Beyond Agreement, Aggregation, and Centrality:

The Role of Influence and Selection in Multilevel Theory

Daniel J. Griffin[1] & Ryan S. Hemsley[2]

[1] Michigan State University, Department of Psychology

[2] Michigan State University, Department of Management

Author Note

Both authors contributed equally.

Abstract

Research on multilevel theory has generated historic advancements in the study of organizations. However, this paradigm did not yield a substantive accumulation of findings until after there was an alignment between theoretical and methodological efforts. Since this surge forward, organizational scholars have primarily relied on multilevel modeling. However, mean aggregation makes the problematic assumption of structural equivalence. To avoid this issue, researchers have used social network analysis as an alternative. However, the dominant use of centrality indices makes a different—but equally problematic—assumption of nodal equivalence. We identify two key theoretical processes that these conventional approaches cannot adequately probe: Influence and Selection. Influence processes focus on how people change because of the networks, while Selection processes focus on how networks change because of the people. We echo previous calls for a paradigm shift in multilevel theoretical and methodological frameworks. Specifically, we articulate the essential role of Influence and Selection processes in addressing the microfoundations of multilevel theory. To facilitate these efforts, we outline statistical models that quantitatively represent these processes while avoiding the problematic assumptions made by conventional conceptual and analytical tools of multilevel researchers.

Beyond Agreement, Aggregation, and Centrality:

The Role of Influence and Selection in Multilevel Theory

Theoretical developments are often downstream of methodological developments. Take, for instance, the historical gridlock between sociological and socio-psychological perspectives of organizational climate (Kozlowski & Hattrup, 1992; Schmidt & Hunter, 1989). Controversy surrounding how to operationalize and study higher level constructs ultimately ended in a consensus about the role of agreement and aggregation in multilevel theory (Bliese, 2000; Klein et al., 2000). This alignment between theory and methods unlocked a new multilevel paradigm, creating a stream of empirical research and rapid theoretical developments regarding teams, multiteam systems, and intra-organizational dynamics (Kozlowski & Bell, 2013; Mathieu et al., 2017). Multilevel modeling (MLM) has been invaluable in this surge of research on multilevel theory; however, researchers have begun to appreciate the potential of social network analysis (SNA) to complement multilevel research by representing nuances in the localized contexts that individuals are embedded within. Indeed, Mathieu and Chen (2011, p. 626) called for a paradigm shift when they stated, “network approaches… may prove valuable for generating alternative paradigms on which new multilevel quantitative investigations and methodologies could be advanced.”

Organizational scholars have primarily relied on MLM practices that overlook structural variance. For example, researchers have criticized the tendency to use agreement and aggregation to dismiss within-team variability as purely a methodological concern, arguing that this practice obscures theory relevant variance (DeRue et al., 2010; Matusik et al., 2019). This is broadly recognized, however, and organizational researchers often use SNA when they believe structural variance is theoretically relevant. Otherwise, they assume *structural equivalence*:

where all members are equally connected and, thereby, members with the same characteristics experience their team in the same way (see Figure 1 for a visualiation of structural equivalence and variance; Granovetter, 1985).[1]

Yet, while appropriate in certain contexts, organizational researchers embrace structural equivalence without fully grasping the consequences. For instance, the team conflict literature was plagued for decades after multiple meta-analyses (De Dreu & Weingart, 2003; de Wit et al., 2012) suggested task conflict may not be as beneficial for team effectiveness as theorized (Jehn, 1995). Despite theory having long recognized the potential for team members to possess asymmetrical perceptions of conflict (Jehn et al., 2010), researchers only recently compared alternative operationalizations of conflict (e.g., individual, dyad, subgroup, and team). Taking a network approach, Shah and colleagues (2021) shed light on the functional task conflict conundrum by showing that discrepancies between theory and findings may largely be due to how researchers nearly exclusively conceptualize and operationalize conflict at the team level.

While network approaches do not necessarily assume structural equivalence, conventional applications of SNA make a different—but equally problematic—assumption of *nodal equivalence*: where members who occupy the same structural position have the same characteristics (see Figure 1 for a visualiation of nodal equivalence and variance; Borgatti & Foster, 2003). This assumption became embedded when the organizational sciences imported SNA techniques from sociology (Kilduff & Lee, 2020). Indices of network structure (e.g., centrality, betweenness, degree) overlook variance in nodal characteristics. For example,

[1] The definition of structural equivalence we employ is distinct from both the structural (two individuals are connected to exactly the same people but not necessarily to each other) and role equivalence (two individuals are connected to different people but occupy equivalent roles in the network) discussed by Kilduff and Lee (2020). In a sense, this structural and role equivalence is more closely related to our discussion on nodal equivalence. Indeed, individuals occupying the same structure or role can have different nodal characteristics.

individuals with three friends (i.e., degree centrality) are treated as mathematically interchangeable and statistically indistinguishable—even if the characteristics of those friends are different (e.g., one person has three expert friends, and another has three non-expert friends).

Importantly, despite the incumbent paradigm, researchers do not need to “pick their poison,” either choosing MLM techniques that make structural equivalence assumptions or SNA techniques that make nodal equivalence assumptions. Instead, researchers can utilize a different theoretical and methodological paradigm which allows them to “have their cake and eat it too.” In this manuscript we introduce the Influence and Selection framework of studying network processes in collectives that avoids both the structural and nodal equivalence assumptions made by conventional MLM and SNA. Influence processes capture how people change due to the characteristics of their network members. For example, leaders can inspire team members to look beyond self-interest and pursue a shared vision (Bono & Judge, 2004). Selection processes, by contrast, capture how people form and maintain social connections. For example, leadership emergence can be conceptualized as a Selection process, where individuals claim and grant leadership forming leadership ties (DeRue & Ashford, 2010). From a social network perspective, Influence processes capture how people change because of their network, while Selection processes capture how networks change because of the people. This perspective suggests novel theoretical directions, including the relationship with team processes and emergent states (Marks et al., 2001). It also creates the potential to change the way researchers view and analyze social network data.

To facilitate the study of these processes and advance research on multilevel theory, we provide tutorials on statistical models able to capture Influence and Selection processes. Specifically, we recommend modeling Influence processes with the Temporal Network

Autocorrelation Model (Dittrich et al., 2020; Frank & Fahrback, 1999; Friedkin, 2011). This Influence model is a straightforward technique that is more conceptually familiar to multilevel researchers compared to the use of centrality indices which often appear conceptually foreign to many researchers. Further, we recommend modeling Selection processes with Multiple Group Additive and Multiplicative Effects Network Models (Hoff, 2021). Beyond providing a tutorial, we employ a multilevel Bayesian approach to estimating this Selection model with multiple networks. This is a crucial extension for studying multiple independent teams. We provide R code for both models, equipping researchers with the tools necessary to empirically assess novel theory regarding Influence and Selection processes and enabling them to move beyond agreement, aggregation, and centrality.

In the following sections, we will first justify the need for a new theoretical and analytical framework by examining in depth the implications of the structural and nodal equivalence assumptions on multilevel theory. As part of this justification, we conduct a systematic review of existing social network techniques used in team research and the extent these techniques avoid these problematic assumptions. Second, we provide a conceptual outline and an analytic tutorial for both the Influence and Selection portions of the framework. Finally, we discuss the implications of the Influence and Selection framework on future theory testing, advancing organizational research methods, and stochastic actor oriented models (SAOM).

Insert Figure 1 about here

The Need for an Influence and Selection Framework

The present manuscript is focused on expanding theoretical lenses through introducing a framework and aligning this framework with accompanying analytic methodologies through a tutorial. To justify the need for this theoretical and analytical framework, we first articulate the blind spots caused by the predominant assumptions of structural and nodal equivalence. Second, we conduct a systematic review to demonstrate that Influence and Selection processes are rarely studied in team research as evidenced by the near exclusive use of certain analytic methodologies.

We proceed on the assumption that readers are familiar with the basics of MLM and SNA (see Raudenbush & Bryk, 2002; Light & Moody 2020). While the approach of our manuscript aligns more closely with SNA, the contribution of the manuscript has broad import for multilevel research. As most multilevel researchers are experienced with MLM, yet are not experts in SNA, the manuscript is intended to be understandable to researchers with varying levels of experience in both camps.

Structural and Nodal Equivalence Assumptions

Structural equivalence and multilevel modeling. Structural equivalence is present when members are equally connected to each other or have identical patterns of interaction. This implies all members with the same characteristics experience their team in the same way. Structural equivalence is frequently assumed in MLM (Roberson et al., 2007), such as when testing cross-level moderation (Raudenbush & Bryk, 2002). While team-level moderators may have differential effects depending on the value of individuals characteristics, out of necessity, this assumes that the team-level variable is itself equivalent for all members of the team, making all individuals with the same characteristics statistically indistinguishable and any deviations from the mean attributed to error. However, what is often dismissed as error is actually theoretically meaningful variance (DeRue et al., 2010; Mathieu & Chen, 2011). This means that the entrenchment of the structural equivalence assumption can limit the study of teams both theoretically and analytically.

To illustrate this, consider the emotional dynamics in teams (Barsade & Knight, 2015). A conventional MLM approach could first justify team affect as an emergent property and then examine the moderating effect on individual affect. This approach could assess research questions such as the role of team positive affect in blunting the effects of individual negative affect on job satisfaction. This approach assumes all individuals with similar affect experience the team in a similar way. However, team members may vary considerably in their exposure to the affect of different team members. In contrast, aligning more with a contagion process than a team climate effect, individuals may interact with some team members more than others leading to different experiences in otherwise comparable teams.

Nodal equivalence and social network analysis. Nodal equivalence is present when all members in a network have the same characteristics. Nodal equivalence is frequently assumed in SNA, such as when using structuralist methodologies (Borgatti & Foster, 2003). To utilize structuralist methodologies, data analysts frequently calculate structural indices that have gained popularity by team researchers in recent years (Park, Grosser, et al., 2020). Out of necessity, these indices assume all members who share similar structural positions are comparable (Borgatti & Foster, 2003). However, there is meaningful theoretical variance in nodal characteristics of network members. As a result, the resources that flow through ties are not addressed by structural indices. This means that the entrenchment of the nodal equivalence assumption can limit the study of teams both theoretically and analytically.

Consider further the emotional dynamics in teams (Barsade & Knight, 2015). A conventional SNA approach could calculate individual degree centrality in a team's social network. This approach could assess research questions such as the role of team social structures in blunting the effects of individual negative affect on job satisfaction. This approach assumes all individuals who occupy structurally similar positions will have similar outcomes, suggesting that the number and nature of social connections is of higher theoretical import than the nature of the people one is connected to. However, team members may vary considerably in their levels of affect. As mentioned in the prior example, change in members' negative affect is likely a function of both the strength of connections to other members as well as the characteristics (e.g., affect) of those members.

Review of Social Analytic Technique Usage in Multilevel Research

Entrenched assumptions of structural and nodal equivalence pose problems for testing developments in multilevel theory. To better gauge the need for a new framework and tutorial, we conducted a systematic review of existing SNA techniques in the team literature. Specifically, we first examined the extent to which existing SNA techniques are used in comparison to other techniques (i.e., the extent existing team research avoids the structural equivalence assumption by adopting a network approach). We then examined the extent to which existing SNA techniques can capture Influence and Selection processes (i.e., the extent existing team network research specifically avoids the nodal equivalence assumption).

