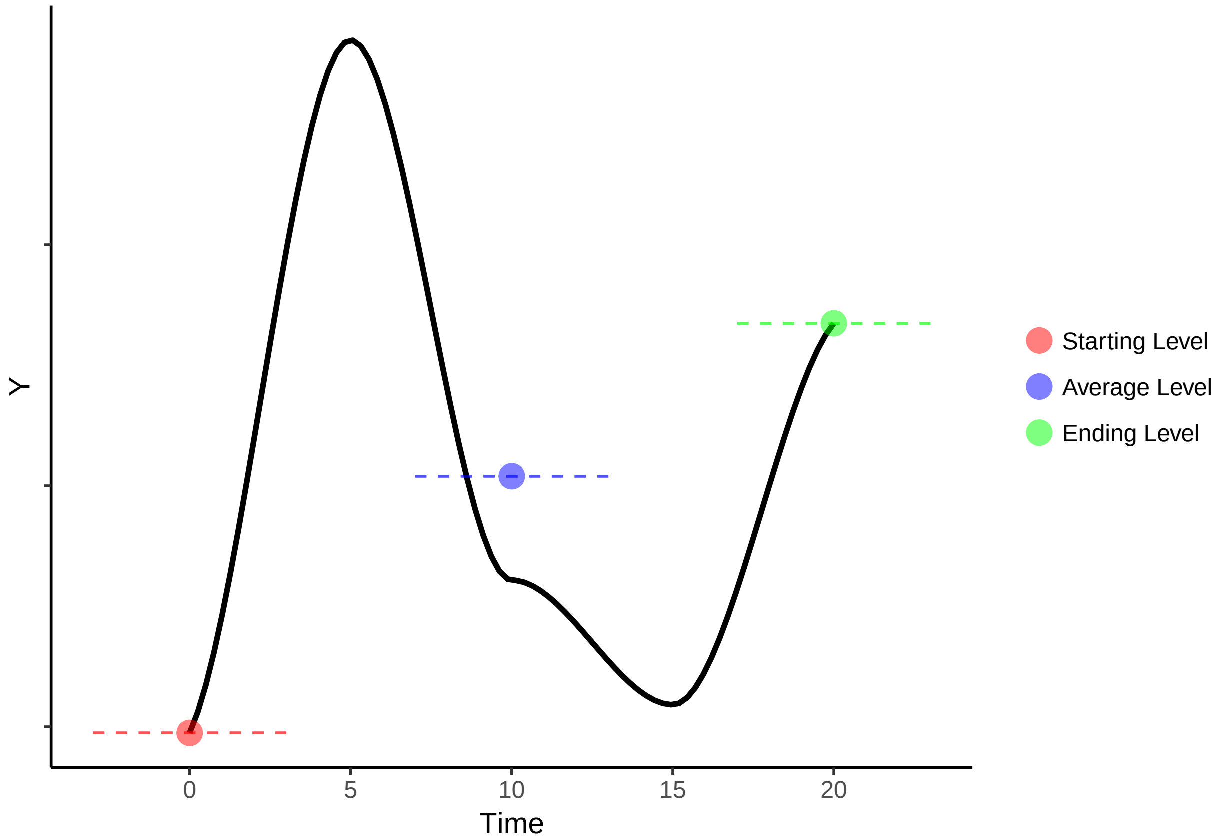

Figure 1. something here

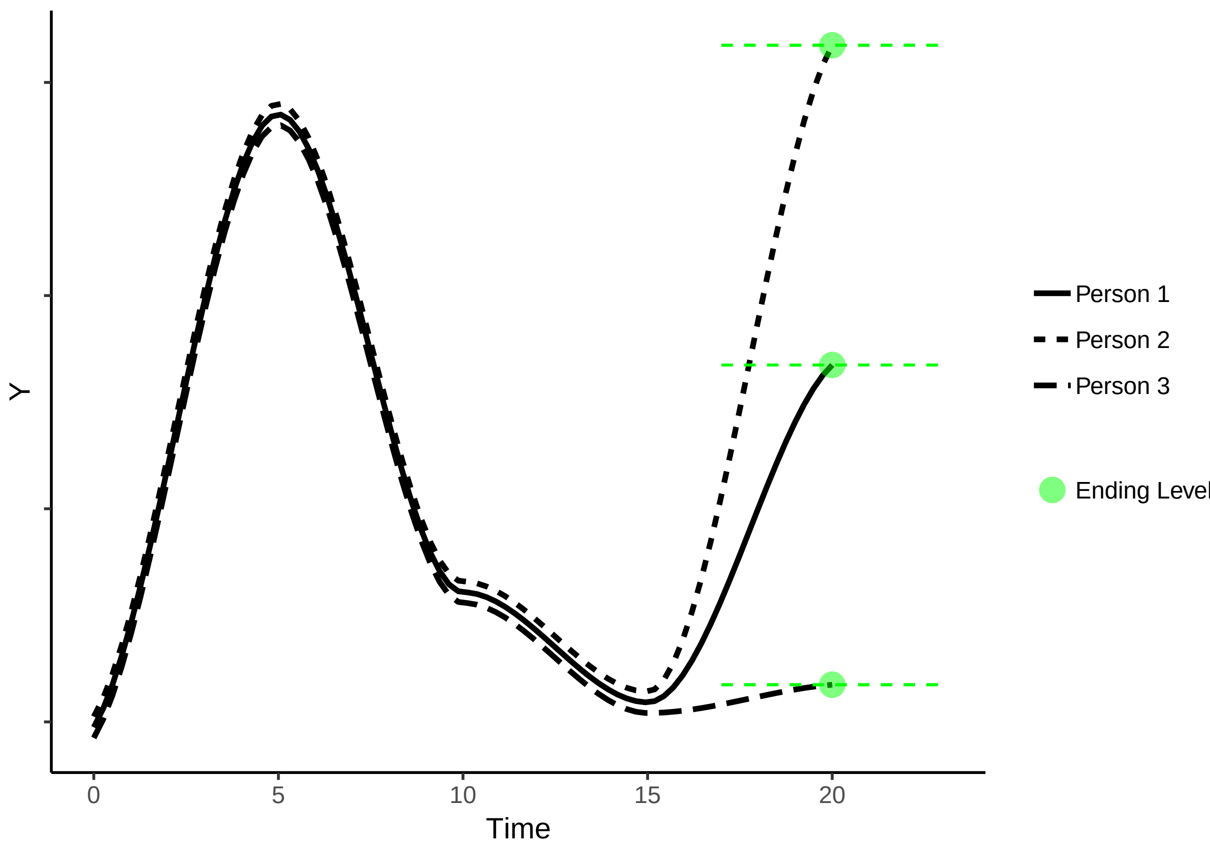

