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 that we attend not to static snapshots of behavior (Ilgen & Hulin, 2000), but to 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).

Option A. . . It's hard for newcomers

Although our field is increasingly interested in exploring patterns in longitudinal data, process-oriented methods are still relatively new to our field and newcomers without much longitudinal modeling training may be unfamiliar with the variety of questions they can ask. Consider a few recent longitudinal studies that all pose different questions. Jones et al. (2016) ask about trend: they want to determine if the trajectories among certain variables increase or decrease over time. Johnson, Lanaj, and Barnes (2014) about change: they are interested in how changes in one variable relate to changes in another across time. Hardy, Day, and Steele (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.

There are then complex statistical models that researchers evoke to examine their 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). Again, researchers without much longitudinal modeling training may not know when to apply each model – or which model is appropriate for a given question.

Finally, the spine of an investigation 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.

In this paper we discuss the common inferences that researchers in our field make when they apply a model to longitudinal data. 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. Moreover, 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 potential models 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.

Option B...Just highlighting the literature

There are many interesting questions researchers can explore with longitudinal data. Consider a few recent longitudinal studies that all pose different questions. 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.

There are then complex statistical models that researchers evoke to examine their 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 et al. (2016). We also find complex hierarchical linear models in many event-sampling studies (e.g., Koopman et al., 2016; Rosen et al., 2016).

Finally, the spine of an investigation 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.

In this paper we discuss the common inferences that researchers in our field make when they apply a model to longitudinal data. 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. Moreover, 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 potential models 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 each i at each t.

Framework

Level. Trend. Dynamics. These are umbrella research foci, each has its own sub-inferences and models. Each section will have several inferences but they all gather into two basic notions: 1) trying to understand the thing itself and variability about the thing itself across units, and 2) correlates or predictors of the thing.

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.

Something about our graphs. We are going to talk about the inferences and provide researchers with phrases that they can use to present hypotheses. But we are also going to show what the inferences look like in data. We feel that graphing the inferences with respect to data is more informative than plotting your usual box and arrow diagrams (and for some of the inferences boxes and arrows would not work). There is a caveat, however, that we want to make sure our readers are aware of. Data are always messy. You will almost never find data in the style that we show here, where the inferences expose themselves simply by plotting. We are using these "data plots" as heuristics, and in our opinion they serve the reader better than box and arrow heuristics – even though our plots carry their own limitations. Again, know that field data will always be messy.

Level

Is employee emotional exhaustion, on average, high across the study? Is trainee skill low at the beginning of a training session? What value are newcomer perceptions of unit climate at the end of a two-week socialization process? These are questions about level, or the specific value of a variable.

Levels either describe the variable at one moment or averaged across a span of time. That is, if we put a variable on the y axis and plot its values against time on the x axis, we can explore the value that it takes at time t, or the value that it takes on average across any span of t.

Figure 1 demonstrates this idea graphically. A variable is plotted across time, and the color labels indicate levels – the red and green describe the variable at a specific moment while the purple, average level, describes it across a window.

Insert Figure 1 about here

Our first level inference, therefore, concerns the value of a variable at a specific time or averaged across a window of time.

Inference 1: What is the level of x at time t, or across a span of t?

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?

We demonstrate this idea in figure 2, where each unit (person) has a similar trajectory but different levels at the last time point.

Insert Figure 2 about here

The second level inference, therefore, is about level variability across units.

> **Inference 2:** Across a span of t or at a specific t there is variability in the level of x.

Inferences one and two concern a single variable, but they can of course be iterated across any or all observed variables in the study. For example, we might discover that affect and performance have high average levels across time, but that affect has greater level variability. Or we might learn that affect has a low initial level whereas performance is initially high. What we are doing here is making descriptive comparisons between the level of one variable and the level of another. We can also produce a quantitative statement about the extent to which levels are related.

Correlating levels provides us with that quantitative statement. A large positive correlation between the initial levels of affect and performance would mean that people with greater initial levels of affect also tend to have greater intial performance, and people with lower initial affect also tend to have lower initial performance.

Figure 3 demonstrates this graphically. Paragraph about the graph.

> **Inference 3:** There is a correlation between the level of x and the level of y at t.

The final level inference is horizontal. Rather than correlating values from a single moment or a single averaged moment, we correlate values across time. For example, we might ask if affect is related to performance across time; i.e., when affect is high is performance also high, and when affect is low is performance also low?

This inference sounds similar to the one just presented, but their difference is important. With inference three we ask about affect and performance at t or at an averaged window of t – we examine, for example, how averaged values across time relate to averaged values across time. Here, we retain all of the information and examine the relationship between affect and performance across all t.

Figure 4 shows this inference graphically. The top panel plots affect and performance trajectories across time. The colored squares represent levels at different points in time. The green squares highlight low values of both variables, the blue high values, and the red middle values. The bottom panel shows how those respective values map onto a graph that describes the relationship between affect and performance across time. Notice that there appears to be a positive relationship in the scatterplot, which tells us that when affect is high performance also tends to be high, and vice versa.

