

RUNNING HEAD: ANALYZING THE CONSEQUENCES OF PERSONALITY

The social consequences and mechanisms of personality: How to analyze longitudinal data  
from individual, dyadic, round-robin, and network designs

Steffen Nestler

University of Münster

Kevin Grimm

Arizona State University

Felix D. Schönbrodt

Ludwig-Maximilians-Universität München

**European Journal of Personality, in press**

*This is an unedited manuscript that has been accepted for publication. The manuscript will undergo copyediting, typesetting, and review of the resulting proofs before it is published in its final form*

Author Note

We thank Sarah Humberg, Marius Leckelt, Mitja D. Back, and four anonymous reviewers for very helpful suggestion concerning earlier drafts of this article.

Correspondence concerning this article should be addressed to either Steffen Nestler, University of Münster, Fliegerstr. 21, 48149 Münster, E-mail: steffen.nestler@wwu.de, or Felix D. Schönbrodt, Ludwig-Maximilians-Universität München, Department Psychologie, Leopoldstraße 13, 80802 München, E-Mail: Felix.Schoenbrodt@psy.lmu.de..

## Abstract

There is a growing interest among personality psychologists in the processes underlying the social consequences of personality. To adequately tackle this issue, complex designs and sophisticated mathematical models must be employed. In this article we describe established and novel statistical approaches to examine social consequences of personality for individual, dyadic, and group (round-robin and network) data. Our overview includes response surface analysis (RSA), autoregressive path models, and latent growth curve models for individual data; actor-partner interdependence models and dyadic RSAs for dyadic data; and social relations and social network analysis for round robin and network data. Altogether, our goal is to provide an overview of various analytical approaches, the situations in which each can be employed, and a first impression about how to interpret their results. Three demo data sets and scripts show how to implement the approaches in R.

*Keywords:* social relations model, social network analysis, dyadic data, response surface analysis, longitudinal data analysis, mediation

**The social consequences and mechanisms of personality: How to analyze longitudinal data from individual, dyadic, round-robin, and network designs**

Psychologists assume that personality has a number of interesting and wide-ranging consequences for the development and maintenance of social relationships. It is posited, for instance, that agreeable individuals become more popular the longer one knows them (see Selfhout, Burk, et al., 2010) or that narcissists are liked less the more one gets familiar with them (see Back, Schmukle, & Egloff, 2010; Kufner, Nestler, & Back, 2013). Similarly, the agreeableness of a member of a couple is believed to influence the relationship satisfaction of the other couple member (see Dyrenforth, Kashy, Donnellan, & Lucas, 2010). Importantly, however, current theorizing does not only focus on such personality-consequence effects, but also on the processes or mechanisms underlying them. Do agreeable individuals, for example, become more popular because others have more valuable communications with them or do narcissists get less popular due to their aggressive communication style?

To answer such sophisticated questions adequately, researchers do not only need longitudinal designs in which multiple measures of personality, process indicators, and indicators of the consequence variable are assessed (see Wrzus & Mehl, 2015), but they also have to employ complex mathematical models to statistically evaluate their research questions. Partially due to this need for more complex statistical models, there has been an explosion of research in the methodological field providing researchers with an ever-growing set of analytic tools. Our goal is to provide a comprehensive overview of the novel statistical approaches that can be used to examine social consequences of personality. We highlight several commonly utilized approaches, briefly discuss their mathematical bases, and describe their usefulness for studying personality and social consequences. Thus, readers will get an overview of possible research questions and analytical approaches (“What can be done?”), the situations in which they can employ an approach (“When can I use it?”), and a first impression about how each approach can be used (“What are the basic results and how can

they be interpreted?"). For all models, we point to relevant literature allowing the reader to get more acquainted with each approach. Furthermore, we provide three demo data sets and scripts that show how to implement all of the approaches in R (R Core Team, 2014). The data and the scripts can be downloaded from <http://osf.io/dny58>.