Figure 2. another

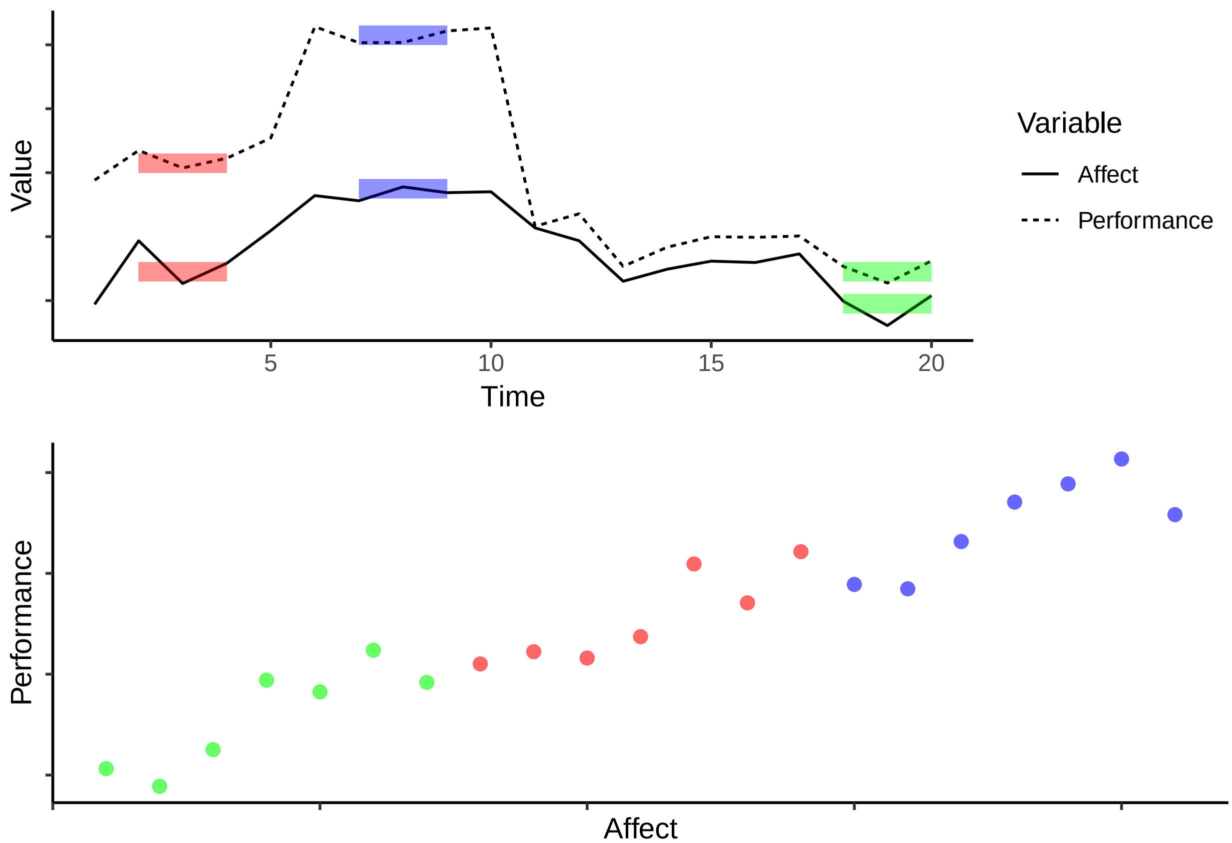

Figure 3. double