Insert Figure 4 about here

Inference 4: There is a relationship between x and y across time.

Level Inference Table

The inference table below provides examples of each level inference. Inference one is about level itself – a single value that describes the variable at one time or averaged across time. Inference two is about variability across units in level. Inference three takes the level in one variable and asks whether it tends to co-occur with the level in another. Think of inference three as creating a latent level variable at a single moment and asking how it relates to another latent variable from a single moment. Inference four, finally, is about the relationship between raw values across time.

Inference	Examples
1	Burnout is high at the last time point. Performance is low, on average, across time.
2	Average affect across time differs across people (units). There is variability in the initial level of turnover across organizations.
3	People with greater initial health status also have greater initial happiness. People with high performance on average across time have lower anxiety on average across time.
4	Affect relates to performance across time. Helping behaviors predict depletion across time.

Models

Level is called intercept in the statistical modeling literature. Typically the mean estimate tells you about the level, and the variance estimate tells you about the variability across units. Intercept only models in HLM or SEM. Time-varying or invariant covariates analyses. Point to references.

Trend

Does affect go up or down across time, or is it relatively stable? Does trainee skill 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 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? It can be thought of as the coarse-grained direction of a trajectory.

Figure 5 demonstrates trend, where the red line shows positive, increasing trend, the blue line shows negative, decreasing trend, and the green line shows a curvilinear trajectory (ADD GREEN LINE). Keep in mind that curvilinear and linear trajectories are both *linear in parameters* and should not be confused with non-linear systems.

Insert Figure 5 about here

Our first trend inference, therefore, concerns the shape of the trajectory.

> **Inference 1:** There is positive/negative/curvilinear trend in a variable across time.

As with the level inferences, when we add more units we can examine trend variability.

Do all trainees develop greater skill across time? Is there variability in the trend of helping behaviors, or counterproductive work behaviors over time?

Figure 6 shows differences in trend variability. In the first panel all units (people) show the same positive trend, whereas everyone in the second panel shows different trend: person one's data appear to increase over time, person two's data remain flat, and person three's data decrease over time. With greater variability there is less consistency in trend across units.

Insert Figure 6 about here

> **Inference 2:** There is variability in the trend of a variable across time; trend differs across units.

Inferences one and two are about one variable, but they can also be iterated across all observed variables. For example, we might discover that affect and performance trends both decrease, but there is greater variability across units 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, but this time we are focused on trends rather than levels. 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 7 shows the inuition behind the inference with a set of graphs. On the top panel, affect and performance are plotted across time for three individuals. For individual one, affect goes up while performance goes down, this pattern is reversed for person two, and person three reports fluctuating affect and performance but with no trend. The bottom panel then maps those pairings onto a figure that shows the relationship between the affect and performance trend. For example, person one has a positive affect trend and a negative performance trend, so their point on the bottom panel goes on the bottom right, whereas person two has the opposite pattern and therefore is placed on the top left (where performance trend is positive and affect trend is negative). Producing this bottom scatter plot tells us that the relationship between affect and performance trend is negative. That is, people with a positive affect trend usually have a negative performance trend, people with a negative affect trend are more likely to have a positive performance trend, and people with no affect trend usually have no performance trend.

Insert Figure 7 about here

> **Inference 3:** There are correlated trends. There is a relationship between two trends.

The final trend inference is about identifying covariates or predictors of trend. Do helping behaviors predict the trend in depletion? Does the trend in unit climate covary with perceptions of leader quality?

Notice the difference between this inference and inference three. Inference three asks how one trend is related to another, whereas this inference asks how raw values relate to or predict the trend in a different variable.

Figure **??** (no figure yet). Paragraph about graph. **Not sure how to graph. Slope is a single term, like 0.5. It is being predicted by the raw values of X across**

time. . .

Inference 4: There are correlates/predictors of trend.

Trend Inference Table

The inference table below provides examples of each trend inference. Inference one is about the general direction or shape of a trajectory across time. Inference two is about variability in that shape across units. Inference three takes the trend in one variable and asks whether it tends to co-occur with trend in another. Inference four, finally, is about the relationship between trend in one variable and the raw values of one or more predictors.

Inference	Examples
1	Burnout decreases over time. Performance increases over time.
2	Affect trends differ across people (units). There is variability in turnover trends across organizations.
3	People with a positive health status trend have a positive happiness trend. People with a positive performance trend have a negative anxiety trend.
4	Affect relates to the performance trend across time. Helping behaviors predict depletion trends.

We want to close this section with a note on phrasing. The inferences we explored in this section have to do with correlating trends, where a statement like "affect and performance trends covary, 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 dynamic statement about how trajectories move across time. That is, the phrase "when affect decreases performance goes up" could refer to 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 have shown here.

