Opening the Black Box: A Multilevel Framework for Studying Group Processes

Jonas W. B. Lang

Ghent University

Paul D. Bliese

University of South Carolina

Amy B. Adler

Walter Reed Army Institute of Research

Lang, J. W. B., Bliese, P. D., & Adler, A. B. (in press). Opening the black box: A multilevel framework for studying group processes. *Advances in Methods and Practices in Psychological Science*.

Author Note

Disclaimer: Material has been reviewed by the Walter Reed Army Institute of Research. There is no objection to its presentation and/or publication. The opinions or assertions contained herein are the private views of the author, and are not to be construed as official, or as reflecting true views of the Department of the Army or the Department of Defense. The investigators have adhered to the policies for protection of human subjects as prescribed in AR 70–25.

Author Contributions: JWBL re-analyzed the Jury data and PDB re-analyzed the Army data. JWBL wrote a first draft of the manuscript with the exception of the sections on the Army data that PDB and ABA wrote. PDB and ABA critically edited the first draft and rewrote several sections. All authors approved the final submitted version of the manuscript.

Abstract

Over time, groups can change in at least two important ways. First, they can display different trajectories (e.g., increases or decreases) on constructs of interest. Second, the configurations of group member responses within groups can change and group members can become more or less similar to each other. Psychologists have historically been interested in understanding changes in groups over time; however, there is currently no comprehensive quantitative framework to study and model group processes over time. We present a multilevel framework for studying group processes—the multilevel group process framework (MGPF). The MGPF builds on a statistical approach developed to capture whether individual group members develop shared climates over time but extends the core ideas in two important ways. First, we describe how researchers can gain insights into group phenomena such as group leniency, group learning, group think, group extremity, group forming, group freezing, and group adjourning through modeling latent mean-level and consensus change processes. Second, we present a sequence of model testing steps that enable researchers to systematically study and contrast different group process. We describe how the MGPF can lead to novel research questions, and illustrate its use in two example datasets.

Opening the Black Box: A Multilevel Framework for Studying Group Processes

Social psychologists (Lewin, 1947; Lewin, Station, & White, 1939), industrial-organizational (I-O) psychologists (Ashforth, 1985), clinical psychologists (Corsini & Rosenberg, 1955) and other fields of psychology have long been interested in understanding changes in groups over time to gain insights into how groups function. Numerous laboratory-based studies have found that groups act as powerful change agents: studies starting in the 1930s demonstrated that individuals often react to the presence of others, changing their perceptions, opinions, and behaviors (Sherif, 1935). Similarly, in applied settings groups often have important effects such as motivating employees (Mayo, 1933) and enhancing decision-making (Kozlowski & Ilgen, 2006).

One obstacle with studying group processes in both laboratory and applied settings is the lack of a comprehensive quantitative framework to study "group processes" – a term we use to refer to changes groups experience over time. Quantitatively capturing group change is challenging for at least two reasons. First, change over time simultaneously occurs for individuals and groups, so group processes are inherently multi-level with a three-level structure (measurement occasions nested in individuals, and individuals nested in groups). Second, change takes at least two different forms. Specifically, perceptions and behaviors in groups include a consensus element (i.e., to what degree do group members share perceptions or converge in their behavior?), and a direction element (i.e., do perceptions and behaviors increase or decrease over time?).

Both consensus and direction effects are dynamic. Nonetheless, researchers in different domains have noted that psychologists know surprisingly little about how group effects dynamically evolve over time (Cronin & Weingart, 2011; Humphrey & Aime, 2014; Kozlowski, 2017; Kozlowski & Bell, 2013; Mohammed, Hamilton, & Lim, 2009). To date, group processes have been examined using qualitative data (Gehman, Trevino, & Garud, 2013), event occurrence data (e.g., Chiu & Lehmann-Willenbrock, 2016), network data (e.g., Huisman & Snijders, 2003), and simulated data (e.g., Kozlowski, Chao, Grand, Braun, & Kuljanin, 2013). Unfortunately, researchers who collect common forms of multilevel data (individual responses over time, nested within groups) have had challenges studying questions such as “Do interactions in groups change individual perceptions and opinions?” or “Do group interactions lead to the emergence of consensus or to increased dissensus over time?” or “What types of individual and environmental predictors explain the absence or presence of consensus forming and shifting in group perceptions and opinions over time?”

We present a multilevel framework for studying group processes referred to as the multilevel group process framework (MGPF). The MGPF is designed to allow researchers to jointly model consensus and directional change over time at the individual and group level. Our paper builds on earlier statistical work (Hedeker & Mermelstein, 2007; Hedeker, Mermelstein, & Demirtas, 2012; Kim & Seltzer, 2011; Rast, MacDonald, & Hofer, 2012; Raudenbush & Bryk, 1987) and a methodological approach (Lang & Bliese, 2018; Lang, Bliese, & de Voogt, 2018) recently developed in the I–O psychology literature. The approach developed for I-O focuses on how individual members of a group develop consensus over time in attitudes or perceptions—a process referred to as climate emergence (Bliese, 2000; Kozlowski & Chao, 2012; Kozlowski et al., 2013). The purpose of the present article is to extend earlier approaches for studying climate emergence (Lang & Bliese, 2018; Lang et al., 2018) into a more general framework.

Our article makes two specific contributions. First, we conceptually describe how the MGPF can track group processes in longitudinal group data by including consensus change (reflecting earlier work on emergence) along with latent group-mean (slope) change and group slope variability. We then conceptually link these specific elements of group process to a set of relevant group phenomena including group leniency, group learning, group think, group extremity, group forming, group freezing, or group adjourning. These types of group phenomena play a central role in group research and theory, so the ability to test these types of ideas using longitudinal multilevel models has the potential to significantly advance group research. Second, we describe a sequence of model testing steps that allow researchers to systematically contrast and study different types of group process. Finally, we empirically demonstrate the use of the MGPF in (a) a freely available dataset on group decision making in mock juries, and (b) a large field dataset of new recruits in the US Army.

The Multilevel Group Process Framework

Studies of group processes fundamentally involve integrating information from individuals, groups, and time. The multilevel group process framework (MGPF) addresses this analytic challenge utilizing a three-tier multilevel structure. More specifically, the MGPF is a combination of a longitudinal growth model with measurement occasions nested within persons and a multilevel model with persons nested in groups.

Using the MGPF, researchers can examine both (a) changes in the direction (or latent means) of perceptions or behavior over time, and (b) the degree to which groups display consensus (or divergence) and changes in consensus over time. Focusing on change in direction is fairly common in multilevel research, but several influential theories and research questions about group processes also revolve around the idea that consensus changes. For instance, the idea of freezing or unfreezing group climates (Lewin, 1947) or the notion that groups go through several stages of development (Gersick, 1988; Tuckman & Jensen, 1977) represent theories about consensus or divergence development. Ultimately, though, the joint modeling of both aspects is important because both types of change are often present and the combination of both provide a rich foundation for theory development and testing.

By jointly and systematically modeling both directional change and consensus change, the MGPF builds on earlier methodological approaches that has focused on these two aspects of change in isolation. In practice, it is often important to model both directional change and consensus jointly to understand group processes. As a simple example of how directional change and consensus effects interact consider the intraclass correlation, type 1 (ICC1; Bliese, 2000). The ICC1 is frequently used in the organizational literature to study consensus; however, ICC1 values are affected by both between and within group variance. Therefore, changes in the ICC1 over time can either result from an increase in consensus within the groups or from changes in the between group variance associated with mean-level change. Lang and Bliese (2018) and Lang, Bliese, and de Voogt (2018) demonstrate how examining ICC1 values over time can lead to misleading conclusions regarding the presence or absence of consensus emergence or divergence.