We base our simple review of specific analytic techniques on a recent and robust review of networks in work teams (Park, Grosser, et al., 2020). Park and colleagues (2020) document the increased use of SNA among multilevel researchers. While increasing in usage, we wanted to gauge the proportion of research that avoids assuming structural equivalence via network approaches. To accomplish this, we slightly modified the search criteria used in this previous review. Specifically, we searched the same 20 journals from 1994 through 2018 in the Scopus database, however we only used “team” or “group” in our search of titles, abstracts, and keywords (i.e., omitted “network” and “social capital”). This yielded 5,477 articles compared to the 489 reported by Park and colleagues (2020). This means that less than one out of ten articles (8.93%) of the total research on teams in this sample referenced social networks. We interpret this as evidence that most team research does not avoid the structural equivalence assumption by adopting a network approach.

Parks and colleagues (2020) further identified that only 116 of these articles empirically assessed social networks within a team context. To assess the extent this existing work has examined Influence and Selection processes, we assessed in depth the types of analytic techniques utilized (see Table 1). Only 16 articles (13.79%) addressed Influence or Selection processes (see Table 2). Of these, only a single article (0.86%) examined Influence directly with 10 (8.62%) examining Selection directly. Five other articles (4.31%) either theorized about both Influence and Selection simultaneously or Selection only, however these relied on SAOMs which simultaneously estimates Influence and Selection. As we will discuss later, such approaches may benefit from separate theoretical development. Finally, the vast majority of the studies utilized a structural index-based approach (81.90%).

Insert Tables 1 and 2 about here

Thus, not only are social networks under-utilized compared to other analytic paradigms in multilevel research, but there is an over-reliance on a single tool within the analytic paradigm. We interpret this to mean most team network research assumes nodal equivalence. Based on this evidence, we assert this methodological myopia constrains multilevel researchers and prevents them from addressing the important theoretical processes of Influence and Selection. As such, we introduce a theoretical framework and provide an analytical tutorial centered on Influence and Selection.

The Influence and Selection Framework

Our proposed framework has both a theoretical component (i.e., Influence and Selection *processes*) and an analytical component (i.e., Influence and Selection *models*). We will initially address Influence and Selection separately. First, we outline what constitutes each process and provide examples. Second, we explain the fundamentals of our modeling approach. Third, we outline how our modeling approach refines existing approaches. Fourth, we provide a tutorial on how to run each model with accompanying R code.

The Influence Process

The Influence process, broadly defined, refers to the extent to which affect, behavior, or cognition of an individual changes as a function of their exposure to or interaction with others, making it one of the most important processes that occurs within collectives (see Figure 2). In essence, the Influence process focuses on the network as a mechanism through which people are affected by those they are connected to. Thus, Influence is a function of both the connections between people as well as the characteristics of those people. These characteristics include the attributes or behaviors that make people particularly influential as well as resilient to the sway of others.

There are numerous examples of Influence in organizational phenomenon. For instance, when a team socializes a new member, repeated interactions with the group will lead the newcomer to adopt a similar perspective of the shared task (Kozlowski et al., 2009). A second example includes coworkers' job search behavior and job embeddedness influencing other employees to resign (Felps et al., 2009). Through such contagion in coworkers' networks, departments experience increased amounts—or even waves—of turnover. A third example includes adopting a business practice either to stay competitive with those already practicing it in the same product market or to counter different competitors' practices (Levine et al., 2017). In summary, the processes of how individuals Influence each other is fundamental to the microfoundations of multilevel theory.

Insert Figure 2 about here

The Influence Model

To capture the Influence process, we recommend researchers use the Temporal Network Autocorrelation Model. This model is a linear model, providing considerable flexibility including estimation within a regression framework. For ease of edification, we will refer to this model as the "Influence Model" (IM). The IM is fundamentally concerned with predicting individual affect, behavior, or cognition accounting for localized network effects. In the following sections, we will outline how the IM addresses local, global, and historical effects as well as explain how our recommended modeling approach refines existing methods.

Fundamentals. Local effects in the IM are represented by an *exposure term* as described by Sun and colleagues (2013; see Figure 3 for the equation and a graphical representation). Individuals are exposed to the characteristics of those they are connected to—the stronger the connection, the stronger the exposure to those characteristics. This means exposure is dependent on the specific dyadic relationship between the individuals of interest, allowing the IM to move beyond the structural equivalence assumption. Further, a strong level of a characteristic coupled with a strong connection maximizes exposure. This allows the IM to move beyond the nodal equivalence assumption.

Although the IM primarily address local effects, it can also incorporate global (i.e., group) effects consistent with traditional MLM. Non-independence of errors among individuals from the same team are likely to occur when studying multiple teams. To account for nested data and accurately estimate standard errors, researchers can specify a random intercept for the team or use clustered standard errors. Further, researchers can examine higher level constructs in conjunction with the exposure term within a MLM framework (level-two and level-one equations respectively), with the exposure term being treated as any other individual-level construct (Frank, 2011). However, many applications of an IM do not need to incorporate global effects as they address phenomenon that occur within a single, large network.

Finally, the IM also address historical effects. Historical effects in an IM represent the extent to which the level of an individual characteristic is impacted by their previous level. An autoregressive term models this simply (see Figure 3). It is important to account for the previous level of the outcome of interest to confidently assess the occurrence of change or Influence. This basic approach to accounting for historical effects can be extended with more intensively longitudinal data, enabling a more dynamic evaluation of the stability or equilibrium of the characteristic (Somaraju et al., 2021).

Insert Figure 3 about here

Refinement of Existing Methods. Borgatti and Foster (2003) identified two different perspectives on the function of ties within the network paradigm in organizational research: structuralist and connectionist. The structuralist perspective focuses on the pattern of ties, with social connections viewed as analogous to “girders.” In contrast, the connectionist perspective focuses on the resources flowing through ties—placing more emphasis on the characteristics of the individual, with social connections viewed as analogous to “pipes.” The structuralist perspective is the more dominant of the two, largely due to the sociological roots of the network paradigm. However, Kilduff and Lee (2020, p. 160) strongly criticized this approach as it relies on the assumption that “network patterns… derive from social structure rather than human agency. Thus, structuralists “shun the ‘person’ construct as polluting” in their search for an individual-free science of networks (White 1992, p. 3).” Recognizing the negative impact this entrenched assumption could have on the study of people in organizations, recent calls have urged social network researchers to develop organizational theory that incorporates nodal characteristics (Kilduff & Lee, 2020).

However, the existing methods to test such theory remain entrenched in the structuralist approach. For example, Fang and colleagues (Fang et al., 2015) examined the mediating effect of network structure (i.e., indegree centrality and brokerage) on the relationship between personality (i.e., Big Five, self-monitoring) on work outcomes (i.e., job performance and career success). Taking a slightly different approach, Sherf and colleagues (2018) examined the moderating effect of the traits of the central voicer on the relationship between team network structure (i.e., voice centralization) on team performance. These approaches, while certainly commendable for incorporating nodal characteristics in their theories and models, rely heavily on centrality indices resulting in independent consideration of connections and characteristics.

In contrast, our proposed IM, via the exposure term, simultaneously considers the connections and characteristics. This simultaneous consideration shifts the focus of the empirical models more squarely on the resources flowing between ties, rather than the ties themselves (Borgatti & Foster, 2003). Thus, our approach facilitates the call for testing theory that incorporates nodal characteristics (Kilduff & Lee, 2020) by outlining a statistical approach more clearly aligned with theory.

Although they are almost completely missing from tests of multilevel theory in the organizational sciences (see Table 1), there are several models used by network scholars in other literatures capable of approximating Influence processes and incorporating nodal characteristics.

These approaches all estimate the network exposure on an individual relative to specific characteristics (e.g., negative affect) and connections (e.g., information sharing). Several approaches estimate network exposure using a single time point, such as Network Autocorrelation Models (Deittrich, 2020; Ord, 1975). However, due to non-independence of error and predictor, estimating exposure with a single time point is only possible by assuming the network members' characteristics are static (Deittrich, 2020; Ord, 1975). Unfortunately, methods for estimating these static network models are complex, reducing the accessibility and applicability of these models.

Estimation becomes much simpler when there is temporal separation between network exposure and the outcome. This temporal perspective is utilized by the Social Influence Network Model (Friedkin, 2011). However, while providing more flexibility in terms of estimation, the Social Influence Network Model poses restrictions reflective of a "closed system" assumption. Specifically, all variance is divided between the autoregressive and exposure terms. This means the total amount of a given characteristic is assumed to be stable in the network over time. For example, the total level of negative affect in a team is constant over time, only the balance between members' negative affect changes. While perhaps appropriate for finite resources, this assumption is not appropriate for the socio-psychological constructs and processes frequently studied in "open system" teams (Katz & Kahn, 1978).

By way of refinement to these existing approaches, we recommend a Temporal Network Autocorrelation Model (TNAM). This model both bypasses the concerns regarding non-independence of error by measuring across two time points and avoids making the "closed system" assumption by freely estimating the variance attributable to the autoregressive and exposure terms. Thus, researchers studying Influence processes can apply this model in a broad array of contexts. Additionally, Frank and Xu (2016, 2020), demonstrated mathematically that the temporal separation incorporated into TNAM overcomes many of the limitations of static models. In a follow up sensitivity analysis, Xu and Frank (2021) further demonstrated that TNAM can be estimated using ordinary least squares. This provides TNAM a considerable advantage over alternative approaches as it allows researchers to assess theoretically compelling influence phenomenon using basic regression techniques.

Influence Model Tutorial

Estimating the IM is a straightforward process that relies primarily on the exposure term (see Figure 4). The exposure term, similar to aggregate team variables, is calculated as the aggregation of several sources. While a team level construct in MLM is typically measured with team-referent reports and aggregated after demonstrating agreement (Bliese, 2000), the exposure term is an aggregate individual level construct that represents the extent to which the focal individual is exposed to the given a construct (usually not a team-referent) via the given dyadic network construct. Exposure is therefore a function of connections between individuals (e.g., other-reports) and characteristics of those individuals (e.g., self-reports). To differentiate between individuals, we follow SNA convention and refer to the focal individual as the “ego” and the individuals connected to the ego as “alters.” This distinction is comparable to the “rater/target” or “actor/partner” distinction made in dyadic MLM.

For this tutorial, we demonstrate a basic IM: a single exposure term, a single level of analysis, as well as weighted and non-symmetric ties. However, the IM can account for multiple exposure terms if they are sufficiently distinguishable (i.e., avoid excessive multi-collinearity). Further, once calculated, the exposure term is treated like any other individual-level variable. Meaning researchers can use the exposure term in conjunction with frequented multilevel techniques (e.g., mediation, cross-level moderation, etc.). Finally, the IM can account for both the weight and symmetry of ties. In the tutorial, we use weighted (i.e., the dyadic construct is on a continuous scale instead of a 0-1 binary, unweighted scale) and non-symmetric ties (i.e., if person A received information from person B, then person B did not necessarily receive information from person A) as many relational ties in teams are not binary or require reciprocal perceptions. This tutorial includes *R* code (2022) and utilizes the *tidyverse* library (Wickham et al., 2019). The code and data used in this example are included in a repository in the online supplemental materials: https://osf.io/qg9uc/?view_only=e6badc6eb9c74109a7a31c58e1d8d165.

Insert Figure 4 about here

Step 1: Research question and design. The first step is to identify an Influence research question, including hypotheses and constructs of interest. Identifying the constructs of interest informs how the study is designed to ensure the necessary data are collected to estimate the model. To estimate the model, at least two time points of data are needed on the dependent variable (*t-1* and *t*). The dependent variable in an IM is typically a characteristic of in individual, including any affect, behavior, or cognition. The exposure term is measured at time *t-1* using both an individual-level *characteristics* of alters (e.g., affect, behavior, cognition, or trait), and a dyad-level measure of *connection* between the ego and alters (e.g., information sharing, advice, friendship, etc. between ego and alter). This connection serves as the mechanism by which individuals (egos) are exposed to the given characteristic of their team members (alters). The exposure term is simply treated as an independent variable.