Models

Trends are called slope estimates in the statistical modeling literature. They are also referred to as growth. Mean estimates of slopes, or trends, or growth will tell you about trend, whereas the variance estimates will tell you about variability across units. Growth curves in SEM or HLM. Bivariate growth curves.

Dynamics

Dynamics refers to a specific branch of mathematics, but the term is used in different ways throughout our literature. It is used informally to mean "change", "fluctuating," "volatile," "longitudinal," or "over time" (among others), whereas formal definitions in our literature are presented within certain contexts. Wang defines 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). Vancouver states 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). Finally, Monge suggests that in dynamic *analyses*, "it is essential to know how variables depend upon their own past history" (p. 409).

The crucial notion to take from dynamics, then, is memory. When the past matters, and future states are constrained by where they were at prior points in time, dynamics are at play. In this section, we unpack a variety of inferences that are couched in this idea.

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?

Figure (no graph yet) shows the concept graphically. Paragraph about the graph.

The statistical term used to describe self-similarity is autoregression, and we use it to put a label on this first inference.

Inference 1: There is autoregression in x.

Inference one was of course about a single variable. When we apply the notion of memory to multivariate systems we move away from asking how variables relate to one another at t, or at an average window of t, or across t, and instead ask how variables relate to one another at different lags. Does affect predict subsequent performance? Do prior counterproductive work behaviors relate to current incivility? When goal discrepancy is large is effort at the subsequent time point high? When prior depletion is low, is current emotional exhaustion high?

We can capture this second inference by relating current values on one variable to future values on another. Equivalently, we can relate prior values on one variable to current values on another. Figure **??** (no figure yet) plots…Paragraph about the graph.

Relating current to future (or prior to current) values from one variable to another is called a “cross lag” relationship.

> **Inference 2:** There is a cross-lag relationship, where one variable relates to another at a different point in time.

Inference two tells us whether the patterns in one variable co-occur with the patterns in another at a subsequent time point. Across time, when affect is low is subsequent performance also low? A related question is as follows: across time, when affect is low does performance increase or decrease? This second question is about change. How does one variable relate to the change in another?

When goal discrepancy is large does effort increase or decrease? When unit climate is low do perceptions of the leader change? When performance is high does self efficacy go up or down?

All of these questions are about change, but notice that change can be construed across different lags. Change from what? Baseline? The prior time point? The last three time points? Typically change is construed with respect to the last time point. When affect is low, does performance from the last to the current time point increase or decrease? How does effort change from the prior to the current time point when goal discrepancy is high?

Figure **??** demonstrates these ideas. Paragraph about the graph.

It is typical to think of change from the prior to the current time point, but researchers are free to move it as they please. Here are the two final inferences that capture change in different locations.

> **Inference 3:** There is a change relationship, where one variable relates to the change in another.

> **Inference 4:** There is a cross-lag relationship of change, where one variable relates to the change of another at a different point in time.

Dynamics Inference Table

Again, we provide an inference table below – this time with respect to dynamic inferences. Inference one is about autoregression, or memory in a single variable. Inference two asks how a variable at one time co-occurs with another at a different time. Inferences three and four focus on change: when one variable is high or low, does it relate to the change (an increase or decrease) in the values of another variable?

Inference	Examples
1	Burnout demonstrates self-similarity across time.
	Performance relates to subsequent performance.
2	Affect predicts subsequent counterproductive work behaviors.
	Turnover relates to subsequent firm performance.
3	Positive health status relates to change in happiness.
	Anxiety relates to change in performance.
4	Affect relates to subsequent change in performance.
	Helping behaviors predict subsequent depletion changes.

Models

Our literature has a history with difference scores and partialling. We debated difference scores so we have converged to partialling models. Typically we create a model with prior values of the response variable as a predictor.

Mediation

Systems of variables. Reciprocal relationships.

Somehow say that mediation kind of goes away when you get into reciprocal dynamics – without making it sound like we don't care about 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. We opened with "we couch ourselves by only discussing studies where constructs were measured on each i at each t." Sometimes this doesn't happen…

References

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.