Likewise, a fair amount of research examines mean-level change without considering consensus. Implicitly, this approach assumes that levels of consensus remain unchanged over time. When consensus change is actually present, however, the omission of consensus effects in statistical models can lead to potential bias because multilevel models assume that the residual variance is homoscedastic (does not change over time) (Bliese & Ployhart, 2002; Pinheiro & Bates, 2000; Singer & Willett, 2003). From a theory development and testing perspective, though, failing to jointly model and understand consensus effects only partially captures the dynamic nature of how group processes unfold over time.

Model Specification

In the three-tier multilevel modeling framework used by the MGPF, changes in direction are captured by examining familiar fixed-effects terms associated with time (e.g., Bliese & Ployhart, 2002; Bryk & Raudenbush, 1987; Singer & Willett, 2003). Consensus emergence or divergence increases are captured by explicitly modeling changes in heteroscedasticity in the error terms over time, and by substantively interpreting changes in heteroscedasticity. Incorporating heteroscedasticity (changes in error variance) over time is fairly common practice, but has commonly been framed as a violation of one of the assumptions of the multilevel growth model (Bliese & Ployhart, 2002; Singer & Willett, 2003). In contrast, the MGPF considers heteroscedasticity as a way to gain meaningful substantive insights into changes in shared group perceptions or behaviors and the emergence of group constructs. Building on the earlier work focused on consensus emergence (Lang & Bliese, 2018), the MGPF relies on the following multilevel model specification.

$$
\begin{aligned}
&\text{Level-1 (observations):} && Y_{tij} = \pi_{0ij} + \pi_{1ij}\mathrm{TIME}_t + e_{tij} \\
&\text{Level-2 (persons):} && \pi_{0ij} = \beta_{00j} + r_{0ij} \\
& && \pi_{1ij} = \beta_{10j} \\
&\text{Level-3 (groups):} && \beta_{00j} = \gamma_{000} + u_{00j} \\
& && \beta_{10j} = \gamma_{100} + u_{10j} \\
& && e_{tij} \overset{iid}{\sim} N(0, \sigma^2\exp[2\delta_1\mathrm{TIME}_t]) \\
& && r_{0ij} \overset{iid}{\sim} N(0, \tau)
\end{aligned}
$$

$$
\begin{pmatrix} u_{00j} \\ u_{10j} \end{pmatrix} \overset{iid}{\sim} N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{pmatrix} \upsilon_{00} & \upsilon_{10} \\ \upsilon_{01} & \upsilon_{11} \end{pmatrix}\right)
$$

This specification explains the response Y_{tij} at timepoint t of person i in group j. TIME is typically coded 0 at the start of the study and increases by 1 with occasion.

The Start of the Study

Before researchers examine group processes, a first step is to examine the characteristics of the data at timepoint 1. The model specification used for the MGFP provides several informative insights for so doing. Specifically, the model specification includes a common intercept γ_{000} that captures the latent mean at the start of the study. Persons differ from this common intercept with r_{0ij} and groups with the group-specific deviation u_{00j}. The variance of the person deviations r_{0ij} is labeled τ, and the variance of the group deviations is labeled υ_{00}. The common intercept and the group and person deviations capture the direction of the group and individual perception or behavior at the start of the study. When the variances of the group deviations (υ_{00}) and person deviations (τ) are considerable, substantial multilevel heterogeneity exists. Researchers commonly assume that the amount of variance at the group-level (υ_{00}) at the beginning of the study will be relatively small particularly when focusing on newly formed groups that have not yet developed a distinct climate. In contrast, the amount of variance at the person-level (τ) is often relatively large because individuals tend to respond consistently over time (Bodner & Bliese, 2017).

Change and Variability in Perception or Behavior Direction

After examining the characteristics of the data at the start of the study, the initial group process modeling steps in the MGFP are to gain insights into change and variability in the direction of the behavior or perception (latent mean change and variability). The common slope captures how the dependent variable changes over time and is labeled γ_{100} (TIME$_i$) in the model specification. As shown in Table 1, Step 1 in the in the MGFP is to test/examine the direction of the common slope γ_{100}.

Groups differ from the common slope with a group-specific slope deviation, u_{10j}. The variance of this slope deviation (slope variability) is labeled υ_{11}. The slope deviation captures the degree to which overall perceptions in a particular group shift over the course of the study. Step 2 tests whether there is substantial variability, υ_{11}, in the slope between groups. In interpreting the slope variability, υ_{11}, it is important to take the covariance, υ_{01}, (see model specification) between the slope variance, υ_{11}, and the intercept variance, υ_{00}, into account (see model specification and earlier discussion). Especially when υ_{01} is negative, the actual variability in the slopes may be small. To make the interpretation of υ_{11} easier, we recommend routinely first fitting a model without υ_{01} (Step 2a) and then a model with υ_{01} (Step 2b). In the former model, υ_{11} can directly be interpreted as the overall amount of slope variability (without paying attention to υ_{01}) and the latter model then provides insights into the association between the group-level intercept and slope.

The common slope and variability in the slope are frequently of interest for group research because group environments have the potential to fundamentally change the trajectories of responses. For instance, in juries, jurors share observations about the prosecution’s case during the deliberation period. Presumably this deliberation influences the overall views of group members so that the group becomes more (or less) favorable towards the defendant or prosecution (significant common slope γ_{100}). As another example, newly formed groups might be expected to show increases in group-level constructs such as cohesion as group members interact and develop relationships and a distinct climate is formed (increase in the common slope γ_{100}). In both examples, there may also be substantial variability between groups in the slope.

Change in Consensus

The third step (Step 3) in the MGFP is to test for consensus change. As noted previously, change in consensus can be captured through modeling change in the error variance over time.

The model specification for the MGPF specifically accounts for this change through the equation $\sigma^2\exp[2\delta_1 TIME_t]$. In this equation, σ^2 is the error variance at the initial measurement occasion at the start of the study. Note that the model specification used by the MGPF does not include an individual slope change term in addition to the person intercept (τ). The reason for this model specification is that it ensures that change within the groups is fully captured by δ_1 and can interpreted in a way which is analogous to estimating ICC1 values at each measurement point (Lang & Bliese, 2018; Lang et al., 2018).

The δ_1 term can be interpreted as an approximate linear percentage increase (when δ_1 is positive) or decrease (when δ_1 is negative) in the error standard deviation (σ) with each increase in the unit of the TIME variable. For instance, $\sigma^2 = 1.6$, $\delta_1 = -0.05$, and TIME = 0, 1, 2, 3 would lead to a change pattern in the residual variance σ^2 of $1.6 \times \exp[2 \times -0.05 \times 0] = 1.6$, $1.6 \times \exp[2 \times -0.05 \times 1] = 1.45$, $1.6 \times \exp[2 \times -0.05 \times 2] = 1.31$, and $1.6 \times \exp[2 \times -0.05 \times 3] = 1.19$. Taking the square root of the variance, this is equivalent to a change pattern in the residual standard deviation of $\sigma = \sqrt{1.6} = 1.26$, $\sigma = \sqrt{1.45} = 1.20$, $\sigma = \sqrt{1.31} = 1.14$, and $\sigma = \sqrt{1.19} = 1.09$ which in turn is roughly equivalent to three five percent decreases. Figure 1 illustrates this example.

Note that the exponential term is included in the model to ensure that the error variance always remains positive in the model so that no model misspecifications can occur. Within the exponential term, an additional 2 is included to bring the δ_1 to the unit of the error standard deviation (Note that $\exp[2\delta_1 TIME_t]$ is identical to $\exp[\delta_1 TIME_t]^2$).

Although the model specification for δ_1 may appear to be complex, the substantial interpretation of δ_1 values is relatively straightforward. For instance, a decrease in residual variances for cohesion ratings in a sample of groups (negative δ_1) suggests that members within the groups are becoming more similar to each other in their perceptions of cohesion. In contrast, a divergence pattern with a positive δ_1 suggest that sub-factions within groups are responding differently over time such that one faction is reporting increases in cohesion while another faction is reporting decreases. The idea of sub-factions responding differently clearly has implications for understanding group processes and potentially group outcomes.