Importantly, the characteristic of alters that helps form an independent variable can either be the same or different from the characteristic of the ego that forms the dependent variable. If the characteristic is the same, this reflects one type of an Influence process: contagion (e.g., ego negative affect increases when exposed to alters' negative affect). If the characteristic is different, this reflects Influence processes more generally (e.g., ego job satisfaction decreases when exposed to alters' negative affect). Either way, to properly account for autoregression, the dependent outcome characteristics must be measured twice—even if it is not the construct incorporated into the exposure term. For instance, if a researcher was interested in exposure to negative affect as a predictor of job satisfaction rather than as a predictor of negative affect, then job satisfaction would be measured at two time points (*t-1* and *t*), with negative affect and information sharing being measured at one time point (*t-1*). This would expand the columns of variables in the "Node_Influence.CSV" file in Figure 4 from two columns to three.

For ease of edification, we continue the negative affect contagion example from earlier. In this example, the characteristic of interest, negative affect, is the same between the focal individual and network members (i.e., negative affect contagion); the connection of interest is receiving information. To test a simple hypothesis with these variables, the research design is required to gather individual data on negative affect from each network member at both time *t-1* and time *t* as well as dyadic data on information sharing at time *t-1* (see Figure 4). This data would allow the test of the following hypothesis:

> *Example Hypothesis 1: An individual's negative affect is positively related to their past exposure to the negative affect of team members from which they receive information.*

Step 2: Data cleaning and formatting. Once the data to test the hypotheses are collected, the data needs to be cleaned and formatted. The IM is an individual-level analysis that relies on the exposure term. In preparation for calculating the exposure term and running the analysis, researchers require two separate types of data. Consistent with social network terminology, one data set will include the individual level data and is referred to as the "node" data set in Figure 4. The other data set will include the dyadic data and is referred to as the "edge" data set in Figure 4. Due to the nested nature of the observations within teams, the node data set needs to include an identifier for individual and team (.i.e., Individual ID and Team ID in Figure 4) and the edge data set needs to include an identifier for ego and alter (i.e., Ego ID and Alter ID in Figure 4).

After the data is formatted correctly and saved as a .csv file, the required packages and data need to be imported into R. This can be done through the *read_csv* command from the *tidyverse* package:

```
# Import Packages
Install.packages("tidyverse")
library(tidyverse)
# Load Data
# Must set path to local directory where the data is stored
# Click: Session > Set Working Directory > Choose Directory
edge = read_csv("./data/edge_influence.csv")
node = read_csv("./data/node_influence.csv")
```

After importing the data into R, we merge the two data sets using the *left_join* function:

```
# Merge Data Sets to the Dyadic Level
merged_data = left_join(edge,node, by = c("alter_ID" =
     "individual_ID"))
```

The resulting merged data will look like the original edge data set with additional columns for the variables in the node data set representing the altar's[2] characteristics. In other words, the merged data is now at the dyadic level with individual level traits for the ego and alter included. This data format is labeled in the MLM domain as a person-pairwise data set (Kenny et al., 2006).

Step 3: Exposure term calculation. The exposure term is necessary to estimating the IM. The exposure term is simply the product of one's teammate's level on the construct of

[2] Note, the segment – by = c("alter_ID" = "individual_ID") – ensures that the new columns represent the alter characteristics not ego characteristics

interest and the strength of connection the focal individual has with that teammate, summed across all their teammates:

$$Exposure_i = \sum_{j \in N(i)} W_{ij} Y_j \tag{1}$$

Where W_{ij} represents the strength of the connection between team members i (ego) and j (alter), and Y_j represents person j's level on some construct of interest, Y (e.g., negative affect). In other words, the exposure for a given ego is the sum of their alters' node characteristics (Yj) weighted by the strength of the connections (Wij).

To calculate the exposure term, we must first calculate the dyadic components of exposure ($W_{ij}Y_j$). Calculating dyadic exposure requires multiplying the connection variable (e.g., "information_receipt_t0") with the alter's characteristic (e.g., "negative_affect_t0"). This can be done following *tidyverse* convention using the *mutate* function:

```
# Calculate the Dyadic Exposure
merged_data = merged_data %>% mutate(exposure_dyad =
     information_receipt_t0 * negative_affect_t0)
```

Calculating the exposure term requires summing the dyadic exposures across each ego. This can be done following *tidyverse* convention using the *group_by* and *summarize* functions:

```
# Calculate the Exposure Term
total_exposure_table = merged_data %>% group_by(ego_ID) %>%
     summarize(exposure_t0 = sum(exposure_dyad,na.rm =
     TRUE))
```

Importantly, the exposure term is not an interaction term; rather it is a distinct, individual-level, and network-based construct. Therefore, the best practices for interaction terms (i.e., requiring variables to be centered prior to estimation to remove interdependence and clarify interpretation; Cohen et al., 2003) do not apply to the IM. Indeed, centering the characteristic and connection components prior to calculating the exposure term will inhibit rather than clarify interpretation.

Once calculated, the exposure term needs to be merged back into the individual data set in preparation for analysis. This can be done using the *left_join* function:

```
# Merge Data Sets to the Individual Level
final_data = left_join(node,total_exposure_table, by =
     c("individual_ID" = "ego_ID"))
```

At this point, the exposure term and data set are at the individual-level. For additional details about the functions used in this section, see the *tidyverse* documentation (Wickham et al., 2019).

Step 4: Running statistical analyses. After the exposure term is calculated, analysts can treat it as any other individual-level variable. Meaning analysts can include the exposure term in ordinary least square regression. Running an IM in regression includes at least three variables: the dependent variable (e.g., ego negative affect at *t*), a lagged control variable[3] (e.g., ego negative affect at *t-1*), and the exposure term as an independent variable (e.g., information receipt from members with negative affect at *t-1*). This can be done using the *lm* function. To calculate standardized coefficients we use the *scale* function to standardize all three variables.

```
# Influence Model in OLS Regression
## Unstandardized
model_OLS = lm(negative_affect_t1 ~ negative_affect_t0 +
     exposure_t0, data = final_data)
## Standardized
model_OLS_std = lm(scale(negative_affect_t1) ~
     scale(negative_affect_t0) + scale(exposure_t0),
                    data = final_data)
```

Step 5: Interpreting results. Interpreting the results is straightforward and comparable to any other results for an individual level variable. The results can be accessed using the *summary* function, with confidence intervals calculated using the *confint.default* function. Lastly, we calculate the effect size for the exposure term using the Exposure Ration Index (ERI). The

[3] When modeling network processes, we note the importance of the autoregressive effect for the dependent variable due to their crucial role in controlling for potential confounds.

ERI is the ratio of the standardized exposure coefficient to the autoregression coefficient, with .10 indicating a small effect, .30 indicating a moderate effect, and .50 indicating a large effect (Frank et al., 2004). The coefficients are extracted using the *coef* function.

```
# View Results
summary(model_OLS
summary(model_OLS_std)
# Calculate Confidence Intervals
confint.default(model_OLS_std)
# Calculate Exposure Ratio Index
coef(model_OLS_std)["exposure_t0_std"]/coef(model_OLS_std)[
     "negative_affect_t0_std"]
```

To illustrate the interpretation, we return to our team negative affect contagion example and display the results of our model using a simulated data set.

Before interpreting the parameters, we first assess model fit and determine the OLS model simulated data is significant (R^2 = .053, F = 9.357, p <.001). Hypothesis 1 predicted an individual's negative affect is positively associated with exposure to the negative affect of team members that individual receives information from. The results are consistent with Hypothesis 1 as demonstrated by a significant positive exposure effect ($\beta =$ .119, $CI_{95\%}$ = [.008, .230], p = .036, ERI = .579). See Table 3 for all estimated parameters in the model. The effect size of the model is calculated using R^2, following standard conventions (R^2 = .053). The effect size for the exposure term is calculated using ERI, following conventions from Frank and colleagues (2004) and indicated a large effect size (ERI = .579).

Insert Table 3 about here

The Selection Process

Selection, broadly defined, refers to the factors that contribute to the formation, maintenance, and dissolution of network ties (see Figure 5). In this sense, the Selection process focuses on how and why individuals construct their social network in particular ways. For instance, individuals experience countervailing strivings for personal uniqueness as well as collective belonging (Smith & Lewis, 2011). This results in efforts to foster group solidarity with similar others as well as to seek diversity in relational and informational resources (Granovetter, 1973). Therefore, Selection processes include the factors that make people more or less likely to establish ties with dissimilar others. This has important implications, for example, in reducing errors that occur in seeking task-relevant information from others (Lu et al., 2012; van Knippenberg et al., 2004) and improving inclusivity initiatives in organizations (Simchi-Levi, 2020).

There are three foci in Selection processes. The formation of ties is concerned with the establishment of relationships, seeking resources, and emergent perceptions of others. For example, in self-managed teams, members develop perceptions of who is responsible for different leadership functions (Morgeson et al., 2010; Zhu et al., 2018). The maintenance of ties focuses on the stability and resiliency of relationships. For example, effective reparation efforts following a trust violation from a supervisor or an organizational infraction (Lewicki & Brinsfield, 2017). The dissolution of ties addresses how characteristics of the individual and surrounding network affect the severing of a tie. For example, relational factors contributing to an impasse in negotiations between a supplier and distributor ultimately resulting in the termination of the business relationship (Jang et al., 2018).

Insert Figure 5 about here

The Selection Model

Fundamentals. Like the IM, the Selection Model (SM) primarily addresses local effects but differs in that it focuses on the change in network ties (see Figure 6 for equation and graphical representation). The SM accounts for the characteristics of the ego and alter (i.e., first-order local effects) as well as their unique dyadic interaction (i.e., second-order local effects) on the probability of forming or trimming a tie. This approach also accounts for the broader interdependence within the network (i.e., third-order local effects) such as the propensity to form transitive relationships (i.e., the friend of my friend is also my friend) and an individual's network role (Hoff, 2009; Hoff et al., 2002). For example, the extent a relationship conflict tie forms between persons A and B is not just a function of the propensity of person A (ego effects) and person B (alter effects) to be in a conflict generally, and their unique relationship with each other (dyadic effects)—but also the broader interdependence in the network. For instance, if a relationship conflict tie exists between persons A and B and person B is friends with person C, this is likely to affect the extent a relationship conflict tie forms between persons A and C—despite the absence of an existing tie between them. Such third order transitive relationships (also known as Simmelian ties) have important impacts on leadership perceptions (Guo et al. 2021) and creativity (Wu et al., 2016) in teams, yet they are not incorporated into most dyadic statistical methods.

The SM can also account for global effects, by including team level constructs as predictors or cross-level moderators. Due to the emphasis of SMs on identifying factors leading to change, addressing historical factors are important for model interpretation. Similar to IMs, autoregressive terms generally should be used to model historical factors. Including previous network ties as a dyadic autoregressive predictor allows researchers to control for historical factors and understand the process of change more clearly.

Insert Figure 6 about here

Refinement of existing methods. In contrast to Influence processes, several models capable of approximating Selection processes exist in the organizational literature. From our systematic review of network techniques (see Table 1), we found that ERGMs are the most common. Notably, many applications of ERGMs are notoriously sensitive to specification of their network effects (Duxbury, 2021) and thus necessarily have a strong focus on structural phenomenon. Such a sociological focus on the structure of networks may be particularly relevant in a specific context or for a given research question. However, when the focus on Selection is more psychological in nature, we recommend alternative methods that better focus on the affect, behavior, and cognition of the individuals involved.