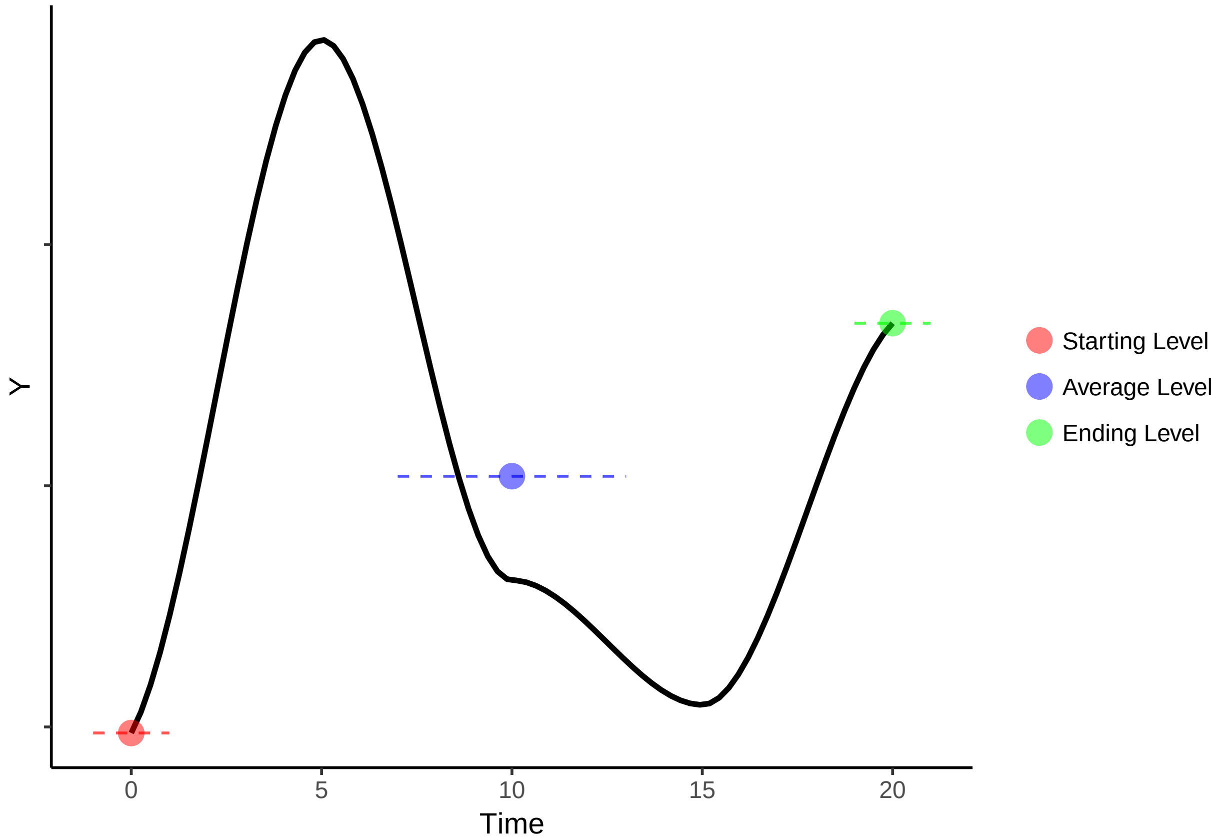

Figure 1. Level examples

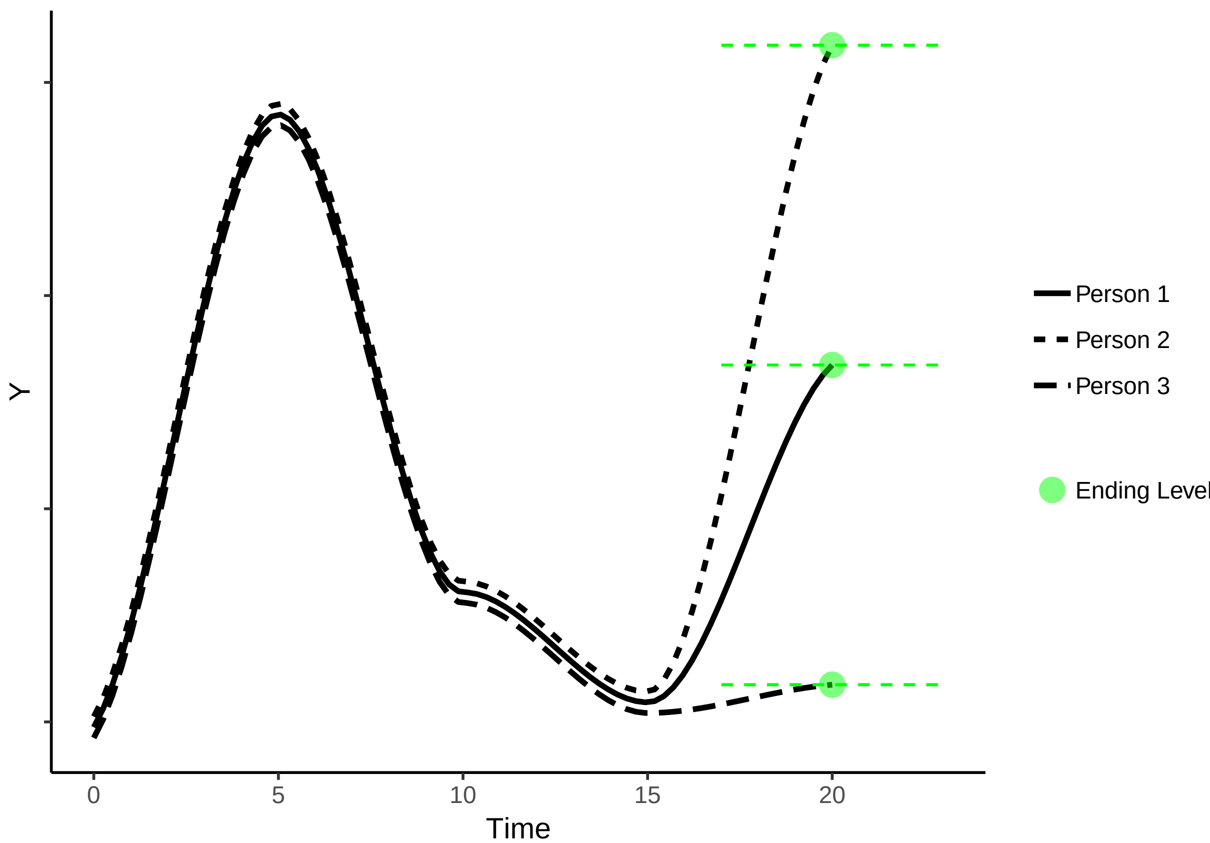

Figure 2. Trajectories with variability in