The upper part of Table 1 includes a summary of the basic change components in the proposed framework and their substantive interpretation and relevance for studying group processes: The slope (γ_{100}), the slope variability (υ_{11}), and the consensus coefficient, δ_1. As indicated by the examples in Table 1, many basic research questions on group processes can be expressed and tested using this framework.

Basic Examples

To illustrate the use of the MGPF, we present two examples with different types of group datasets. Our first example uses data from a mock jury experiment (Dann, Hans, & Kaye, 2007). Data are freely available on the ICPSR website (Dann, Hans, & Kaye, 2006), and detailed code for this example is provided in the appendix. The mock juries were recruited from a realistic pool of available jurors who watched a hypothetical robbery trial (based on a true case). The jurors then deliberated in juries of eight to reach a verdict. Before the deliberation (Time 1), each juror provided his or her opinion on the case. Jurors specifically provided their personal verdict (0 = not guilty, 1= undecided, 2 = guilty), their confidence in their personal verdict (0 = *not at all* to 10 = *very much*), and the likelihood of the guilt of the defendant (scale from 0 to 100). After the deliberation (Time 2), each jury member again provided their personal opinion on the case using these questions. Finally (Time 3), the jury as a whole provided a final verdict (0 = not guilty, 1= undecided, 2 = guilty) and each juror indicated to what degree he or she agreed with the final verdict and was satisfied with the final verdict (1 = not at all to 5 = very much). To analyze these data using the MGPF, we converted all items to a common scale from 0 (not guilty) to 100 (guilty) using the Percent of Maximal Possible score method (POMP; Cohen, Cohen, Aiken, & West, 1999) and then took the average of the items at each timepoint. POMP is a useful metric because it can be generalized to all types of metrics and be interpreted as the range of the possible scores on a scale. For instance, for a 1-item 5-point Likert scale POMP 100 corresponds to "strongly agree" and POMP 0 to "strongly disagree". Clearly, the fact that the study used different scales between the first two measurement occasions and the third measurements and the fact that Time 2 and Time 3 were not really spaced in time is a limitation; nonetheless, these data illustrate typical group process in a highly realistic experimental setup.

Figure 2 provides an illustration of this dataset. The graph also includes information on two different experimental conditions that we consider later in the paper. Table 2 and Table 3 provide details on the MGPF analyses. Results for Step 1 revealed that there was a significant decrease in the average opinion of jurors toward less guilty (slope: $\gamma_{100} = -7.37$, $p < .001$, see also the estimate in the final model in Table 3). This effect is in line with earlier findings suggesting that jury deliberations typically decrease judgments of guilt (Bornstein & Greene, 2011; MacCoun, 1989). Importantly, however, Step 2 results revealed that juries also significantly varied in their average latent mean trajectories over time. Step 2a indicated a substantial and significant amount of slope variability, $\upsilon_{11} = 176.60$, $\chi^2(df = 1) = 288.58$, $p < .001$ (also see Table 2 and 3). Step 2b also indicated that the covariance was significant, $\chi^2(df = 1) = 13.69$, $p < .001$ suggesting that groups that started at a higher level also showed more variability. The covariance estimate was $\upsilon_{11} = 78.76$ in the Step 2b model and $\upsilon_{11} = 82.60$ in the final model in Table 3. The Step 2 analyses show that the slope of jury opinion change during the deliberation systematically differed among juries despite deliberated about the same case. Finally, Step 3 of the MGFP analysis revealed considerable evidence that juries reached consensus over time. The average personal opinion of the jurors moved closer to the latent means of their groups, $\delta_1 = -.33$, $\chi^2(df = 1) = 86.03$, $p < .001$ (also see Table 2 and 3) and the residual variance decreased from $\sigma^2 = 642.80$ to $\sigma^2 = 171.71$ over the course of the study. These Step 3 findings suggest that juries in this study did not simply take the average of the pre-deliberation opinions. Instead, the analysis suggests that group interactions shaped the nature of the decisions and enabled the juries to reach a final verdict. Jurors appeared to change personally held beliefs to arrive at group consensus.

Our second example involves 1,872 new recruits in basic training in the US Army. The new recruits were members of 42 platoons and represented a subset of data reported in Adler et al. (2015). The example data differ from the original because in the current sample we omitted 67 respondents who provided only one rating over the three trials, and we focus here on cohesion as the outcome variable. Cohesion was assessed at 3, 6 and 9 weeks into the 10-week basic training course using a three-item cohesion measure (Britt, Dickinson, Moore, Castro, & Adler, 2007; Williams et al., 2016). An example item is "The members of my platoon are cooperative with each other". The items were answered on a 5-point Likert scale.

The MGFP results for the Army data are provided in Table 2 and Table 3. As indicated by Table 2, Step 1 indicated that time had a main effect on cohesion ($\gamma_{100} = 0.22$, $p < .001$). Overall, the groups increased their perceived levels of cohesion. Results for Step 2 provided evidence of a significant amount of variance in the trajectory, $\chi^2(df = 1) = 145.45$, $p < .001$ (Step 2a). These findings again suggest that some groups had a stronger change in levels of cohesion than others. The addition of the covariance term (Step 2b) revealed a positive and significant covariance, $\upsilon_{11} = 0.03$, $\chi^2(df = 1) = 8.36$, $p < .001$. The analysis of the patterns in the residual variance (Step 3) revealed that δ_1 was positive with a value of 0.03. However, a comparison of -2log likelihood values indicated that a model that included δ_1 did not significantly fit better than a model that omitted this estimate, $\delta_1 = 0.03$, $\chi^2(df = 1) = 2.11$, $p = .15$. The δ_1 estimate was therefore not significant at the 95% percent confidence level suggesting that groups as a whole showed no discernible variance patterns. As we later illustrate, however, leadership differences among the groups identify distinct variance patterns that are masked here when we examine groups as a whole.

Explaining Change Using Predictors

The framework discussed until now only includes time and the nesting structure as predictors. In many cases, researchers may also be interested in explaining group and individual processes over time. To do so, it is possible to add predictors at the person and group levels to explain direction and consensus processes in two additional steps (Step 4 and Step 5) to the MGFP (see Table 1). The predictors used for Step 4 and Step 5 can be dummy codes for experimental manipulations as well as continuous measurements of stable characteristics.

Group-level Predictors

A model that includes a group-level predictor can be written as follows (Equations for r_{0ij}, u_{00j}, and u_{10j} are identical to the previous model specification and therefore not shown again).

$$
\begin{aligned}
&\text{Level-1 (observations):} && Y_{tij} = \pi_{0ij} + \pi_{1ij}\text{TIME}_t + e_{tij} \\
&\text{Level-2 (persons):} && \pi_{0ij} = \beta_{00j} + r_{0ij} \\
& && \pi_{1ij} = \beta_{10j} \\
&\text{Level-3 (groups):} && \beta_{00j} = \gamma_{000} + \gamma_{010}\text{PRED}_j + u_{00j} \\
& && \beta_{10j} = \gamma_{100} + \gamma_{110}\text{PRED}_j + u_{10j} \\
& && e_{tij} \overset{iid}{\sim} N(0, \sigma^2\exp[2\delta_1\text{TIME}_t] \\
& && \qquad \exp[2\delta_2\text{PRED}_j] \\
& && \qquad \exp[2\delta_3\text{TIME}_t\text{PRED}_j])
\end{aligned}
$$

In this model, several additional effects are included. The model accounts for the fact that a group-level predictor (PRED_j) may explain baseline differences in the intercept at the start of the study (γ_{010}). The effect of the predictor on the intercept is primarily included to make the test of the next effect in Step 4 of the MGFP possible: A group-level effect of the predictor on the slope of the latent group means, (γ_{110}). The effect of the predictor on the slope tests whether the predictor can explain differences in the slope between groups and will frequently be of interest for researchers. In particular, researchers may wish to test the idea that groups that received a particular intervention or have a particular characteristic have a stronger increase in the criterion (Y_{tij}) over time than other groups. For instance, in the Army cohesion data the 42 platoons significantly differed in terms of mean cohesion trajectories so these differences could presumably be explained by some attribute of the group.