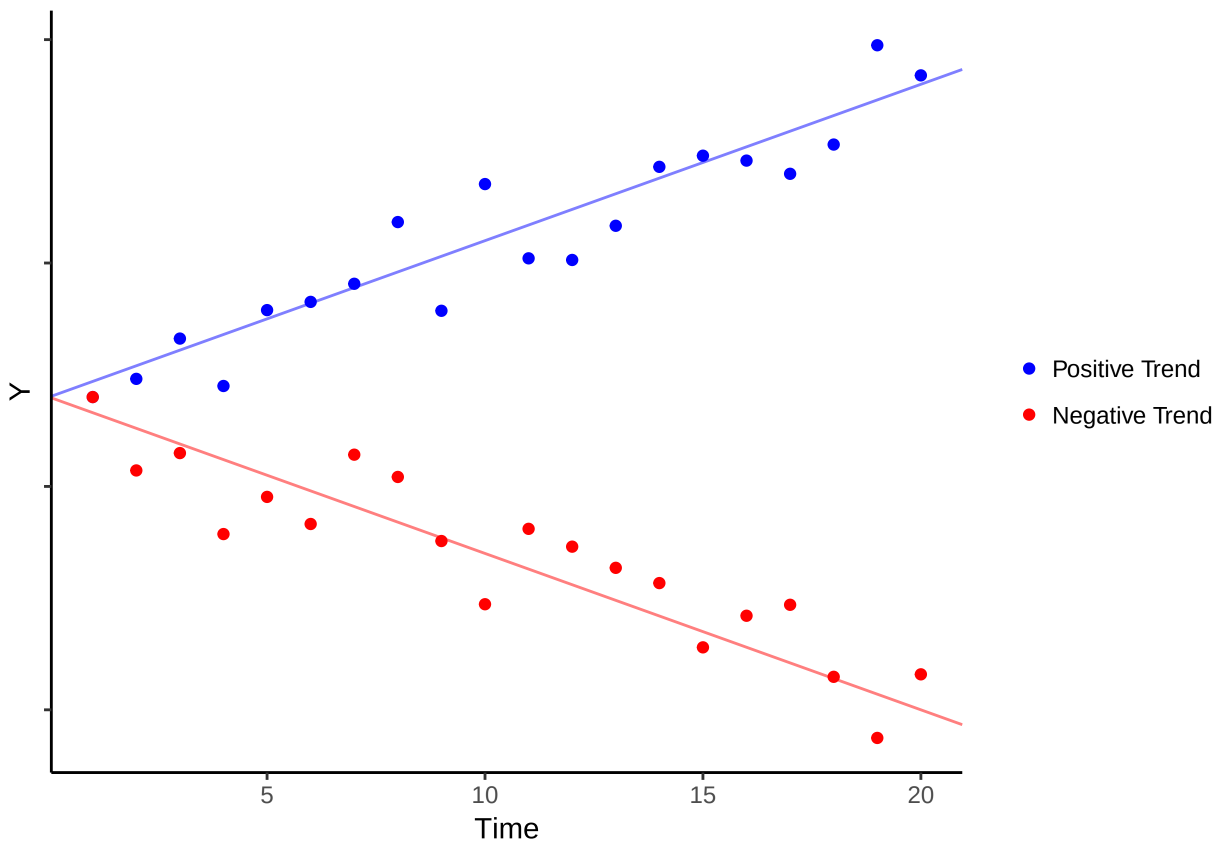

Figure 4. t

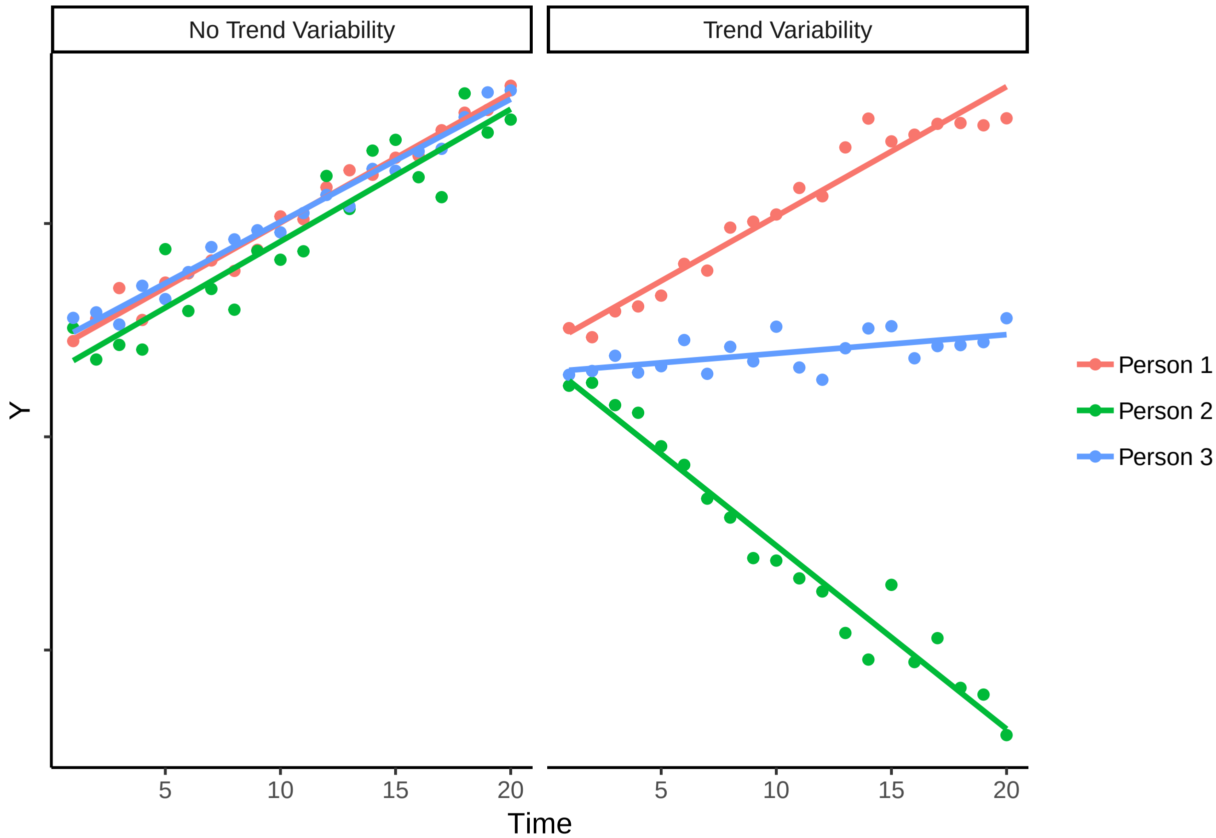
No Trend Variability
Trend Variability
Y
Time
5
10
15
20
5
10
15
20
Person 1
Person 2
Person 3