ending level across units

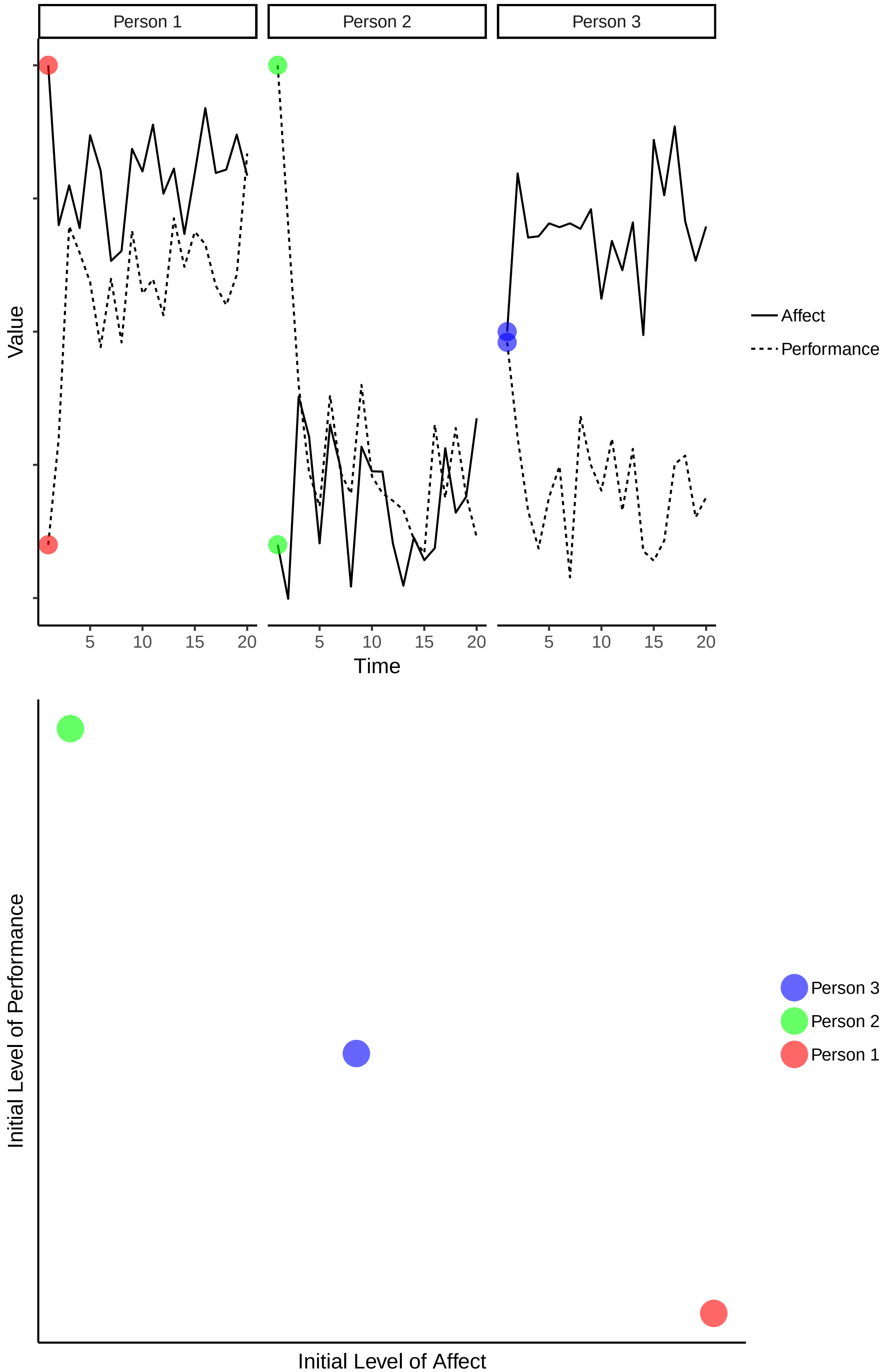

Figure 3. Correlating starting levels, or relating initial affect to initial performance