To this end, we recommend researchers use the Additive and Multiplicative Effects Network Model (AMEN; Hoff, 2021). Not only does it better align theretically with the Selection process, but AMEN has several analytic advantages over other similar network techniques. This advantage largely relates to accounting for the broader interdependence in the network through the use of latent factors (Hoff, 2009; Hoff et al., 2002), which other approaches frequently fail to account for. A deeper mathematical discussion of the differences between various approaches that can approximate Selection processes is beyond the scope of the present research and is found elsewhere in the literature (see Minhas et al. (2019) for a more exhaustive comparison).

Beyond SNA approaches, AMEN also refines conceptually similar MLM approaches. For example, the Social Relations Model and Actor-Partner-Interdependence Model (Kenny et al., 2006) focus on factors impacting dyadic variables and account for multiple layers of nesting in dyadic data structures. While similar, these approaches conceptualize dyads, not networks. As such they only account for first (i.e., ego and alter) and second order (i.e., dyad) effects, not the third order effects (i.e., latent factors). Indeed, AMEN is a direct extension of the Social Relations Model (Hoff, 2021) through the inclusion of the broader interdependence in the network. Thus, our recommended analytical approach is not conceptually foreign to SNA or MLM researchers, and yet provides substantive refinements over existing approaches.

Not only do our current efforts extend MLM and SNA approaches, but they also further refine AMEN specifically for the study of teams. AMEN, like many other network models (e.g., ERGMs, SAOMs), typically analyze all the data as a single, large network. In a given network, the absence of ties between members can be as informative as the presence of ties. However, multilevel researchers frequently study independent teams where the absence of ties between members of different teams is not informative. It is more appropriate to assess each network independently, rather than indicating missing data or the absence of a tie between members of different teams. Unfortunately, most existing network approaches are not equipped to study multiple networks simultaneously.

Therefore, we introduce a Bayesian updating approach to estimating AMEN models on multiple networks. This approach estimates a single model for all observed teams by sequentially evaluating each team independently and using the resulting posterior distribution as an informed prior for the evaluation of the next team. Each step essentially takes the value of regression coefficients that mostly likely would have resulted in the previous teams' data and adjusts it to incorporate the data from the current team. The result is a posterior distribution which appropriately provides estimates of the values for regression coefficients most likely to have resulted in the observed data from the teams. This allows team researchers to combine data collected across numerous teams into a single statistical model, allowing for tests of generalized inference in a way that has heretofore been difficult. This makes our multiple-group extension of AMEN ideal for studying Selection processes in teams. While we recognize there are multiple approaches that can approximate Selection processes, for ease of edification, we refer to our multiple-group extension of AMEN as “the Selection Model” henceforth.

Selection Model Tutorial

The SM follows a similar process as the IM (see Figure 7). However, due to the complexity of the SM, we developed an R package for our multiple-group additive multiplicative effects approach (*MGAME*). This package simplifies and streamlines the process of running and interpreting the model. As before, the code and data for this tutorial are provided in the supplemental material: https://osf.io/qg9uc/?view_only=e6badc6eb9c74109a7a31c58e1d8d165.

Insert Figure 7 about here

Step 1: Research question and design. The first step is to identify a Selection research question, including hypotheses and constructs of interest. To illustrate how these factors affect the formation, maintenance, or dissolution of ties, we will continue the information sharing and negative affect in teams example. The SM can assess factors that predict the formation of information sharing ties in a network. For instance, egos with negative affect may be less likely to share information, while alters with negative affect may be less likely to have information shared with them. In contrast, information sharing may be attributable to the unique interaction between an ego and alter. An “opposites attract” theory suggests a team member is more likely to share information with someone with the opposite levels of negative affect, while a “birds of a feather flock together” theory suggests a team member is more likely to share information with someone with the same levels of negative affect. While the SM can test each of these four hypotheses, for sake of simplicity we will illustrate just one:

> *Example Hypothesis 2: The amount of information an individual shares with a team member is positively related to the similarity in their level of negative affect.*

While the estimation of the SM is possible with only one time point, we recommend at least two time points to better align with the theorized Selection process and control for history effects. In contrast to the IM, which requires two time points of “node” data and one time point of “edge” data, the SM requires one time point of “node” data and two time points of “edge” data (see Figure 7). This temporal separation is essential to studying change in the network structure and eliminating potential confounds.

Step 2: Data cleaning and formatting. Once the data to test the hypothesis are collected, we need to clean and format the data. We illustrate two data sets in Figure 7 to reinforce the necessary data structure for the SM. However, the statistics package only requires a single data set in person-pairwise format. Therefore, we only load the merged person-pairwise data set (Merge_Selection.CSV), which is created using the same process outlined in Step 2 of the IM tutorial. The R package, *MGAME*, can be downloaded and the data loaded using *devtools*:

```
# Install and Load MGAME from GitHub Directly
install.packages("devtools")
devtools::install_github("[MASKED FOR BLIND REVIEW]/MGAME",
upgrade = "never")
library(MGAME)
library(tidyverse)
```

Next, we load the data using the read_csv function from the tidyverse package.

```
# Load Data
# Must set path to local directory where the data is stored
# Click: Session > Set Working Directory > Choose Directory
data = read_csv("./data/selection_data.csv")
```

Step 3: Relational covariate calculation. Although not required for all SMs, our hypothesis includes a relational covariate. Relational covariate in SMs include any dyadic or team variables. For our example Hypothesis 2, we need to calculate the dyadic similarity in negative affect:

```
# Calculate Relational Covariates
```

```
     data = data %>% mutate(negative_affect_sim_t0 = 5 -
abs(negative_affect_ego_t0 - negative_affect_alter_t0))
```

This code captures similarity by first calculating the absolute difference between the ego and alter on a given individual characteristic (i.e., negative affect), and then subtracting this measure of dissimilarity from the maximum posible value of negative affect (e.g., five on a Likert-type scale). Similarly, if not done previously, at this point researcher need to calculate variables for any interaction terms used in there analysis including dyadic or cross level interactions.

Step 4: Running statistical analyses. After the relational covariates are calculated, using the *MGAME* package is straightforward. First, specify the criterion (*Y*) and the level of each predictor (*Xego*, *Xalter*, and *Xdyad*). In addition to dyadic predictors, the *Xdyad* line is where to specify the autoregressive term for the criterion as well as any team-level predictors. Second, specify the grouping variable, which is the identifier for a specific network (e.g., Team ID). Third, indicate which variables (if any) to center and standardize at the group level (*group_standard*) and/or the data set level (*grand_standard*).

Note that to get interpretable standardized coefficients for a given variable both the given predictor and the dependent variable must be group-mean or grand-mean standardized. While group-mean centering and standardizing may be appropriate with certain variables or contexts, due to the implications for structural equivalence, we generally recommend grand-mean centering and standardizing variables for the SM.

Lastly, specify other model options. We specify the number of latent factors (R = 2), the number of iterations (*nscan* = 10000), and the nature of the dependent variable (e.g., *family* = “nrm” for a continuous DV). The rationale for these decisions, as well as different options available for these models, are outlined by Hoff (2015; see also the *help* function for this package). The model can be estimated using the *mgame* command:

```
# Run Model
fit = mgame(
   data = data,
   Y = "info_share_t1",
   Xego = c("negative_affect_ego_t0"),
   Xalter = c("negative_affect_alter_t0"),
   Xdyad = c("info_share_t0", "negative_affect_sim_t0"),
   group = "team_ID",
   grand_standard = c("negative_affect_ego_t0",
      "negative_affect_alter_t0", "negative_affect_sim_t0",
      "info_share_t0", "info_share_t1"),
   R = 2,
   nscan = 10000,
   family = "nrm")
```

Step 5: Interpreting results. Interpreting the results is also straightforward. As with the models estimated in the Influence tutorial, a focused output—comprising of model fit indices, parameter estimates, proportion of variance explained, and variances—can be requested using the *summary* command:

```
# Get Results
summary(fit)
```

To illustrate the interpretation of this simulated data, we return to our information sharing and negative affect example. Before interpreting the parameters, we must first assess model fit. Assessing model fit per the guidelines of Hoff (2015) is complicated and requires considerable expertise to make viable visual judgment calls across five indices. These indices represent the extent to which the observed data fits the resulting model in terms of distinct factors. These factors represent two first order effects (i.e., ego and alter effects), one dyadic effect (i.e., reciprocity), and two higher order effects (i.e., transitivity and cycle-closure). Values for each effect are simulated in conjunction with the Bayesian estimation procedure (see, Hoff, 2009, 2015 for detail).

We extend Hoff's work, simplifying his process considerably by generating a single numerical goodness-of-fit value that we refer to as the Structural Selection Fit Index (SSFI).

Calculating the SSFI involves calculating the team score for each of the five factors, averaging across teams to create the sample score for each of the five factors, and then taking the minimum of those five factors. First, the observed value for the five fit factors is compared against simulated values for each team. We then compare the absolute distance of the simulated values from the simulated value mean and the distance of the observed values from the simulated value mean. For example, a value of .75 in the transitivity factor indicates the given team has an observed propensity for transitivity closer to the mean of the simulations than 75% of the simulated propensity for transitivity values. The closer this value is to 1, the better the model recreates the structure found in the observed data from the given team. Second, the five values for each team are then averaged across teams to create an index for each of the five factors, representing the proportion of observed team network structures that fit the model at least as well as the simulated structures.

Third, to standardize assessing fit and to make the interpretation more accessible, we use the minimum across the five factors as a single, conservative omnibus index of structural goodness-of-fit. The greater the index value is, the more effectively the model recreates the structure of the observed data. We propose using .10 as a threshold indicating poor fit given this indicates that the observed data fits worse than more than nine in ten simulated cases. In our example, our model demonstrates acceptable fit (SSFI = .249) and we can now interpret individual parameters

Hypothesis 2 predicted the extent an individual shares information with a team member is positively related to the similarity in their level of negative affect. The results of the dyadic similarity in negative affect are consistent with Hypothesis 2: β = .321, $CI_{95\%}$ = (.133, .474), $p = .001$. See Table 4 for all estimated parameters in the model. We determined the effect size of the overall model using pseudo R^2 which is presented beside the fit indices (Pseudo R^2 = .162). Following conventions described by LaHuis et al., (2014), Pseudo R^2 (also referred to as R^2_1(*approx.*)) is calculated by comparing the variance in the outcome explained by a full model (i.e., the model with all the predictors) with the variance explained by a null model (i.e., the model with random additive and multiplicative effects, or latent factors, but no predictors).

Lastly, the SM output provides the necessary data to decompose the variance. This can be done following the process addressed in detail in Hoff (2009). Variance decomposition informs researchers how much variance in the dependent variable is generally attributable to ego, alter, and dyad effects. Researchers might use this to investigate theoretical implications of the model. For example, it may be theoretically important to note that a larger portion of the variance is accounted for by unique ego vs. alter effects. Notably, the multi-group AMEN model does not incorporate a random group effect. This is consistent with existing methodological findings that suggest that when accounting for dyadic and individual contributions, a group random effect becomes redundant (Kenny & La Voie, 1984; Nestler, 2018) and as such may result in convergence issues.

Insert Table 4 about here

Discussion

To advance multilevel theory, researchers have identified that “network approaches… may prove valuable for generating alternative paradigms on which new multilevel quantitative investigations and methodologies could be advanced” (Mathieu & Chen, 2011; p. 626). We identified the limitations and prevalence of structural and nodal equivalence assumptions in multilevel research. To redress this, we introduced a novel theoretical and analytical framework centered on the network processes of Influence and Selection. This framework has important implications for multilevel theory and methods.

Contributions

Theory. The Influence and Selection framework provides an alternative paradigm to conventional approaches to multilevel phenomenon. For instance, individuals are not influenced solely by a collective emergent state, rather team members influence each other to varying extents over time. Further, network patterns are not derived entirely by a broader social structure, rather the psychology of actors plays a critical role in forming network ties. This approach moves away from a purely sociological space and more squarely positions network processes in a socio-psychological domain. In doing so, like historical developments in multilevel theory, network processes are poised to advance research on the microfoundations of multilevel phenomenon.