Before the interaction between time and the predictor on consensus change can be examined, it is necessary to add the δ_2 effect to the model (Step 5a). This effect accounts for potential differences in consensus between groups explained by the predictor. Step 5b then focuses on the δ_3 effect, and tests the idea that the predictor explains the degree to which groups show consensus or divergence. In other words, this effect allows researchers to test the idea that a particular group characteristic is associated with groups developing consensus.

Person-level Predictors

While group research typically focuses on group-level predictors, the MGPF can also incorporate person-level predictors in Step 4 and Step 5 instead of or in addition to group-level predictors. Person-level predictors may be relevant because they provide insights into how persons with different characteristics interact within their groups. For instance, person characteristics like being a leader or being in a minority may be associated with differential change patterns within the group. A model that incorporates a person-level predictor can be written as follows:

Level-1 (observations): $Y_{tij} = \pi_{0ij} + \pi_{1ij}\text{TIME}_t + e_{tij}$
Level-2 (persons): $\pi_{0ij} = \beta_{00j} + \beta_{01k}(\text{PRED}_{ij}) + r_{0ij}$
$\pi_{1ij} = \beta_{10j} + \beta_{11k}(\text{PRED}_{ij})$
Level-3 (groups): $\beta_{00j} = \gamma_{000} + u_{00j}$
$\beta_{01j} = \gamma_{010}$
$\beta_{10j} = \gamma_{100} + u_{10j}$
$\beta_{11j} = \gamma_{110}$

$$e_{tij} \overset{iid}{\sim} N(0, \sigma^2 \exp[2\delta_1\text{TIME}_t] \exp[2\delta_2\text{PRED}_{ij}] \exp[2\delta_3\text{TIME}_t\text{PRED}_{ij}])$$

As with models including group-level predictors, we recommend that researchers first examine the degree to which the predictor can explain latent mean-level change in Step 4 (adding only β_{01k} and β_{11k}) and then go on and predict consensus change in Step 5 (also adding δ_2 in Step 5a and δ_3 in Step 5b). The interpretation of the effects with person-level predictors at Step 4 and Step 5 is somewhat different from models with group-level effects as the effects now do not refer to the group as a whole but to the position of individuals within their groups. Specifically, the Step 4 and Step 5 effects test to what degree the predictor is associated with persons' position relative to the group-mean level. For instance, minority group members may generally show less consensus with their groups (δ_2) and this level of consensus may increase or decrease over time so that the minority members become less or more consistent with others within the group (δ_3). Likewise, a reasonable assumption is that group leaders in a sample of groups are generally more central (i.e. shows higher levels of consensus) and may typically also become more central in their groups over time.

The lower part of Table 1 summarizes the explanations of group processes that can be tested using the MGPF. This table also summarizes each step in investigating the effects of a group-level or person-level predictor.

Examples for an Explanatory Analysis

Group-Level Predictor in the Jury Dataset

Table 4 and 5 provide examples for analyses with explanatory predictors using the jury dataset. The purpose of the original jury study was to evaluate whether the introduction of DNA evidence could improve the jury process. Therefore, 40 of the 60 juries received additional information about DNA evidence before and during their deliberation. As a basic test of the idea that this additional information altered the deliberation process, we added condition (0 / 1 = No or Yes for additional DNA evidence provided) as a predictor. Results shown in Table 5 indicate that condition was not associated with the intercept or with latent mean-level change for the groups. However, the model comparisons for Step 5 in Table 4 and the model estimates in Table 5 indicate that groups in the DNA condition showed stronger consensus than other groups, $\delta_2 = -0.10$, $\chi^2(df = 1) = 4.7$, $p < .05$. DNA evidence thus generally increased agreement between the jurors across all three time points. Interestingly, the level of consensus did not differentially increase over time for the two conditions, as indicated by the non-significant interaction between condition and time in the model, $\chi^2(df = 1) = 0.0$, n.s.. In summary, providing additional DNA information generally increased consensus, but did not accelerate the emergence of consensus in the juries over time.

Person-Level Predictor in the Jury Dataset

Another research question that is frequently of interest on jury deliberation is whether jurors that are members of a minority (Hispanic, Black, Native American or non-white in the United States) behave differently from white/majority jurors in a group context (Bornstein & Greene, 2011; MacCoun, 1989; Sommers, 2007). To study this research question using the MGPF, we studied non-white minority (yes = 1; no = 0) as a person-level predictor.

We began by adding minority as a predictor of the latent mean-level and latent mean-level change to the model in Step 4. As shown in Table 5, minority jurors differed in their mean level and the amount of change in their latent means over time. Specifically, minority jurors were less convinced that the defendant was guilty at the start of the study at T1 ($\beta_{01k} = -14.30$, $p < .001$). This finding is in line with earlier research suggesting that the experiences of non-white Americans with the justice system differ from the white majority (Sommers, 2007). Minority jurors also had a less steep slope over time, $\beta_{01k} = 4.47$, $p < .01$. Thus overall, minority members showed less change from their original view of less guilt than non-minority jurors. A potential explanation for this pattern is that minority jurors had already identified or thought about potential mitigating circumstances on their own, and thus additional information on potentially mitigating circumstances led to less change in their opinions.

As indicated by Table 4, results for Step 5 revealed that minority jurors also differed from white/majority jurors in their level of consensus with the group. As shown in Table 5, minority jurors generally showed less consensus with their jury, $\delta_2 = 0.13$, $\chi^2(df = 1) = 4.24$, $p < .05$. This pattern remained constant across time as indicated by a missing interaction of the predictor with time, $\chi^2(df = 1) = 1.3$, n.s.

Group-Level Predictor in the Basic Training Dataset

In our third example, we examined whether having highly rated drill sergeants in the beginning of basic training (modeled as a shared platoon-level rating) would be related to changes in the mean trajectories and to changes in the residual variances. Despite evidence suggesting that platoons differed in terms of mean trajectories, we found no evidence that these differences were related to leadership, $\beta_{01k} = -0.03$, n.s., in Step 4.

Likewise, with respect to residual variances (Step 5a), we found no evidence to suggest that ratings of leadership were related to mean levels of residual variance, $\delta_2 = 0.01$, $\chi^2(df = 1) = 0.48$, n.s. suggesting that overall levels of consensus were unrelated to how the groups rated their leadership. Interestingly, though, we found that leadership interacted with time with respect to the residual variance, $\delta_3 = .04$, $\chi^2(df = 1) = 4.07$, $p < .05$. The residual variance pattern indicated that at the first time period, groups with leadership ratings one-standard deviation below the mean and groups with leadership ratings one standard deviation above the mean had similar residual variances ($\sigma^2 = 0.37$ and $\sigma^2 = 0.34$, low and high leadership, respectively). In contrast, at the last time period, groups with initially low leadership ratings were relatively unchanged at $\sigma^2 = 0.36$ but groups with initially high leadership ratings had a variance increase to $\sigma^2 = 0.43$. This effect is a bit surprising as one might normally expect that positive leadership would help groups develop strong shared climates. In this case, though, units with positive leadership tended to have members who showed more (not less) variability in terms of cohesion over time. A potential explanation for this type of finding is that only some soldiers resonated with the strong leadership that was provided and other soldiers who struggled in basic training felt more estranged by highly rated leaders. Ultimately, though, we caution readers not to over-interpret these exploratory findings, but rather to appreciate the types of questions that can be addressed.

Advantages of the Multilevel Group Process Framework

The described multilevel framework addresses several issues and challenges in studying group processes and also allows researchers to study novel research questions. In this section, we highlight some of these specific advantages.