Many of the phenomena examined by personality psychologists refer to the *individual* (e.g., how agreeableness affects the popularity of a person). But there are also many phenomena that occur in a *relational* context, such as dyads (e.g., how the agreeableness of the member of a couple influences the other couple's members relationship satisfaction) or groups (e.g., how the agreeableness of single group members is related to how much they are liked in the group on average), and personality psychologists are increasingly interested in investigating personality effects in such contexts. This motivation stems from the fact that many consequence variables refer to two or more persons (e.g., relationship satisfaction, see Kenny, Kashy, & Cook, 2006). They are thus *interpersonal*, so that the effect of personality on the consequence variable may also depend on the characteristics of the other people that are embedded in the social context where the consequence variable is occurring. To disentangle these different influences, relational data in addition to individual data have to be assessed.

From a statistical point of view, the analysis of relational data is more complicated than the analysis of individual data as relational data are *dependent* (see Gonzales & Griffin, 1997; Kenny et al., 2006). For example, the relationship satisfaction scores from a married couple tend to be positively correlated because of their mutual influence (e.g., satisfaction of one member will affect the satisfaction of the other member). Similarly, personality scores of the members of a couple are likely to be correlated because of compositional effects, such as assortative mating. Given that most statistical procedures (e.g., regression) assume independent observations, specific methods have to be employed to analyze relational data.

To accommodate the interest of personality psychologists in relational data, the present article is therefore not only devoted to a description of established and novel statistical approaches to examine social consequences of personality at the *individual* level, but also to statistical approaches used to analyze data at the *relational* level. To enhance the clarity of our explanations, we will discuss two forms of relational data separately: dyadic data and group data. Specifically, this article is organized as follows: The first section is devoted to the analysis of individual data. We start with a description of an example we use throughout the section to illustrate the different statistical approaches. The paragraphs after show how to examine the effects of personality and personality similarity on social consequence variables, how to examine longitudinal processes and their determinants, and how to test whether a variable mediates the personality-consequence relation. In the second section, we concentrate on the analysis of dyadic data. After the description of an example, we show how the effect of personality on consequences can be examined in a dyadic context, how to analyze longitudinal dyadic data, and present mediation models for cross-sectional and longitudinal dyadic data. In the final section, we will explain statistical approaches that can be used to analyze group data. Specifically, we introduce how the social relations model and social network analysis tackle this kind of data. This is followed by explaining how the effect of personality on social consequences can be examined, how one can handle longitudinal group data, and how to test mediation models with group data.

### **Analyzing Individual Data**

In this section, we concentrate on statistical approaches to analyze individual data. We first describe an example together with illustrative research questions. This is followed by remarks concerning data management and preparation. Then, we describe how the social consequences of personality can be analyzed, how one can analyze longitudinal data, and how one can test mediation.

### **Typical Example and Illustrative Research Questions**

To make the analytical approaches concrete, we refer to a hypothetical study in which data on the popularity of 100 freshmen were collected during the first semester at three consecutive measurement occasions (e.g., the first week of the first semester, 2 months later and another 2 months later). On each occasion, each freshman was asked to indicate the number of other freshmen with which they are friends. Furthermore, on each occasion, the participants' agreeableness was measured by asking them to fill out an agreeableness scale. At the first occasion, a friend was also asked to answer this scale (i.e., a peer-report was assessed). Finally, in an attempt to examine mediators underlying the potential effect of agreeableness on popularity, all participants were asked to indicate the number of freshman they interacted with in the week before the measurement. Table 1 gives a summary of the example. The (simulated) dataset and the R codes to analyze the data with the described statistical approaches can be downloaded from <http://osf.io/dny58/>.