Indeed, Influence and Selection at the dyadic level are similar in meaningful ways to the processes and emergent states at the team level (Marks et al., 2001). Selection pertains to the decision to interact with individuals in a particular way, while team processes reflect the nature of interdependent interactions. Influence pertains to the effect the alters have on the ego through interactions, while emergent states reflect shared properties of the team that arise from interactions. Network processes undoubtedly play an important role in the emergence of collective constructs (Kozlowski & Klein, 2000), meaning Influence and Selection may undergird numerous processes and emergent states.

By providing a parsimonious way to conceptualize network processes in collectives, the Influence and Selection framework has the potential to inspire new theory and research in various domains. In this manuscript alone we generated examples involving numerous phenomena, including leadership, trust, decision-making, self-managed teams, socialization, negotiations, inclusivity initiatives, career decisions, turnover contagion, and diffusion of innovation to name a few. Indeed, the team conflict literature has begun to embrace the microfoundations of multilevel phenomenon with noteworthy success (Humphrey et al., 2017; Park, Mathieu, et al., 2020; Shah et al., 2021). While we have focused on team research, as this literature dominates the work on multilevel theory (Kozlowski & Bell, 2013), Influence and Selection are also relevant within larger collectives, such as multiteam systems, department dynamics, and inter-organizational interactions. Importantly, once new theory is developed, there are readily available and clearly aligned statistical models to facilitate empirical assessment.

Methods. The present framework does not require researchers to pick between assuming structural or nodal equivalence. Rather, our recommended approach bypasses both problematic assumptions through techniques used in numerous other fields, and yet remains to be fully utilized in the organizational sciences. These techniques are conceptually familiar to multilevel researchers who utilize MLM. This contrasts with the structural indices of SNA, which often appear conceptually foreign to multilevel researchers at first blush. This initial lack of familiarity might partially explain the apparent reluctance for the field to adopt SNA techniques more robustly. The accessibility of IM and SM may increase the likelihood of researchers incorporating the advantages of SNA into the study of multilevel phenomenon in organizations.

Beyond providing a tutorial on powerful, but largely unfamiliar and underutilized techniques, we also provide refinements to existing methodologies. First, we provide objective indices to aid in standardizing the assessment of these models: the ERI for assessing parameter effect sizes in the IM and the SSFI for assessing model fit in the SM. The SSFI, in contrast to previous approaches to assessing model fit in these models, is a conservative index utilizable by researchers of various degrees of SNA expertise. Second, we developed a Bayesian updating approach to study multiple networks. This approach is a methodological advancement for network models generally and is crucial for the study of network processes in teams. Finally, we introduce the *MGAME* package in *R* to aid researchers in running these models.

Limitations and Future Directions

Theory. Like all research, what we present here has limitations that need to be addressed by future research. Our framework presents Influence and Selection as separate process and identify models to assess them separately. We recognize Influence and Selection are related processes, however the seminal work of Simon and Ando (1961) demonstrates it is often ideal to assess interrelated processes separately when they are distinguishable by time period and timescale. In many instances Influence and Selection processes occur at different time periods and on different timescales (Xu & Frank, 2016). For example, the development of close relationships with work colleagues (i.e., Selection) may take years to fully develop, while the diffusion of a new, innovative work practice among close colleagues (i.e., Influence) could occur on a much faster timescale. Thus, we recommend researchers strongly consider empirically assessing Influence and Selection processes separately.

Recommending separate assessment has implications for the use of SAOMs (Kalish, 2020). SAOMs are powerful tools able to simultaneously account for changes in individual characteristics (i.e., Influence) and network connections (i.e., Selection). On the one hand, this explicitly recognizes that the process are interrelated. On the other hand, estimates of one process might be influenced by the specification and estimation of the other. This has the potential to obfuscate interpretability and hinder theory testing efforts. Further, when there is little existing theoretical and empirical work about a phenomenon, predominantly examining Influence and Selection simultaneously may hinder knowledge accumulation. We recommend the use of

SAOMs to simultaneously estimate Influence and Selection processes only when there is strong reason to believe Influence and Selection are occurring at the same time and timescale as well as strong theory about how the Influence and Selection processes are unfolding.

Ideally, this strong reasoning and theory is supported by previous empirical research. However, the limited extent of existing empirical research on Influence and Selection processes in teams (see Table 1) presents challenges to providing such justifications in a compelling way. Even in instances when the characteristics and connection are changing at the same time and on the same timescale, we recommend researcher run models of each process separately prior to running them simultaneously to ensure the theoretical mechanisms are unfolding as predicted. This practice can help avoid misinterpreting SAOM results and further clarify the relationship between Influence and Selection processes.

While separately assessing Influence and Selection frequently facilitates theory testing, jointly considering Influence and Selection frequently strengthens theory building. Future research is needed to better understand how and why these network processes are related. Developing theory on the relationship between these processes hold the potential to spur research into exciting, new directions. Further, understanding the relationship between these processes would help theorists identify when one process is more central than another in a given context.

Methods. There are important methodological refinements to be made to both the IM and SM. One area for future extension for the IM is the incorporation of latent factors. While we identified latent factors as important in the estimation of Selection processes and included them in the SM, latent factors are likely also important in the estimation of Influence processes. Future work could incorporate a latent factor extension into the IM. One area for refinements to the SM is greater work on assessing model fit. We introduced the SSFI as a conservative index of model fit. Future work should investigate the implications for this index compared to more holistic approaches. Further, methodologists could explore the implications of the different benchmarks we propose here to refine recommendations to researchers.

Future methodological research could include more dynamic extensions of both the IM and SM, as we include only two time points in our example. However, integrating these models with more sophisticated longitudinal methodologies could allow researchers to assess Influence and Selection at a more granular timescale, resulting in a higher resolution understanding of network processes. This is especially important considering advancements in less resource intensive means to capture social network data (e.g., wearable sensors, text analysis, and facial recognition software; Mathieu et al., 2022; Matusik et al., 2019; Park, Grosser, et al., 2020). There is a need for the development of methodological tools to analyze this intensive longitudinal data about the important network processes of Influence and Selection.

Conclusion

We outline a framework centered on two key network processes and specify empirical models to study these processes. This framework answers calls for the development of network approaches more focused on individuals in collectives rather than pure social structure (Kilduff & Lee, 2020; Mathieu & Chen, 2011). This emphasis on Selection and Influence avoids the structural equivalence assumption of MLM as well as the nodal equivalence assumption of typical SNA approaches. Our hope is that, equipped with these tools, researchers are better able to address the microfoundations of multilevel theory and advance our understanding of the social dynamics in organizational phenomena.

References

Aguinis, H., Gottfredson, R. K., & Culpepper, S. A. (2013). Best-practice recommendations for estimating cross-level interaction effects using multilevel modeling. *Journal of Management, 39*(6), 1490-1528.

Barsade, S. G., & Knight, A. P. (2015). Group affect. *Annual Review of Organizational Psychology and Organizational Behavior,* 2(1), 21-46.

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, Chapter xxix, 605 Pages). Jossey-Bass, San Francisco, CA.

Bono, J. E., & Judge, T. A. (2004). Personality and transformational and transactional leadership: a meta-analysis. *Journal of Applied Psychology, 89*(5), 901.

Borgatti, S. P., & Foster, P. C. (2003). The network paradigm in organizational research: A review and typology. *Journal of Management*, *29*(6), 991-1013.

Brewer, M. B. (1991). The social self: On being the same and different at the same time. *Personality and Social Psychology Bulletin, 17*(5), 475-482.

Cable, D. M., & Edwards, J. R. (2004). Complementary and supplementary fit: A theoretical and empirical integration . *Journal of Applied Psychology*, *89*(5), 822-834.

Chatman, J. A. (1991). Matching People and Organizations: Selection and Socialization in Public Accounting Firms Jennifer A. Chatman. *Administrative Science Quarterly, 36*, 459-484.

Cox, T. H., Lobel, S. A., & McLeod, P. L. (1991). Effects of ethnic group cultural differences on cooperative and competitive behavior on a group task . *Academy of Management Journal, 34*(4), 827-847.

De Dreu, C. K. W., & Weingart, L. R. (2003). Task versus relationship conflict, team performance, and team member satisfaction: A meta-analysis . *Journal of Applied Psychology*, *88*(4), 741-749.

de Wit, F. R. C., Greer, L. L., & Jehn, K. A. (2012, 03//). The paradox of intragroup conflict: A meta-analysis . *Journal of Applied Psychology*, *97*(2), 360-390.

DeRue, D. S., & Ashford, S. J. (2010). Who will lead and who will follow? A social process of leadership identity construction in organizations. *Academy of Management Review*, *35*(4), 627-647.

Dittrich, D., Leenders, R. T. A., & Mulder, J. (2020). Network autocorrelation modeling: Bayesian techniques for estimating and testing multiple network autocorrelations. *Sociological methodology*, *50*(1), 168-214.

Duxbury, S. W. (2021). Diagnosing multicollinearity in exponential random graph models. *Sociological Methods & Research*, *50*(2), 491-530.

Emich, K. J. (2020). Well, I feel differently: The importance of considering affective patterns in groups. Social and Personality Psychology Compass, 14(4), e12523.

Emich, K. J., & Vincent, L. C. (2020). Shifting focus: The influence of affective diversity on team creativity. *Organizational Behavior and Human Decision Processes*, *156*, 24-37.

Fang, R., Landis, B., Zhang, Z., Anderson, M. H., Shaw, J. D., & Kilduff, M. (2015). Integrating personality and social networks: A meta-analysis of personality, network position, and work outcomes in organizations. *Organization Science*, *26*(4), 1243-1260.

Feldman, D. C., & Ng, T. W. (2007). Careers: Mobility, embeddedness, and success. *Journal of Management*, *33*(3), 350-377.

Felps, W., Mitchell, T. R., Hekman, D. R., Lee, T. W., Holtom, B. C., & Harman, W. S. (2009). Turnover contagion: How coworkers' job embeddedness and job search behaviors influence quitting. *Academy of Management Journal, 52*(3), 545-561.

Frank, K. A., & Xu, R. (2020). Causal inference for social network analysis. *The Oxford Handbook of Social Networks*, 288-310.

Frank, K. A. and K. Fahrbach (1999). "Organization culture as a complex system: Balance and information in models of influence and selection." *Organization Science* 10(3), 253-277.

Frank, K. A., Zhao, Y., & Borman, K. (2004). Social capital and the diffusion of innovations within organizations: The case of computer technology in schools. *Sociology of education*, *77*(2), 148-171.

Granovetter, M. (1985). Economic action and social structure: The problem of embeddedness. *American Journal of Sociology, 91*(3), 481-510.

Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology, 78*(6), 1360-1380.

Guo, Z. A., Heidl, R. A., Hollenbeck, J. R., Yu, A., & Howe, M. (2021). When discretionary boundary spanning relationships cease becoming discretionary: The impact of closed ties on informal leadership perceptions. *Journal of Applied Psychology*.

Heider, F. (1958). *The Psychology of Interpersonal Relations*. John Wiley and Sons.

Hirschman, A. O. (1970). *Exit, voice, and loyalty: Responses to decline in firms, organizations, and states* (Vol. 25). Harvard university press.

Hoff, P. (2021). Additive and multiplicative effects network models. *Statistical Science, 36*(1), 34-50.

Hoff, P. D. (2009). Multiplicative latent factor models for description and prediction of social networks. *Computational and Mathematical Organization Theory, 15*(4), 261.