Systematic Approach for Studying Group Processes

The MGPF translates several theoretical group phenomena into specific hypotheses within a statistical model and thus allows researchers to study these hypotheses empirically in laboratory or field data. We have also presented a step-by-step procedure that researchers can follow to systematically map and explore how different change phenomena are present in their data.

Jointly Studying Consensus Emergence vs. Direction Shift

One core issue in many earlier approaches for studying groups is that these approaches frequently focus on group phenomena of directional change versus consensus change in isolation. The MGPF models both types of processes within one framework and thus helps avoid potential bias when researcher do not account for the respective other process in the context of growth models (Bliese & Ployhart, 2002; Singer & Willett, 2003) or models that account for consensus change (Lang & Bliese, 2018; Lang et al., 2018).

Formal Tests

One important advantage of the methods described in this article is that they provide formal tests for the presence of consensus emergence and unbiased tests for mean-level change. Extant approaches for studying consensus emergence using, for instance, the ICC1 did not directly provide researchers with a way to statistically test whether a pattern of change in their data was significant or not. Likewise, extant approaches for studying mean-level change typically aggregate group member responses and ignore potential within-group changes in consensus. Notably, viewing group processes and dynamics from a multilevel perspective also allows researchers to add predictor variables of different types of change and can thus be used with both correlational and also experimental data.

Flexible Estimation

Although the framework described is not based on basic mixed-effects multilevel model, the types of models can nevertheless be estimated in a variety of software programs like nlme

(Pinheiro & Bates, 2000), glmmTMB (Brooks et al., 2017), and the Bayesian MCMCglmm (Hadfield, 2010) in R, NLMIXED in SAS (Wolfinger, 1999) or Mplus (Muthén & Muthén, 2018). A somewhat different parametrization of the described models (see Lang et al., 2018 for detail) can also be fitted using flexible standard multilevel software like lme4 in R (Bates, Maechler, Bolker, & Walker, 2015), mlwin (Goldstein, 2011) or HLM (Raudenbush & Bryk, 2002). In the appendix to this paper, we provide code for the Jury data analyses using nlme in R.

Conclusion

We started this article by noting that psychologists are frequently interested in understanding the complex change processes that take place in groups in a variety of applied settings. However, when researchers seek to study and model these processes they are confronted with the fact that several different processes (e.g., group leniency, group extremity, and group freezing, also see examples in Table 1) can simultaneously unfold in groups over time. In this article, we have described a framework that allows researchers to track mean-level change and consensus forming in groups and to test predictors of both processes. We believe that this framework provides a basis for gaining new insights into how group processes unfold over time.

References

Adler, A. B., Bliese, P. D., Pickering, M. A., Hammermeister, J., Williams, J., Harada, C., … Ohlson, C. (2015). Mental skills training with basic combat training soldiers: A group-randomized trial. *Journal of Applied Psychology*, *100*(6), 1752–1764. https://doi.org/10.1037/apl0000021

Ashforth, B. E. (1985). Climate formation: Issues and extensions. *Academy of Management Review*, *10*, 837–847. https://doi.org/10.2307/258051

Bates, D. M., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models using lme4. *Journal of Statistical Software*. Retrieved from http://arxiv.org/abs/1406.5823

Bliese, P. D. (2000). Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In K. J. Klein & S. W. J. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions* (pp. 349–381). San Francisco, CA: Jossey-Bass.

Bliese, P. D., & Ployhart, R. E. (2002). Growth modeling using random coefficient models: Model building, testing, and illustrations. *Organizational Research Methods*, *5*, 362–387. https://doi.org/10.1177/109442802237116

Bodner, T. E., & Bliese, P. D. (2017). Detecting and differentiating the direction of change and intervention effects in randomized trials. *Journal of Applied Psychology*, *103*, 37–53. https://doi.org/10.1037/apl0000251

Bornstein, B. H., & Greene, E. (2011). Jury decision making: Implications for and from psychology. *Current Directions in Psychological Science*, *20*(1), 63–67.

https://doi.org/10.1177/0963721410397282

Britt, T. W., Dickinson, J. M., Moore, D. W., Castro, C. A., & Adler, A. B. (2007). Correlates and consequences of morale versus depression under stressful conditions. *Journal of Occupational Health Psychology*, *12*(1), 34–47. https://doi.org/10.1037/1076-8998.12.1.34

Brooks, M. E. J. K. K., van Benthem, K., Magnusson, A., Berg, C. W., Nielsen, A., Skaug, H. J., … Bolker, B. M. (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R Journal*, *9*, 378–400. Retrieved from https://journal.r-project.org/archive/2017/RJ-2017-066/index.html

Bryk, A. S., & Raudenbush, S. W. (1987). Application of hierarchical linear models to assessing change. *Psychological Bulletin*, *101*, 147–158. https://doi.org/10.1037//0033-2909.101.1.147

Chiu, M. M., & Lehmann-Willenbrock, N. (2016). Statistical discourse analysis: Modeling sequences of individual actions during group interactions across time. *Group Dynamics: Theory, Research, and Practice*, *20*, 242–258. https://doi.org/http://dx.doi.org/10.1037/gdn0000048

Cohen, P., Cohen, J., Aiken, L. S., & West, S. G. (1999). The problem of units and the circumstance for POMP. *Multivariate Behavioral Research*, *34*, 315–346. https://doi.org/10.1207/S15327906MBR3403_2

Corsini, R. J., & Rosenberg, B. (1955). Mechanisms of group psychotherapy: Processes and dynamics. *Journal of Abnormal and Social Psychology*, *51*(3), 406–411. https://doi.org/10.1037/h0048439

Cronin, M. A., & Weingart, L. R. (2011). Dynamics in groups: Are we there yet ? *The Academy of Management Annals*, *5*, 37–41. https://doi.org/10.1080/19416520.2011.590297

Dann, B. M., Hans, V. P., & Kaye, D. H. (2006). *Testing the effects of selected jury trial innovations on juries' comprehension of DNA evidence in New Castle County, Delaware, 2003. ICPSR04356-v1.* Washington, DC: National Institute of Justice. https://doi.org/10.3886/ICPSR04356.v1

Dann, B. M., Hans, V. P., & Kaye, D. H. (2007). Can Jury Trial Innovations Improve Juror Understanding of DNA Evidence? *National Institute of Justice Journal*, *255*, 2–6. Retrieved from http://heinonlinebackup.com/hol-cgi-bin/get_pdf.cgi?handle=hein.journals/judica90&section=43

Gehman, J., Trevino, L., & Garud, R. (2013). Values work: A process study of the emergence and performance of organizational values practices. *Academy of Management Journal*, *56*, 84–112. https://doi.org/10.5465/amj.2010.0628

Goldstein, H. (2011). *Multilevel statistical models (4th ed.)*. *Book*. Chichester, UK: Wiley. https://doi.org/10.2307/1534624

Hadfield, J. D. (2010). MCMC methods for multi-response generalized linear mixed models: The MCMCglmm R Package. *Journal of Statistical Software*, *33*, 1–22.

Hedeker, D., & Mermelstein, R. (2007). Mixed-effects regression models with heterogenous variance: Analyzing ecological momentary assessment (EMA) data of smoking. In T. D. Little, J. A. Bovaird, & N. A. Card (Eds.), *Modeling contextual effects in longitudinal studies* (pp. 183–206). Mahwah, NJ: Lawrence Erlbaum.