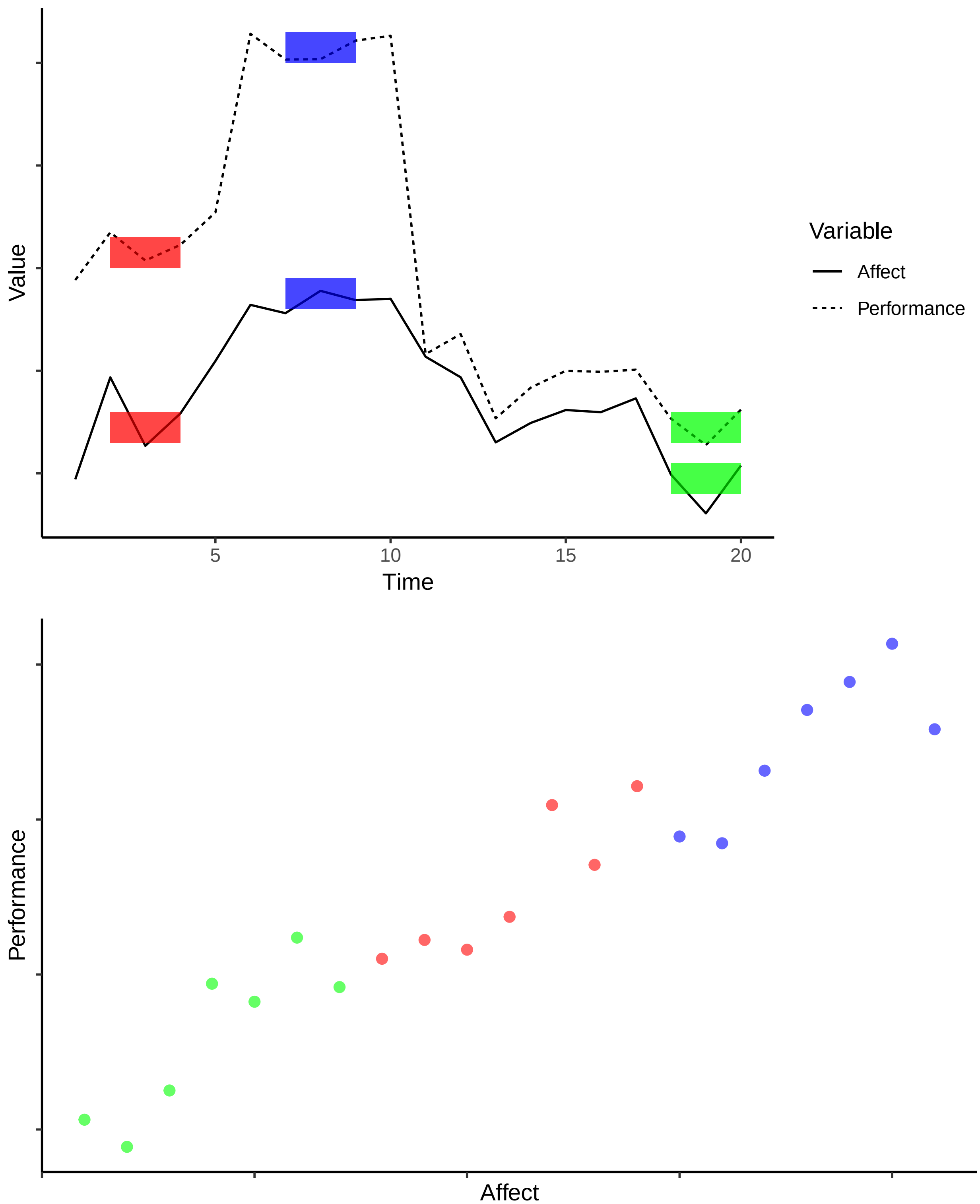

Figure 4. Relating affect to performance levels

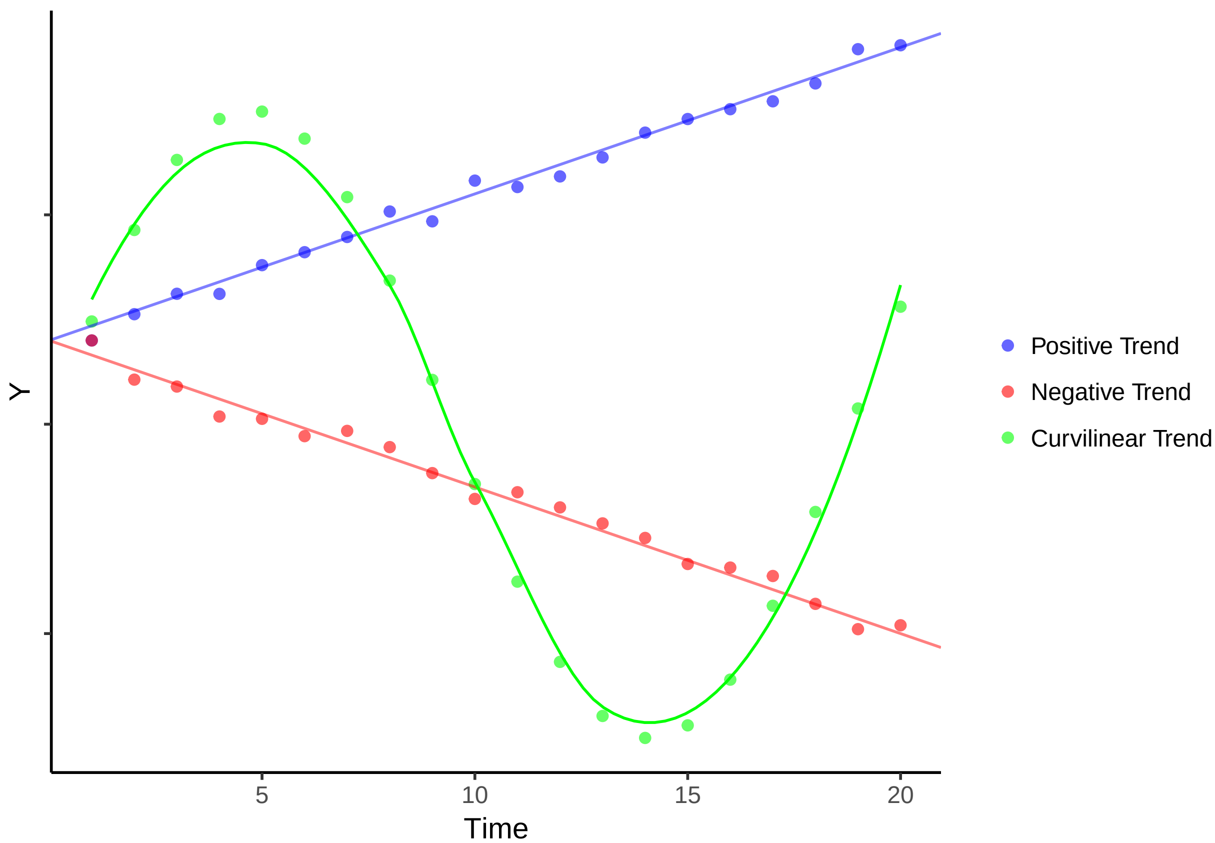