Hoff, P. D. (2015). Dyadic data analysis with amen.

Hoff, P. D., Raftery, A. E., & Handcock, M. S. (2002). Latent space approaches to social network analysis. *Journal of the American Statistical Association, 97*(460), 1090-1098.

Hoffman, L. (2015). *Longitudinal Analysis: Modeling Within-Person Fluctuation and Change*. Routledge.

Hollenbeck, J. R., DeRue, D. S., & Nahrgang, J. D. (2015). The opponent process theory of leadership succession. *Organizational Psychology Review, 5*(4), 333-363.

Humphrey, S. E., Aime, F., Cushenbery, L., Hill, A. D., & Fairchild, J. (2017). Team conflict dynamics: Implications of a dyadic view of conflict for team performance. *Organizational Behavior and Human Decision Processes, 142*, 58-70.

Jackson, S. E., Brett, J. F., & Sessa, V. I. (1991, 10//). Some differences make a difference: individual dissimilarity and group heterogeneity as correlates of recruitment, promotions, and turnover . *Journal of Applied Psychology, 76*, 675-689.

Jang, D., Elfenbein, H. A., & Bottom, W. P. (2018). More than a phase: Form and features of a general theory of negotiation . *Academy of Management Annals, 12*(1), 318-356.

Jehn, K. A. (1995). A multimethod examination of the benefits and detriments of intragroup conflict. *Administrative Science Quarterly*, 256-282.

Jehn, K. A., Rispins, S., & Thatcher, S. M. (2010). The effects of conflict asymmetry on workgroup and individual outcomes. *Academy of Management Journal, 53*, 596-616.

Kalish, Y. (2020). Stochastic actor-oriented models for the co-evolution of networks and behavior: An introduction and tutorial. *Organizational Research Methods, 23*(3), Kenny, D. A., Kashy, D. A., & Cook, W. L. (2006). Social relations designs with indistinguishable members. In *Dyadic data analysis*. Guilford press.

Kenny, D. A., & La Voie, L. (1984). The social relations model. In *Advances in experimental social psychology* (Vol. 18, pp. 141-182). Elsevier.

Kilduff, M., & Lee, J. W. (2020). The integration of people and networks. *Annual Review of Organizational Psychology and Organizational Behavior, 7*, 155-179.

Klein, K. J., Bliese, P. D., Kozolowski, S. W. J., Dansereau, F., Gavin, M. B., Griffin, M. A., Hofmann, D. A., James, L. R., Yammarino, F. J., & Bligh, M. C. (2000). Multilevel analytical techniques: Commonalities, differences, and continuing questions. In K. J. Klein & S. W. J. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions* (pp. 512-553). Jossey-Bass, San Francisco, CA.

Kozlowski, S. W. J., & Bell, B. S. (2013). Work groups and teams in organizations. In W. C. Borman, D. R. Ilgen, & R. J. Klimoski (Eds.), *Handbook of psychology: Industrial and organizational psychology, Vol. 12* (pp. 333-375). John Wiley & Sons Inc, Hoboken, NJ.

Kozlowski, S. W. J., & Hattrup, K. (1992). A disagreement about within-group agreement: Disentangling issues of consistency versus consensus . *Journal of Applied Psychology, 77*(2), 161-167.

Kozlowski, S. W. J., Watola, D. J., Jensen, J. M., Kim, B. H., & Botero, I. C. (2009). Developing adaptive teams: A theory of dynamic team leadership. In E. Salas, G. F. Goodwin, & C. S. Burke (Eds.), *Team effectiveness in complex organizations: Cross-disciplinary perspectives and approaches* (pp. 113-155, Chapter xxxiv, 589 Pages). Routledge/Taylor & Francis Group, New York, NY.

Levine, S. S., Bernard, M., & Nagel, R. (2017). Strategic intelligence: The cognitive capability to anticipate competitor behavior. *Strategic Management Journal, 38*(12), 2390-2423.

Lewicki, R. J., & Brinsfield, C. (2017). Trust repair. *Annual Review of Organizational Psychology and Organizational Behavior, 4*, 287-313.

Light, R., & Moody, J. (Eds.). (2020). *The Oxford handbook of social networks*. Oxford University Press.

Lu, L., Yuan, Y. C., & McLeod, P. L. (2012). Twenty-five years of hidden profiles in group decision making: A meta-analysis. *Personality and Social Psychology Review, 16*(1),

Lucas, B. J., Galinksy, A. D., & Murnighan, K. J. (2016). An intention-based account of perspective-taking: why perspective-taking can both decrease and increase moral condemnation. *Personality and Social Psychology Bulletin, 42*(11), 1480-1489.

Marks, M. A., Mathieu, J. E., & Zaccaro, S. J. (2001). A temporally based framework and taxonomy of team processes . *Academy of Management Review, 26*(3), 356-376.

Mathieu, J. E., & Chen, G. (2011). The etiology of the multilevel paradigm in management research. *Journal of Management, 37*(2), 610-641.

Mathieu, J. E., Hollenbeck, J. R., van Knippenberg, D., & Ilgen, D. R. (2017). A century of work teams in the Journal of Applied Psychology. *Journal of Applied Psychology, 102*(3),

Mathieu, J. E., Wolfson, M. A., Park, S., Luciano, M. M., Bedwell-Torres, W. L., Ramsay, P. S., Klock, E. A., & Tannenbaum, S. I. (2022). Indexing dynamic collective constructs using computer-aided text analysis: Construct validity evidence and illustrations featuring team processes. *Journal of Applied Psychology*.

Matusik, J. G., Heidl, R., Hollenbeck, J. R., Yu, A., Lee, H. W., & Howe, M. (2019). Wearable bluetooth sensors for capturing relational variables and temporal variability in relationships: A construct validation study. *Journal of Applied Psychology, 104*(3), 357.

Minhas, S., Hoff, P. D., & Ward, M. D. (2019). Inferential approaches for network analysis: Amen for latent factor models. *Political Analysis, 27*(2), 208-222.

Morgeson, F. P., DeRue, D. S., & Karam, E. P. (2010). Leadership in teams: A functional approach to understanding leadership structures and processes . *Journal of Management, 36*(1), 5-39.

Nestler, S. (2018). Likelihood estimation of the multivariate social relations model. *Journal of Educational and Behavioral Statistics, 43*(4), 387-406.

Park, S., Grosser, T. J., Roebuck, A. A., & Mathieu, J. E. (2020). Understanding work teams from a network perspective: a review and future research directions. *Journal of Management, 46*(6), 1002-1028.

Park, S., Mathieu, J. E., & Grosser, T. J. (2020). A network conceptualization of team conflict. *Academy of Management Review, 45*(2), 352-375.

R Core Team (2022). *R: A language and environment for statistical computing.* In R Foundation for Statistical Computing. https://www.R-project.org/

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). sage.

Roberson, Q. M., Sturman, M. C., & Simons, T. L. (2007). Does the measure of dispersion matter in multilevel research? A comparison of the relative performance of dispersion indexes. *Organizational Research Methods*, *10*(4), 564-588..

Schmidt, F. L., & Hunter, J. E. (1989). Interrater reliability coefficients cannot be computed when only one stimulus is rated . *Journal of Applied Psychology, 74*(2), 368.

Shah, P. P., Peterson, R. S., Jones, S. L., & Ferguson, A. J. (2021). Things are not always what they seem: the origins and evolution of intragroup conflict. *Administrative Science Quarterly*, 0001839220965186.

Sherf, E. N., Sinha, R., Tangirala, S., & Awasty, N. (2018). Centralization of member voice in teams: Its effects on expertise utilization and team performance. *Journal of Applied Psychology*, *103*(8), 813.

Simchi-Levi, D. (2020). From the Editor: Diversity, Equity, and Inclusion in Management Science. *Management Science, 66*(9), 3802-3802.

Simon, H. A., & Ando, A. (1961). Aggregation of variables in dynamic systems. *Econometrica: journal of the Econometric Society*, 111-138.

Smith, W. K., & Lewis, M. W. (2011). Toward a theory of paradox: A dynamic equilibrium model of organizing . *Academy of Management Review, 36*(2), 381-403.

Sun, M., Frank, K. A., Penuel, W. R., & Kim, C. M. (2013). How external institutions penetrate schools through formal and informal leaders. *Educational Administration Quarterly*, *49*(4), 610-644.

Turner, J. C. (1985). Social categorization and the self-concept: A social cognitive theory of group behavior. In E. J. Lawler (Ed.), *Advances in group processes: Theory and research* (Vol. 2, pp. 77-122). JAI Press.

Turner, M. E., & Pratkanis, A. R. (1998). A social identity maintenance model of groupthink . *Organizational Behavior & Human Decision Processes, 73*(2/3), 210-235.

van Knippenberg, D., De Dreu, C. K. W., & Homan, A. C. (2004). Work group diversity and group performance: An integrative model and research agenda . *Journal of Applied Psychology*, *89*(6), 1008-1022. https://doi.org/10.1037/0021-9010.89.6.1008

Wasserman, S., & Faust, K. (1994). Social network analysis in the social and behavioral sciences. In *Social network analysis: Methods and applications* (pp. 3-27). Cambridge University Press.

Wasserman, S., & Pattison, P. (1996). Logit models and logistic regressions for social networks: I. An introduction to Markov graphs andp. *Psychometrika*, *61*(3), 401-425.

Westphal, K. R. (2003). *Hegel's epistemology: a philosophical introduction to the Phenomenology of spirit*. Hackett Publishing.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D. A., François, R., Grolemund, G., Hayes, A., Henry, L., & Hester, J. (2019). Welcome to the Tidyverse. *Journal of open source software*, *4*(43), 1686.

Wu, Y., Chang, K., & Sha, X. (2016). Creative performance in the workplace: The roles of Simmelian ties and communication media. *Computers in Human Behavior*, *63*, 575-583.

Xu, R., & Frank, K. A. (2016). The Implications for Network Structure of Dynamic Feedback between Influence and Selection. International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation,

Xu, R., & Frank, K. A. (2021). Sensitivity analysis for network observations with applications to inferences of social influence effects. *Network Science*, 9(1), 73-98.

Zhu, J., Liao, Z., Yam, K. C., & Johnson, R. E. (2018). Shared leadership: A state-of-the-art review and future research agenda. *Journal of Organizational Behavior*, *39*(7), 834-852.

Table 1

Publications in Organizational Science Outlets Employing Social Network Analysis from 1994-2018

Focus of Analysis	Total	Count	Percentage
1. Influence	1		.86%
Variation of Influence Model		1	
2. Selection	10		8.62%
ERGM		5	
QAP		1	
*P**		1	
Other		3	
3. Both (SAOM)	5		4.31%
Theorized Influence and Selection		3	
Theorized Selection Only		2	
4. Centrality	95		81.90%
5. Other (Computational Model, etc.)	5		4.31%
Total	116		100.00%

Note. Numbers were obtained by tabulating the methods reported from 116 empirical articles identified in a recent systematic review that identified studies that use social network analysis in team contexts across 20 journals (Park, Grosser, et al., 2020).