Hedeker, D., Mermelstein, R. J., & Demirtas, H. (2012). Modeling between-subject and within-subject variances in ecological momentary assessment data using mixed-effects location scale models. *Statistics in Medicine*, *31*, 3328–3336. https://doi.org/10.1002/sim.5338

Huisman, M., & Snijders, T. A. B. (2003). Statistical analysis of longitudinal network data with changing composition. *Sociological Methods & Research*, *32*, 253–287. https://doi.org/10.1177/0049124103256096

Humphrey, S. E., & Aime, F. (2014). Team microdynamics: Toward an organizing approach to teamwork. *The Academy of Management Annals*, *8*, 443–503. https://doi.org/10.1080/19416520.2014.904140

Kim, J., & Seltzer, M. (2011). Examining heterogeneity in residual variance to detect differential response to treatments. *Psychological Methods*, *16*, 192–208. https://doi.org/10.1037/a0022656

Kozlowski, S. W. J. (2017). Advancing research on team process dynamics: Theoretical, methodological, and measurement considerations. *Organizational Psychology Review*. https://doi.org/10.1177/2041386610376255

Kozlowski, S. W. J., & Bell, B. S. (2013). Work groups and teams in organizations. In *Handbook of psychology: Industrial and organizational psychology* (pp. 412–469). Hoboken, NJ: John Wiley.

Kozlowski, S. W. J., & Chao, G. T. (2012). The dynamics of emergence: Cognition and cohesion in work teams. *Managerial and Decision Economics*, *33*, 335–354. https://doi.org/10.1002/mde.2552

Kozlowski, S. W. J., Chao, G. T., Grand, J. A., Braun, M. T., & Kuljanin, G. (2013). Advancing multilevel research design: Capturing the dynamics of emergence. *Organizational Research Methods*, *16*, 581–615. https://doi.org/10.1177/1094428113493119

Kozlowski, S. W. J., & Ilgen, D. R. (2006). Enhancing the effectiveness of work groups and teams. *Psychological Science in the Public Interest*, *7*, 77–124. https://doi.org/10.1111/j.1529-1006.2006.00030.x

Lang, J. W. B., & Bliese, P. D. (2018). A temporal perspective on emergence: Using three-level mixed effects models to track consensus emergence in groups. In S. E. Humphrey & J. M. LeBreton (Eds.), *The Handbook for multilevel theory, measurement, and analysis*. Washington, DC: APA.

Lang, J. W. B., Bliese, P. D., & de Voogt, A. (2018). Modeling consensus emergence in groups using longitudinal multilevel models. *Personnel Psychology*, *71*, 255–281. https://doi.org/10.1111/peps.12260

Lewin, K. (1947). Frontiers in group dynamics: Concept, method and reality in social science; social equilibria and social change. *Human Relations*, *1*, 5–41. https://doi.org/10.1177/001872674700100103

Lewin, K., Station, L., & White, R. K. (1939). Patterns of aggressive behavior in experimentally created social climates. *Journal of Social Psychology*, *10*(2), 271–299. https://doi.org/10.1080/00224545.1939.9713366

MacCoun, R. J. (1989). Jury decision-making. *Science*, *244*(4908), 1046–1050.

Mayo, E. (1933). *The human problems of an industrial civilization*. New York, NY: Macmillan.

Mohammed, S., Hamilton, K., & Lim, A. (2009). The incorporation of time in team research: Past, current, and future. In E. Salas, G. F. Goodwin, & C. S. Burke (Eds.), *Team effectiveness in complex organizations: Cross-disciplinary perspectives and approaches* (pp. 321–346). Mahwah, NJ: Erlbaum.

Muthén, L. K., & Muthén, B. O. (2018). *Mplus user's guide. Eighth edition.* Los Angeles, CA: Muthén & Muthén.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS*. New York, NY, NY: Springer.

Rast, P., MacDonald, S. W. S., & Hofer, S. M. (2012). Intensive measurement designs for research on aging. *GeroPsych: The Journal of Gerontopsychology and Geriatric Psychiatry*, *25*, 45–55. https://doi.org/10.1024/1662-9647/a000054

Raudenbush, S. W., & Bryk, A. S. (1987). Examining correlates of diversity. *Journal of Educational and Behavioral Statistics*, *12*, 241–269. https://doi.org/10.3102/10769986012003241

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models* (2nd ed.). Newbury Park, CA: Sage.

Sherif, M. (1935). A study of some social factors in perception. *Archives of Psychology*, *27*, 1–60.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurence*. New York, NY: Oxford University Press.

Sommers, S. R. (2007). Race and the decision making of juries. *Legal and Criminological*

Psychology, *12*(2), 171–187. https://doi.org/10.1348/135532507X189687

Williams, J., Brown, J. M., Bray, R. M., Anderson Goodell, E. M., Rae Olmsted, K., & Adler, A. B. (2016). Unit cohesion, resilience, and mental health of soldiers in basic combat training. *Military Psychology*, *28*(4), 241–250. https://doi.org/10.1037/mil0000120

Wolfinger, R. D. (1999). *Fitting nonlinear mixed models with the new NLMIXED procedure (paper 287)*. Cary, NC: SAS Institute Inc.

Table 1
Group Processes, Explanations for Group Processes, and Model Testing Sequence in the Multilevel Group Process Framework (MGPF)

Model component	Description	Interpretation	Examples for research questions	Step in model evaluation
Group Processes				
γ_{100}	Slope	Latent means for groups increase/decrease in the sample	General group leniency/toughness, performance increases/decreases (e.g., group learning)	Step 1
υ_{11}	Slope variability	When υ_{01} is positive or constrained to 0, υ_{11} provides evidence for group extremity	Group think, group extremity	Step 2a (υ_{11} added) and Step 2b (υ_{01} added)
δ_1	Consensus change	Increase or decrease in consensus	Group forming, group freezing, consensus emergence, group adjourning	Step 3
Group-level Predictors as Explanations for Group Processes				
γ_{110}	Group-level predictor effect on slope	Groups with higher/lower predictor values show stronger/less increases in their latent group means over time	An intervention or stable group characteristic predicts which groups learn a new task or change their perceptions	Step 4
δ_3	Group-level predictor effect on consensus change	Groups with higher/lower predictor values develop more/less consensus/divergence	An intervention or stable group characteristic predicts which groups develop consensus/divergence	Step 5a (main δ_3 effect for predictor added) and Step 5b (δ_3 added)
Person-level Predictors as Explanations for Group Processes				
β_{11k}	Person-level predictor effect on slope	Persons with higher/lower predictor values show stronger/less increases in their latent group means over time	Leaders showing more/less decline than other group members; minority participants in a group show less increase/more increase in their mean opinion than other participants	Step 4
δ_3	Person-level predictor effect on consensus change	Persons with higher/lower predictor values become more/less central in the group over time	Leaders becoming more central for group climate over time; minorities adapting to the majority opinion or vice versa	Step 5a (main δ_3 effect for predictor added) and Step 5b (δ_3 added)

Table 2

Basic Modeling Steps in the Jury and the Basic Training Datasets

Data, model	AIC	BIC	*logLik*	*df*	χ^2 vs. previous model
Jury data					
Step 1	13,573.1	13,599.4	-6,781.5	5	
Step 2a (υ_{11} added)	13,286.5	13,318.1	-6,637.2	6	288.6***
Step 2b (υ_{01} added)	13,274.8	13,311.7	-6,630.4	7	13.7***
Step 3 (δ_1 added)	13,190.8	13,233.0	-6,587.4	8	86.0***
Basic training data					
Step 1	13,425.7	13,458.8	-6,707.8	5	
Step 2a (υ_{11} added)	13,282.2	13,322.0	-6,635.1	6	145.4***
Step 2b (υ_{01} added)	13,275.9	13,322.3	-6,630.9	7	8.4**
Step 3 (δ_1 added)	13,275.7	13,328.8	-6629.9	8	2.1

Note. For the jury data, N = 1,440 observations nested in 480 jurors and 60 juries. For the basic training data, N = 5,635 observations nested in 1,945 group members and 42 groups.