Several research questions can be examined with these data. A first question is whether agreeable freshmen are indeed more popular and second, whether self-other agreement in agreeableness predicts their popularity. Given the longitudinal design, one may also examine whether there are any individual changes in the freshmen's popularity over time, and if so, whether these changes can be explained by agreeableness. Finally, one can examine whether the effect of agreeableness on changes in popularity is mediated by the quantity of their interactions. Note that this is only a short list of potential research questions that can be evaluated with this data. Nevertheless, several analytical approaches are already needed to answer these exemplary questions.

Before we describe these approaches, we note that data collected at multiple time-points (as well as dyadic and group data) do not conform to the typical common two-dimensional data structure ('each row is one participant'). In contrast, these are multidimensional data structures that call for a more flexible way of storing. We strongly recommend storing data in the *long format* (Kenny et al., 2006) where there are multiple

records per person (one per occasion). An example of data in long format is given in Table 2. In this format each measurement occasion goes into one row, and multiple variables that have been assessed at this observation are placed into columns.

Additionally, we note two problems that often occur when designing studies and applying statistical models. These are (1) how many participants have to be assessed to have sufficient power to detect the effect of interest, and (2) how to treat missing values. A general treatment of both issues is beyond the scope of this article. Concerning power, however, we note that for some statistical approaches described here, necessary sample sizes can be computed (i.e., with G\*Power, for example; Faul, Erdfelder, Buchner, & Lang, 2009), whereas for the more sophisticated models (e.g., growth models), Monte Carlo simulations are typically conducted (see Muthén & Muthén, 2002, for an accessible introduction). With regard to missing values, we would like to mention that effective missing data algorithms exist (e.g., full information maximum likelihood, multiple imputation) that should be used in the analyses instead of case-wise deletion (e.g., omit a participant who has one missing value; see Schafer & Graham, 2002, Enders , 2010).

### **Examine the effects of personality or similarity in personality judgments on social consequences**

We now describe how the effects of personality or the effects of self-other agreement in personality judgments on social consequences can be investigated. Concerning the effect of personality on social consequences, first, the most straightforward way is to compute a single or multiple (depending on the number of predictors) linear regression. Here, the outcome variable is the social consequence variable (e.g., popularity of the freshmen) and the predictor is the personality trait (e.g., agreeableness). Given that most personality researchers have good knowledge of regression, we will not explain this method in detail here (see Tabachnick & Fidell, 2013, and Gelman & Hill, 2007, for introductions), but rather we move to discussing

how the effect of personality similarity on social consequences can be examined using Response Surface Analysis.

**Response Surface Analysis.** Many research questions contain the notion that the similarity (also called "fit", as in person-environment-fit) between two variables predicts an outcome. As an example, a fit hypothesis could be that popularity (e.g., in terms of indicated number of friends) is lower when the self-concept of agreeableness deviates from the way one is seen by his or her peers (i.e., low self-other agreement). Such research questions typically have been investigated with moderated regression or with absolute or squared difference scores which quantify the amount of (mis)fit of the two predictor variables. However, both analytical approaches are problematic. Fit hypotheses are inherently nonlinear hypotheses that usually predict an impaired outcome for both types of misfit (see also Figure 1). Moderated regression belongs to the family of linear models, and therefore is, in principal, unable to adequately test a nonlinear hypothesis. Difference scores, in turn, have been heavily criticized for several reasons (Blanton, Jaccard, Gonzales, & Christie, 2006; Edwards, 2001; Johns, 1981; Peter, Churchill, & Brown, 1993). Most importantly, they impose untested constraints on the data, for example, that (in)congruence on higher level ( $9-8 = 1$ ) is equivalent to (in)congruence on a lower level ( $2-1 = 1$ ). As an alternative to moderated regression and difference scores, response surface analysis (RSA) has been proposed (see Edwards & Parry, 1993; Edwards, 2002; Schönbrodt, 2014a, 2014b; Shanock, Baran, Gentry, Pattison, & Heggestad, 2010, for introductions). With this approach, fit hypotheses are modeled using a second-degree polynomial, such that<sup>1</sup>

$$z_n = b_0 + b_1x_n + b_2y_n + b_3x_n^2 + b_4x_ny_n + b_5y_n^2 + e_n, \quad (1)$$

where  $z_n$  is the outcome variable of person  $n$  (in our example: popularity),  $x_n$  (self-rating of agreeableness) and  $y_n$  (peer-rating of agreeableness) are the predictors,  $x_ny_n$  is their product, and  $x_n^2$  and  $y_n^2$  are their squared values.