Figure 5. Trend across time

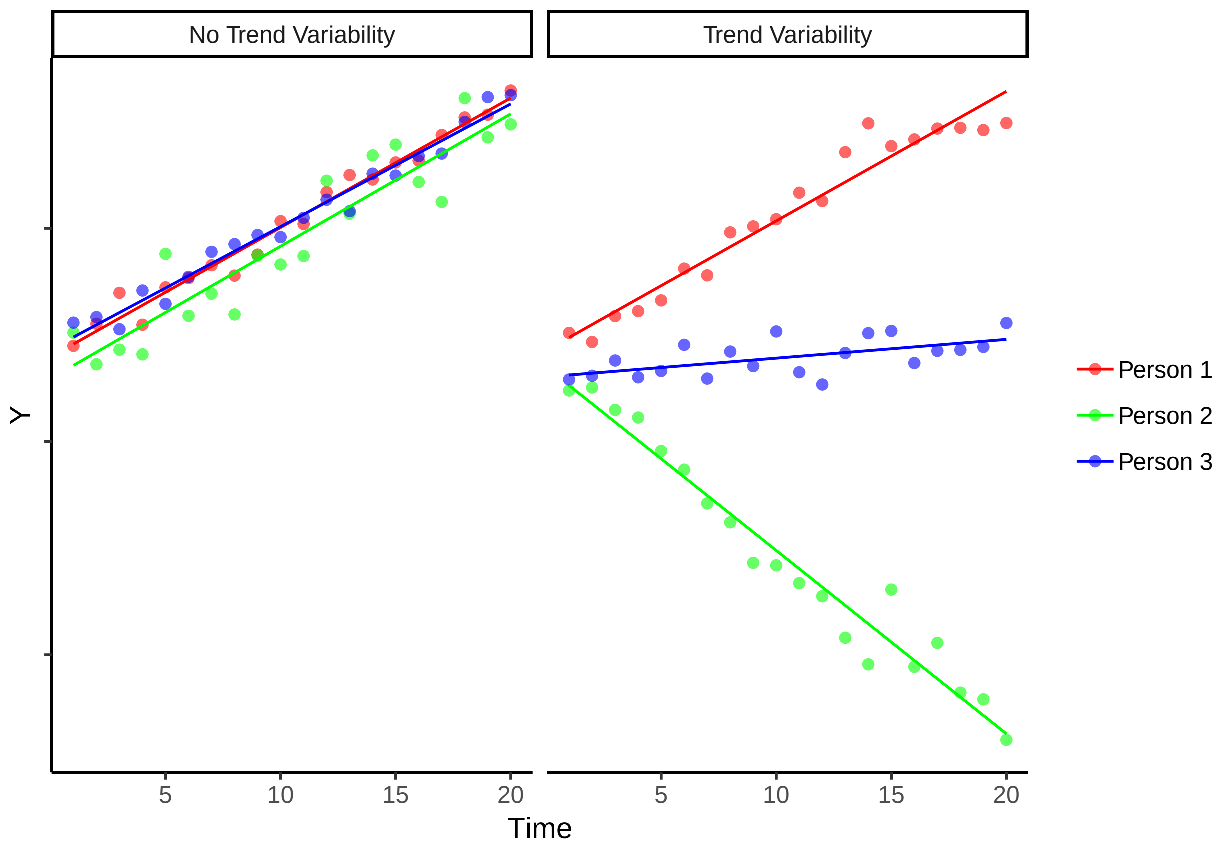

Figure 6. Differences in trend variability across units