**p<.01; ***p < .001

Table 3

Model Estimates For Step 3 in the Jury and the Basic Training Datasets

Parameters	Jury Data	Basic Training Data
Model estimates		
Intercept, γ_{000}	68.89***	2.73***
TIME, γ_{100}	**-9.34*****	**0.22*****
Group intercept variance, υ_{00}	83.88	0.12
Group slope variance for TIME, υ_{11}	**154.74**	**0.03**
Covariance, υ_{01}	**82.60**	**-0.03**
Person intercept variance, τ	192.21	0.37
Residual variance, σ^2	642.80	0.35
TIME, δ_1	**-0.33**	**0.03**

Note. Estimates that are the focus in the basic steps of the multilevel group process framework (see Table 1) are shown in bold. For the jury data, N = 1,440 observations nested in 480 jurors and 60 juries. For the basic training data, N = 5,635 observations nested in 1,945 group members and 42 groups.

* $p < .05$ *** $p < .001$

Table 4

Predictors as Explanations For Group Processes in the Jury and the Basic Training Datasets

Data, model	AIC	BIC	*logLik*	*df*	χ^2 vs. previous model
Jury data with CONDITION predictor					
Step 4 ($\gamma_{010} + \gamma_{110}$ added)	13,185.0	13,237.7	-6,582.5	10	
Step 5a ($\gamma_{010} + \gamma_{110} + \delta_2$ added)	13,182.3	13,240.2	-6,580.1	11	4.7*
Step 5b ($\gamma_{010} + \gamma_{110} + \delta_2 + \delta_3$ added)	13,184.2	13,247.5	-6,580.1	12	0.0
Jury data with MINORITY predictor					
Step 4 ($\beta_{01k} + \beta_{11k}$ added)	13,163.2	13,215.9	-6,587.6	10	
Step 5a ($\beta_{01k} + \beta_{11k} + \delta_2$ added)	13,161.0	13,218.9	-6,569.5	11	4.2*
Step 5b ($\beta_{01k} + \beta_{11k} + \delta_2 + \delta_3$ added)	13,161.7	13,224.9	-6,568.8	12	1.3
Basic training data with LEADER predictor					
Step 4 ($\beta_{01k} + \beta_{11k}$ added)	13,287.2	13,353.5	-6,633.6	10	
Step 5a ($\beta_{01k} + \beta_{11k} + \delta_2$ added)	13,288.7	13,361.7	-6,633.3	11	0.5
Step 5b ($\beta_{01k} + \beta_{11k} + \delta_2 + \delta_3$ added)	13,286.6	13,366.3	-6,631.3	12	4.1*

Note. For the jury data, N = 1440 observations nested in 480 jurors and 60 juries. For the basic training data, N = 5,635 observations nested in 1,945 group members and 42 groups.

*p<.05

Table 5

Model Estimates For Explanatory Models With Predictors in the Jury and the Basic Training Datasets

Parameters	Jury data with CONDITION predictor (Step 5a)	Jury data with MINORITY predictor (Step 5a)	Basic training data with LEADER predictor (Step 5b)
Intercept, γ_{000}	67.03***	71.99***	2.73***
TIME, γ_{100}	**-11.58*** **	**-10.32*** **	**0.22*** **
CONDITION, γ_{010}	**2.82**		
TIME × CONDITION, γ_{110}	**3.35**		
MINORITY, β_{01k}		**-14.30*** **	
TIME × MINORITY, β_{11k}		**4.47** **	
LEADER, β_{01k}			**0.07**
TIME × LEADER, β_{11k}			**-0.03**
Group intercept variance, υ_{00}	87.94	77.62	0.11
Group slope variance for TIME, υ_{11}	**155.07**	**156.62**	**0.03**
Covariance, υ_{01}	**80.46**	**82.59**	-0.03
Person intercept variance, τ	193.32	183.86	0.37
Residual variance, σ^2	734.12	608.93	0.35
TIME, δ_1	**-0.33**	**-0.34**	**0.03**
CONDITION, δ_2	**-0.10**		
TIME × CONDITION, δ_3			
MINORITY, δ_2		**0.13**	
TIME × MINORITY, δ_3			
LEADER, δ_2			**-0.03**
TIME × LEADER, δ_3			**0.04**

Note. Estimates that are the focus of the multilevel group process framework are shown in bold. For the jury data, N = 1440 observations nested in 480 jurors and 60 juries. For the basic training data, N = 5,635 observations nested in 1,945 group members and 42 groups.

* $p < .05$ *** $p < .001$

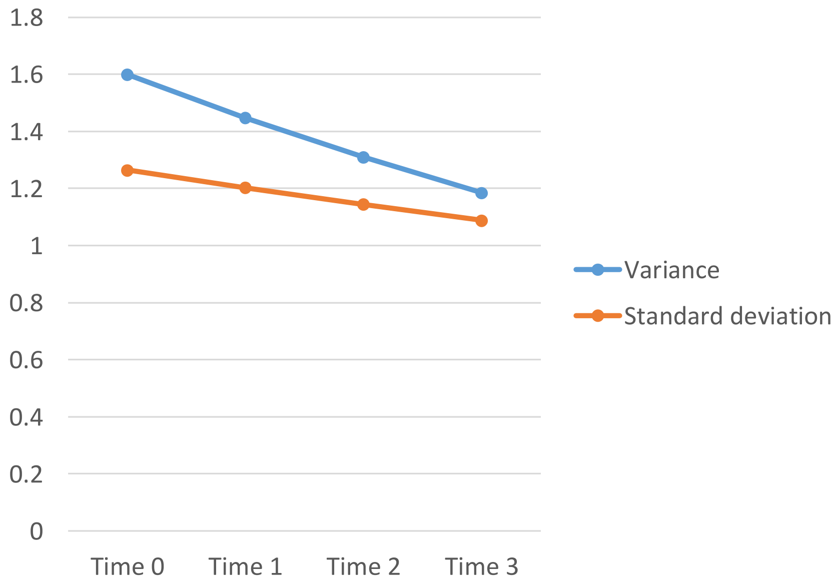

Figure 1. Change in the residual standard deviation and variance when the variance is σ^2 = 1.6 at Time 0 and δ_1 = -0.05.