Figure 5. tr

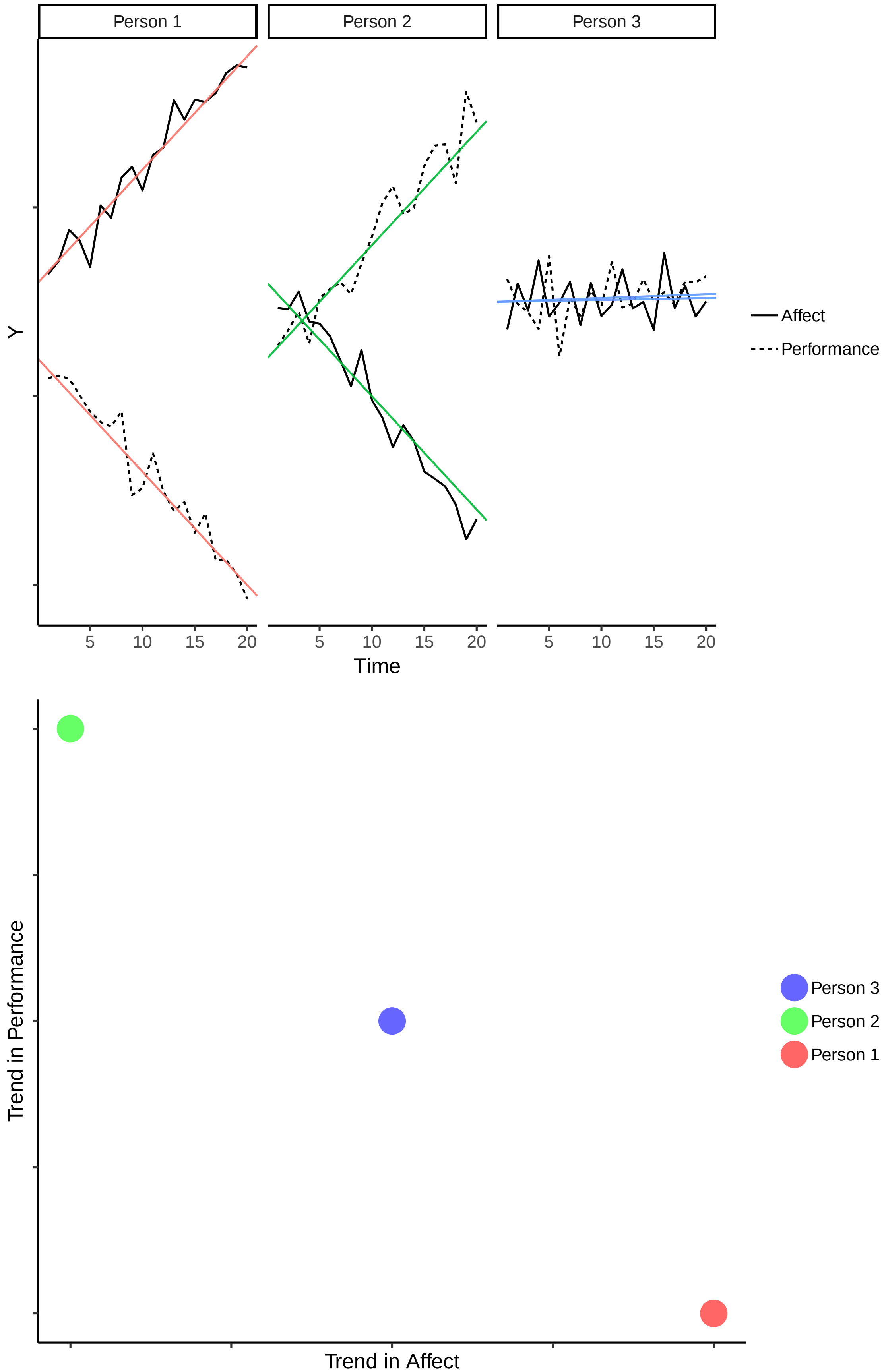

Figure 6. trend to trend