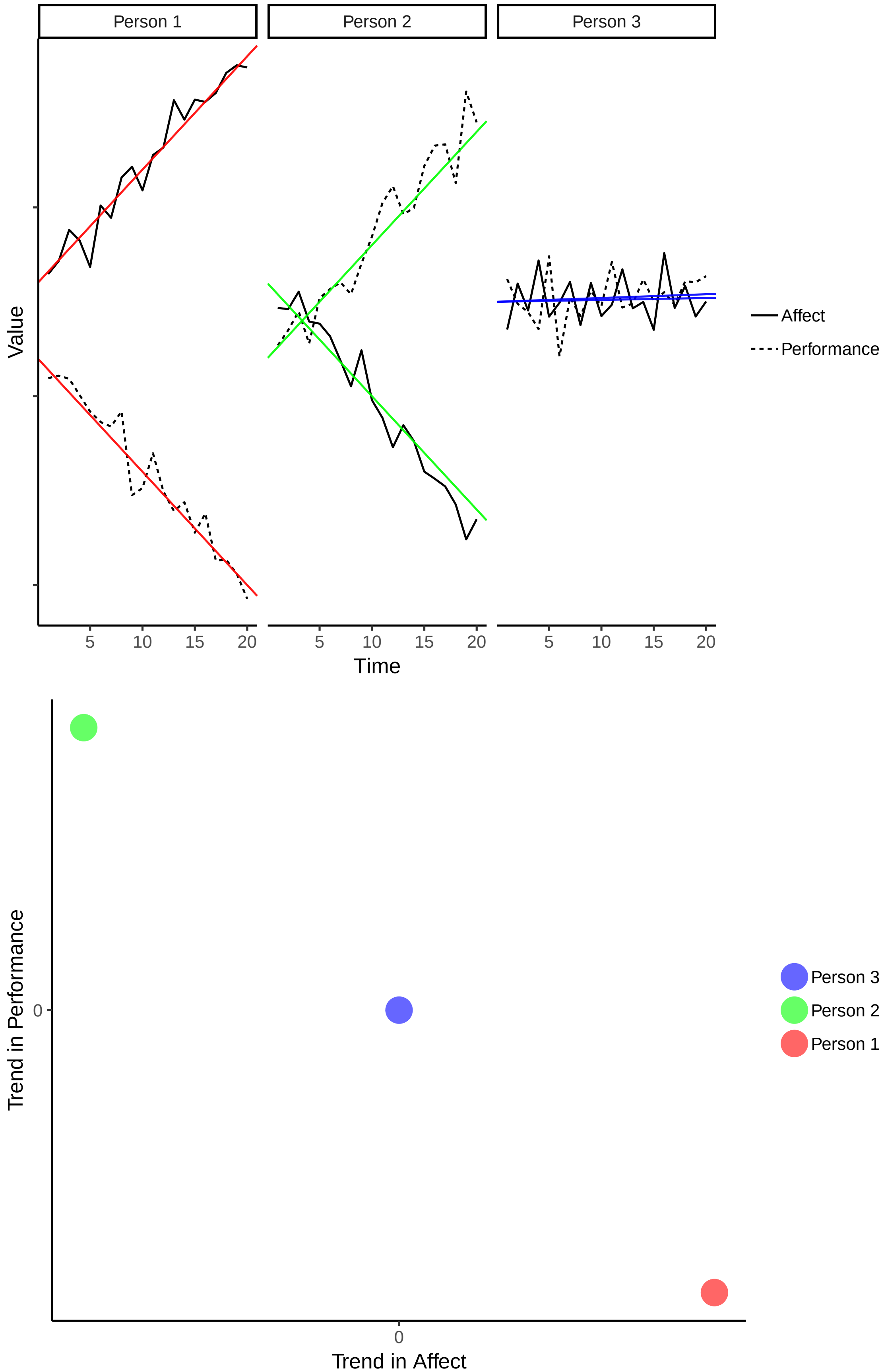

Figure 7. Correlating slopes, or relating the affect to performance trend