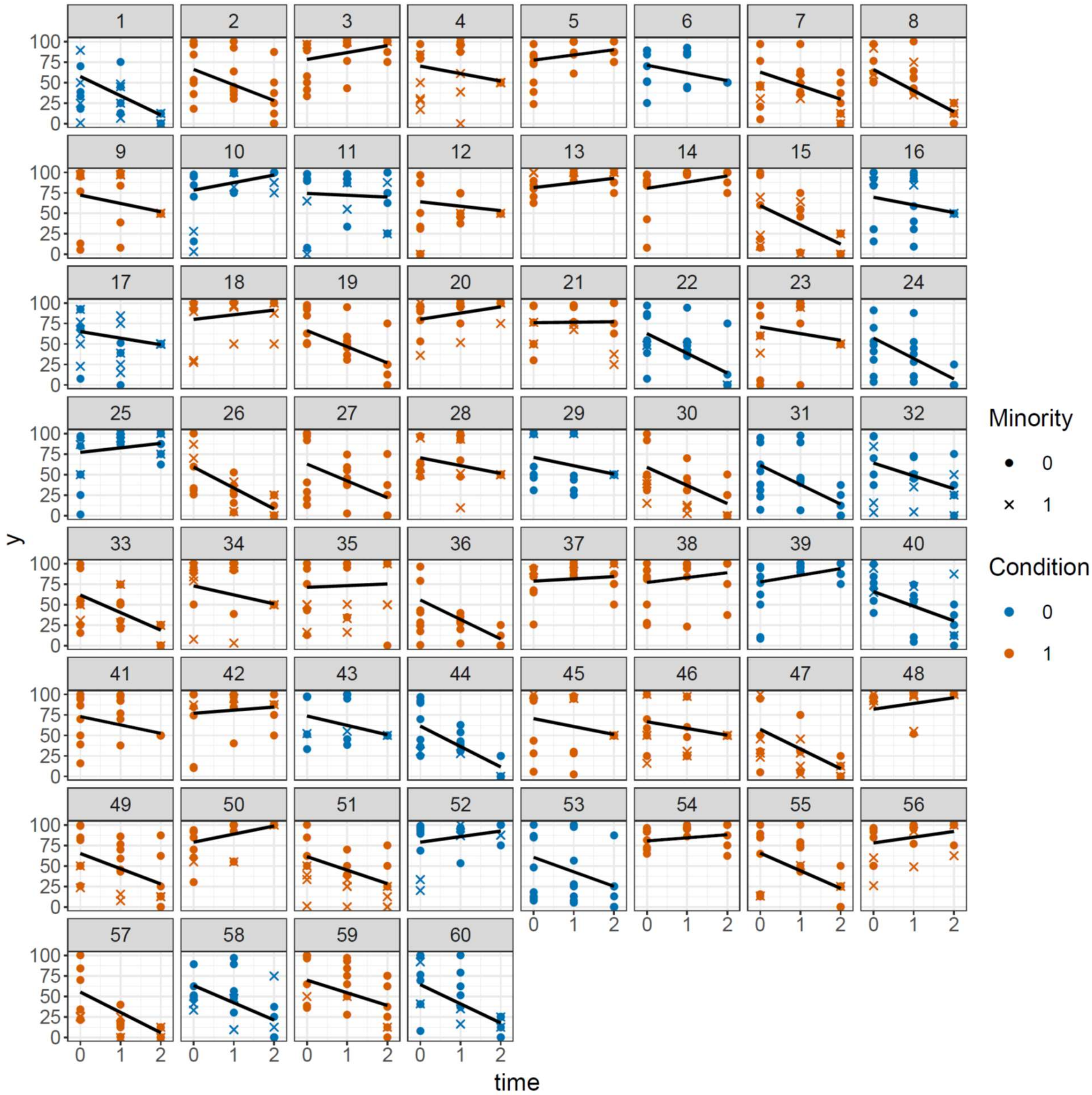

Figure 2. Change patterns in the Jury data (Dann et al., 2006). Observations for groups provided with additional information on DNA (1) are shown in red and groups not provided with additional information are graphed in blue (0).

Appendix

```
# read in the 04356-0001-Data.dta file

idat<-"C:\\mypath\\04356-0001-Data.dta"

library(foreign)
jdatnum<-read.dta(idat,convert.factors = FALSE,missing.type=TRUE)

rdat<-subset(jdatnum,select=c(CASEID,CONDITON,JURY,JUROR,
         QN3AGE,QN3GENDE,QN3RACE,
         QN2VERDI,QN2CONF,QN2KEVIN,
         QN3VERDI,QN3CONFI,QN3LIKEL,
         QN3SATVE,QN3AGREE,VERDICT))


########################### DATA PREPARATION ###############################

library(car)

rdat$FEMALE<-recode(rdat$QN3GENDE,"-1=NA;1=0;2=1")
rdat$AGE<-ifelse(rdat$QN3AGE==-1,NA,rdat$QN3AGE)
rdat$MINORITY<-recode(rdat$QN3RACE,"-1=1;1=1;2=0;3=1;4=1;5=1;6=1;7=1")
rdat$CONDITION<-recode(rdat$CONDITON,"1=0;2=0;3=1;4=1;5=1;6=1")


#POMP = 100*(raw - min)/(max - min)
# T1
rdat$QN2VERDI2<-recode(rdat$QN2VERDI,"-1=NA;1=3;3=2;2=1")
rdat$QN2CONF2<-recode(rdat$QN2CONF,"-1=NA")
rdat$T1I1<-ifelse(rdat$QN2VERDI2==2,50,
                ifelse(rdat$QN2VERDI2==1,abs(50*(rdat$QN2CONF2-10)/(10-1)),
                ifelse(rdat$QN2VERDI2==3,10*(5+5*(rdat$QN2CONF2-1)/(10-1)),NA)))
rdat$T1I2<-rdat$QN2KEVIN

#POMP = 100*(raw - min)/(max - min)
# T2
rdat$QN3VERDI2<-recode(rdat$QN3VERDI,"-1=NA;1=3;3=2;2=1")
rdat$QN3CONF2<-recode(rdat$QN3CONFI,"-1=NA")
rdat$T2I1<-ifelse(rdat$QN3VERDI2==2,50,
                ifelse(rdat$QN3VERDI2==1,abs(50*(rdat$QN3CONF2-10)/(10-1)),
                ifelse(rdat$QN3VERDI2==3,10*(5+5*(rdat$QN3CONF2-1)/(10-1)),NA)))
rdat$T2I2<-rdat$QN3LIKEL

#POMP = 100*(raw - min)/(max - min)
# T3
rdat$QN3SATVE2<-recode(rdat$QN3SATVE,"-1=NA")
rdat$QN3AGREE2<-recode(rdat$QN3AGREE,"-1=NA")
cor(rdat$QN3SATVE2,rdat$QN3AGREE2,use="complete.obs")
rdat$VERDICT2<-recode(rdat$VERDICT,"-1=NA;1=3;3=2;2=1")
rdat$T3I1<-ifelse(rdat$VERDICT2==2,50,
                ifelse(rdat$VERDICT2==1,abs(100*(rdat$QN3SATVE2-5)/(5-1)),
                ifelse(rdat$VERDICT2==3,100*(rdat$QN3SATVE2-1)/(5-1),NA)))
rdat$T3I2<-ifelse(rdat$VERDICT2==2,50,
                ifelse(rdat$VERDICT2==1,abs(100*(rdat$QN3AGREE2-5)/(5-1)),
                ifelse(rdat$VERDICT2==3,100*(rdat$QN3AGREE2-1)/(5-1),NA)))

rdat$T1<-rowMeans(subset(rdat,select=c(T1I1,T1I2)),na.rm=T)
rdat$T2<-rowMeans(subset(rdat,select=c(T2I1,T2I2)),na.rm=T)
rdat$T3<-rowMeans(subset(rdat,select=c(T3I1,T3I2)),na.rm=T)

library(reshape)
longdat<-reshape(subset(rdat,select=c(CASEID,CONDITION,JURY,JUROR,
         AGE,FEMALE,MINORITY,T1,T2,T3)),idvar="CASEID",
         varying=list(c("T1","T2","T3")),direction="long",
         v.names = "verdict")
longdat$time<-longdat$time-1
```

```
############## ANALYSIS #########################


longdat$DV<-longdat$verdict
longdat$GROUP<-longdat$JURY
longdat$PERSON<-longdat$JUROR
longdat$TIME<-longdat$time
longdat$PRED<-longdat$MINORITY
# longdat$PRED<-longdat$CONDITION # uncomment for analysis with CONDITION


library(nlme)

step1<-lme(DV ~ TIME, random =
         list(GROUP=pdDiag(~1),PERSON=pdSymm(~1)),
         data=longdat,na.action=na.omit)
step2a<-update(step1,random = list(GROUP=pdDiag(~TIME),PERSON=pdSymm(~1)))
step2b<-update(step2a,random = list(GROUP=pdSymm(~TIME),PERSON=pdSymm(~1)))
step3<-update(step2b,weights=varExp( form = ~ TIME))

anova(step1,step2a,step2b,step3)


summary(step3)$tTable
VarCorr(step3)
summary(step3)$modelStruct$varStruct


step4<-lme(DV ~ TIME*PRED, random =
         list(GROUP=pdSymm(~TIME),PERSON=pdSymm(~1)),
         data=longdat,na.action=na.omit,weights=varExp( form = ~ TIME))
step5a<-update(step4,weights=
         varComb(varExp( form = ~ TIME),varExp( form = ~ PRED)))
step5b<-update(step4,weights=
         varComb(varExp( form = ~ TIME),varExp( form = ~ PRED),
                varExp( form = ~ PRED*TIME)))

anova(step4,step5a,step5b)

summary(step5a)$tTable
VarCorr(step5a)
summary(step5a)$modelStruct$varStruct
```