Table 2

Publications in Organizational Science Outlets That Study Influence and Selection Processes

Authors	Year	Journal
Influence		
1. Friedkin, N.E.	2011	Administrative Science Quarterly
Selection		
2. Schecter, A., Pilny, A., Leung, A., Poole, M. S., & Contractor, N.	2017	Journal of Organizational Behavior
3. Brennecke, J., & Rank, O. N.	2016	Social Networks
4. Lomi, A., Lusher, D., Pattison, P. E., & Robins, G.	2014	Organization Science
5. Lusher, D., Kremer, P., & Robins, G.	2014	Small Group Research
6. Quintane, E., Pattison, P. E., Robins, G. L., & Mol, J. M.	2013	Social Networks
7. Ellwardt, L., Labianca, G. J., & Wittek, R.	2012	Social Networks
8. Lusher, D., Robins, G., Pattison, P. E., & Lomi, A.	2012	Social Networks
9. Sosa, M. E., Eppinger, S. D., & Rowles, C. M.	2004	Management Science
10 Hinds, P. J., Carley, K. M., Krackhardt, D., & Wholey, D.	2000	Organizational Behavior and Human Decision Processes
11. Van den Bulte, C., & Moenaert, R. K.	1998	Management Science
Both		
12. de Klepper, M. C., Labianca, G., Sleebos, E., & Agneessens, F.	2017	Journal of Management Studies
13. Kalish, Y., Luria, G., Toker, S., & Westman, M.	2015	Journal of Applied Psychology
14. Sosa, M. E., Gargiulo, M., & Rowles, C.	2015	Organization Science
15. Schulte, M., Cohen, N. A., & Klein, K. J.	2012	Organization Science
16. Emery, C.	2012	Social Networks

Note.

Table 3

Example Results from Influence Model Predicting Negative Affect (T1)

	b	***SE***	***β***	***CI$_{95\%}$***	***p***
Intercept	2.326	.252	-.005	[-.116, .105]	.924
Negative Affect (*T0*)	.245	.067	.206	[.095, .316]	< .001
Network Exposure (*T0*)	.005	.003	.119	[.008, .230]	.036
					R^2 = .053

Note. F = 9.357, p < .001. b is the unstandardized regression coefficient work. SE is the standard error. β is the standardized regression coefficients. Confidence intervals and p values are reported for standardized coefficients. Exposure = ego's exposure to the alters negative affect through dyadic information sharing. The ratio between exposure and autoregression (i.e., the Exposure Ratio Index) is .579.

Table 4

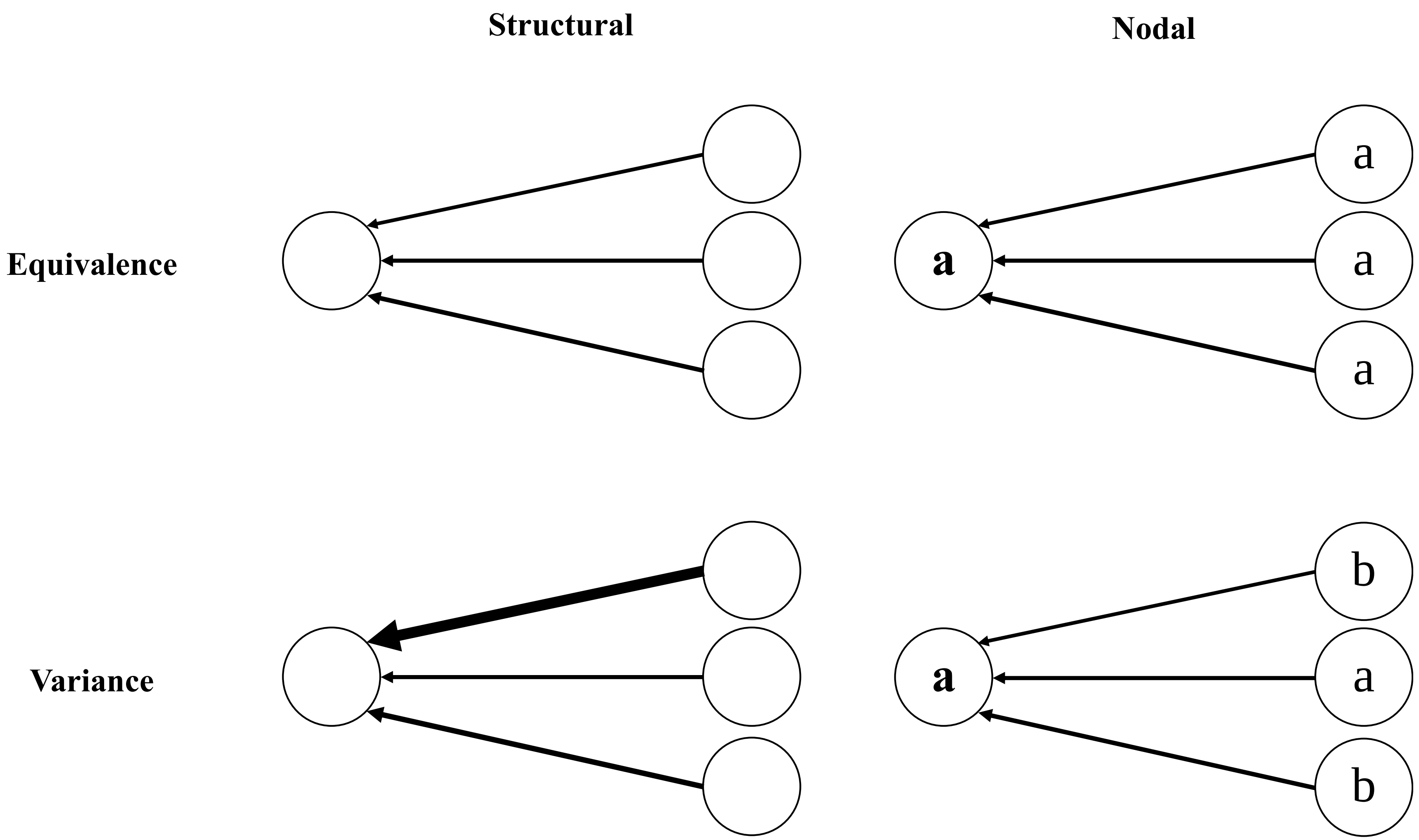
Example Results from Selection Model Predicting Information Sharing (T1)

	b	***SD***	***β***	***CI***$_{95\%}$	***p***
Constant	2.162	.814	.535	[.041, 1.105]	.057
Ego Effects					
Negative Affect (T0)	-.125	.147	-.149	[-.488, .151]	.371
Alter Effects					
Negative Affect (T0)	.045	.136	.047	[-.327, .366]	.787
Dyadic Effects					
Information Sharing (T0)	.416	.081	.386	[.234, .540]	< .001
Affective Similarity (T0)	.321	.084	.297	[.133, .474]	.001

Note. Pseudo R^2 = .162. SSFI = .249. *b* and SD represent the mean and standard deviation of the posterior parameter distribution. β represents the standardized mean of the posterior parameter destitution. Confidence intervals and *p* values are presented for the standardized model. Information sharing at *T0* and *T1* were standardized using within group variance due to potential group contextual factors determining the frequency of communication. Ego, alter, and dyadic similarity effects for negative affect were standardized using the full sample variance.

Figure 1

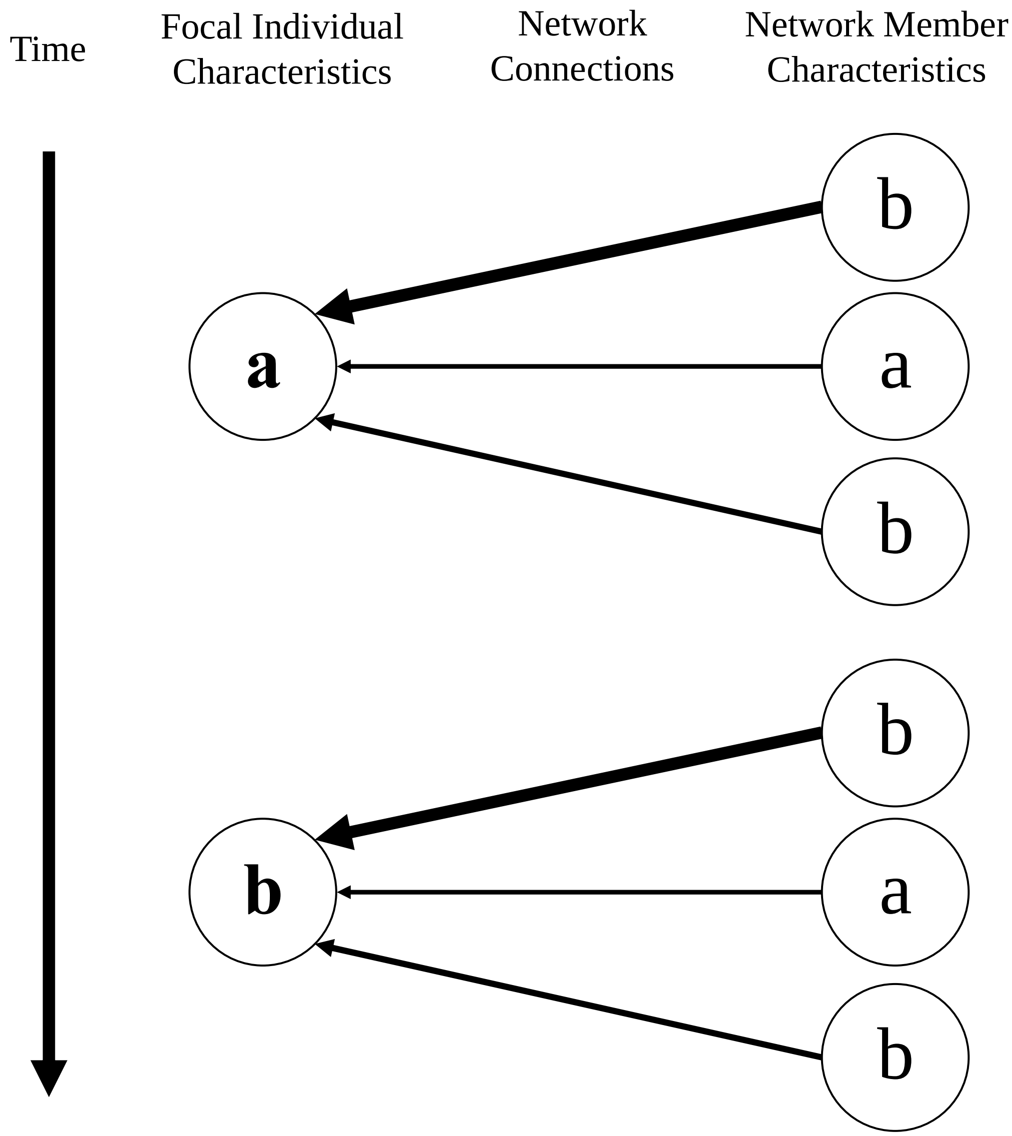
Graphical Representation of the Equivalence and Variance in Structure and Nodes



Note. Where a = one characteristic (e.g., low negative affect), b = an alternative characteristic (e.g., high negative affect), thicker lines = a stronger connection (e.g., high amounts of information receipt), and thinner lines = a weaker connection (e.g., low amounts of information receipt).

Figure 2

The Influence Process: Change in People

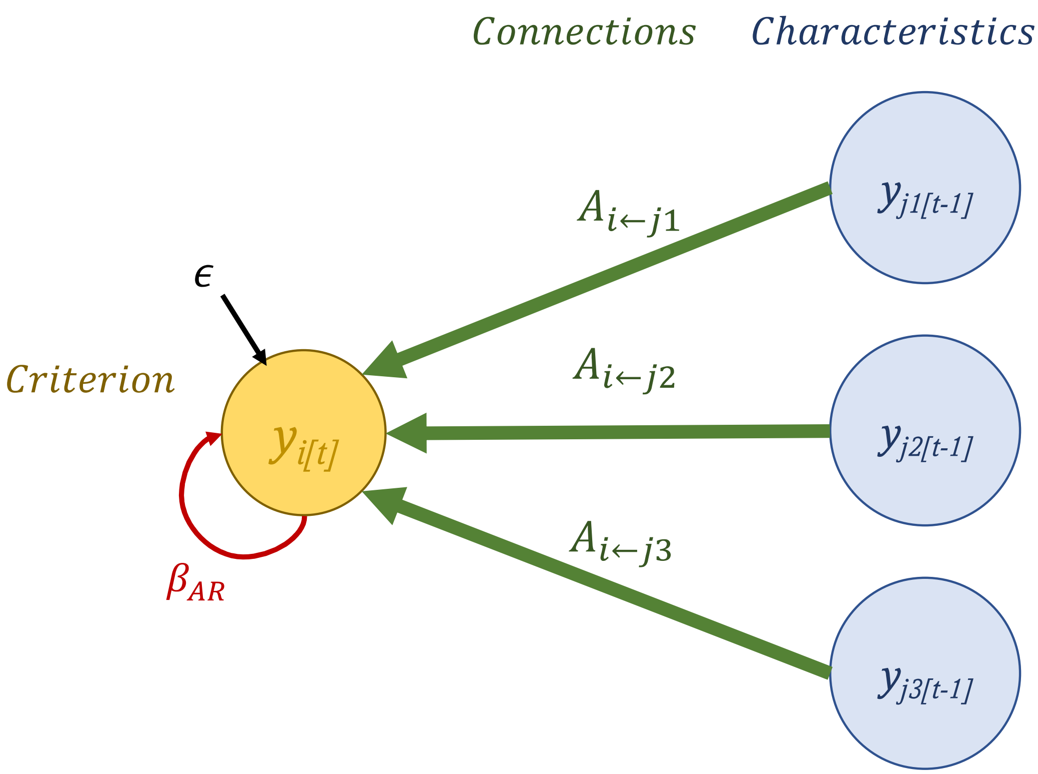


Note. Where a = one characteristic and b = an alternative characteristic. The thickness of the lines represents the strength of the tie. Thus, characteristics of the focal individual change over time as a function of their network connections and member characteristics.

Figure 3

The Influence Model: Exposure through the Network

$$y_{i[t]} = \beta_{AR} y_{i[t-1]} + \beta_{ex} \sum_{j \in N(i)} A_{i \leftarrow j} \; y_{j[t-1]} + \epsilon$$



Note.

Figure 4:

Overview of the Data Structure and Process of Evaluating an Influence Model

Steps 1 and 2

Node_Influence.CSV

One row per person with two characteristic measures (times *t-1* and *t*)

Individual ID	Team ID	Negative Affect T0	Negative Affect T1
1	1	3.29	4.21
2	1	4.58	3.41
3	1	1.19	2.74
4	1	2.38	2.23
5	1	1.46	1.20
6	1	4.84	4.53
7	1	3.15	3.24
8	1	2.57	1.37
9	1	4.33	4.31
10	1	1.69	1.48
11	2	1.24	2.88
12	2	3.59	3.61
13	2	1.06	3.57
14	2	2.51	4.45
15	2	2.57	4.42
16	2	1.13	3.72
17	2	1.27	3.94
18	2	4.26	3.04
19	2	3.41	2.27
20	2	1.41	4.20

+

Edge_Influence.CSV

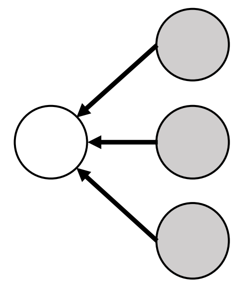
One row per dyad with one connection measure (time *t-1*)

Ego ID	Alter ID	Information Receipt T0
1	2	4.03
1	3	3.82
1	4	3.71
1	5	4.14
1	6	1.85
1	7	3.37
1	8	2.14
1	9	4.19
1	10	3.61
2	1	2.23
2	3	1.42
2	4	4.39
2	5	1.15
2	6	3.08
2	8	2.11
2	9	4.87
2	10	3.42
3	1	1.01
3	2	2.56
3	4	3.23

Step 3

Network Exposure

Calculate network exposure for each individual



Step 4

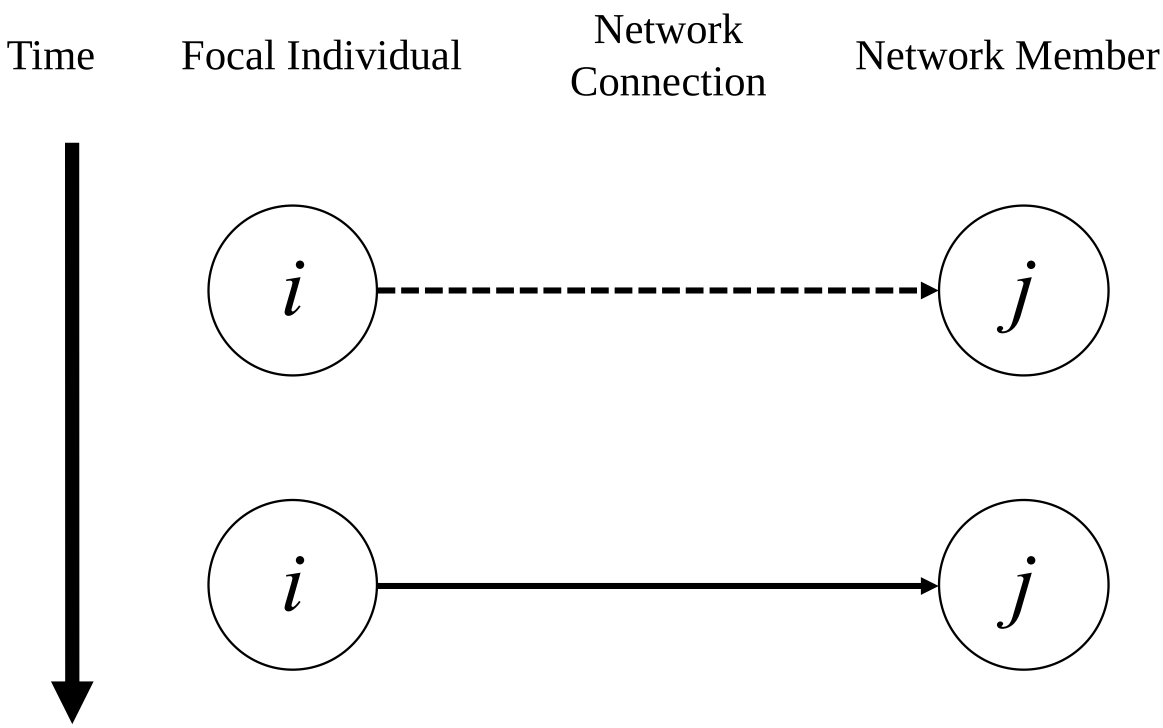
Statistical Analysis

Predict outcome with autoregressive term, exposure, and other predictors

$$y_{i[t]} = \beta_{AR} y_{i[t-1]} + \beta_{ex} \sum_{j \in N(i)} A_{i \leftarrow j} \; y_{j[t-1]} + \epsilon$$

Figure 5

The Selection Process: Change in the Network

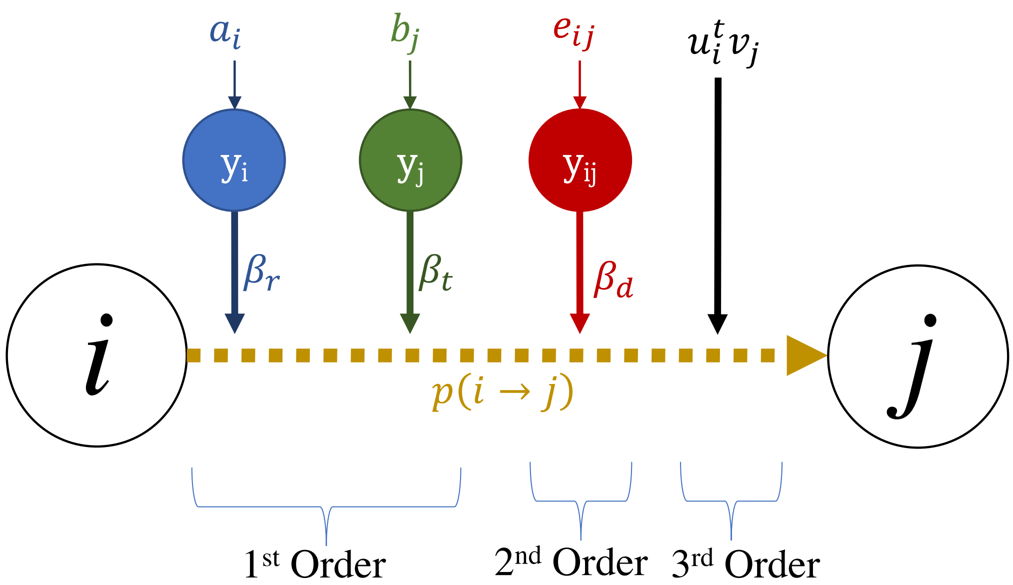


Note. Where the dashed line represents a potential tie and a solid line represents an established tie. Thus, a connection (i.e., tie) forms over time due to the characteristics of person *i* and person *j*.

Figure 6

The Selection Model: Social Relations and Latent Factors

$$p(i \rightarrow j) = \beta_r y_i + a_i + \beta_t y_j + b_j + \beta_d y_{ij} + e_{ij} + u_i' v_j$$



Note. Where $p(i \rightarrow j)$ represents the probability of forming a tie from person *i* to person *j* when the network is non-weighted, requiring a logistic transformation. In a weighted network analysis this will represent the strength of the given tie. y_i represents ego effects (i.e., how likely person *i* is to form ties in general), y_j alter effects (i.e., how likely people are to form ties with person *j*), and y_{ij} the dyadic effect (i.e., unique interaction between person *i* and person *j*). a_i, b_j, and e_{ij} are random intercepts for the rater, target and dyad respectively.
u_i and v_j represent latent factors or multiplicative effects for the ego and alter respectively (see Hoff, 2021). β_r, β_t, and β_d represent regression coefficients for the rater, target, and dyadic effects respectively.

Figure 7:

Overview of the Data Structure and Process of Evaluating a Selection Model

Steps 1 and 2

Node_Selection.CSV

One row per person with one characteristic measure (time *t-1*)

Individual ID	Team ID	Negative Affect T0
1	1	3.29
2	1	4.58
3	1	1.19
4	1	2.38
5	1	1.46
6	1	4.84
7	1	3.15
8	1	2.57
9	1	4.33
10	1	1.69
11	2	1.24
12	2	3.59
13	2	1.06
14	2	2.51
15	2	2.57
16	2	1.13
17	2	1.27
18	2	4.26

+

Edge_Selection.CSV

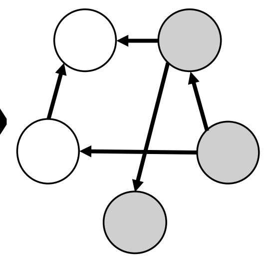
One row per dyad, with two connection measures (times *t-1* and *t*)

Ego ID	Alter ID	Information Sharing T0	Information Sharing T1
1	2	1.84	2.39
1	3	1.76	3.70
1	4	3.75	2.32
1	5	2.40	1.86
1	6	2.28	4.97
1	7	4.59	1.82
1	8	3.17	4.26
1	9	3.76	3.73
1	10	2.81	4.25
2	1	2.22	2.29
2	3	3.08	4.43
2	4	4.94	3.80
2	5	2.02	2.94
2	6	3.78	3.16
2	7	1.66	4.22
2	8	2.40	2.66
2	9	4.32	3.13
2	10	4.92	4.18

Step 3

Relational Covariates

Calculate relational covariates for each dyad or interaction terms



Step 4

Statistical Analysis

Selection Model

$$p(i \rightarrow j) = \beta_r y_i + a_i + \beta_t y_j + b_j + \beta_d x_{ij} + e_{ij} + u_i' v_j$$