Polynomial regression is a special case of the general multiple regression framework, and can be seen as an extension of the commonly used moderated regression. Due to the squared predictors, this regression model is able to accommodate a large range of linear and nonlinear functional forms. After computing the five regression coefficients, they can be used to plot a *response surface* in the three-dimensional space which in turn can be employed to analyze the validity of the fit hypothesis (hence the name response surface analysis). Typical response surface plots are shown in Figure 1. All congruent predictor values lie on the diagonal of the predictor plane (the "floor" of the cube). This diagonal is called the *line of congruence* (LOC) and goes from (-2, -2) to (+2, +2) in the plots. RSA provides statistical descriptions and hypothesis tests of certain features of that response surface. For example, one can test whether the ridge of the surface is significantly shifted away from the LOC (as in Figure 1b), or whether the LOC has a significant upwards slope towards the high/high predictor combination (as in Figure 1c). The latter example refers to a situation where congruence on a high level is not equivalent to congruence on a low level. Most importantly, one can directly test whether the surface bends along the *line of incongruence* (LOIC; see Figure 1c), which runs orthogonal to the LOC. This is a direct test of the fit hypothesis, which predicts that the outcome gets worse for combinations of predictor variables that approach the corners of misfit (i.e., (-2, +2) and (+2, -2)).

To summarize, RSA makes the major implicit assumptions of discrepancy scores testable and covers a much wider range of theoretically relevant functional forms than simple discrepancy scores or moderated regressions. Visualizing the results as a response surface plot makes the results of the polynomial regression, which are often hard to understand numerically, easily accessible. Statistical tests of surface features, such as the curvature or slope of the LOC, give guidance which features of the surface to interpret, and directly test specific hypotheses. As RSA is a special case of general multiple regression framework, the usual effect sizes, such as  $R^2$  or standardized regression weights, can be reported.

In practice, RSA can be computed using standard statistical software (e.g., SPSS, SAS, Excel; for an introduction see Shanock et al., 2010). Alternatively, a comprehensive solution for computing and plotting response surfaces can be found in the *RSA* package for *R* (Schönbrodt, 2015), which includes several convenient plotting options, bootstrapped parameter estimates, and model comparisons.

This section showed how the effects of personality on social consequences and the effect of personality similarity on social consequences can be examined *cross-sectionally* (although the social consequence variable may be assessed later in time than the personality variable). If a researcher has collected data on consecutive measurement occasions, longitudinal models have to be employed.

### **Examining longitudinal changes in social consequences and how personality is related to these changes**

In the example data that we have discussed, personality (e.g., agreeableness), the consequence variable (e.g., popularity), and the mediators were assessed longitudinally (over time). Such longitudinal data allow us to test hypotheses more strongly because we can assess how these variables change over time, the determinants of such changes, and whether these variables interact and impact one another over time. In terms of our example, we can thus examine whether there are longitudinal changes in agreeableness or the number of friends, whether agreeableness, measured at the first occasion, predicts subsequent changes in the number of friends, or finally whether agreeableness leads to subsequent changes in friendship number and/or vice versa whether the number of friends leads to subsequent changes in agreeableness.

Given the richness of longitudinal data and the plethora of research questions that can be answered with longitudinal data, there exist, in turn, numerous statistical models. These models can be used to study within-person individual changes and between-person differences in those changes (e.g., growth models), within-person fluctuations, time-sequential dynamic

associations (e.g., autoregressive and cross-lag model, latent difference score models), and within-person acceleration and its associations with velocity and state (e.g., differential structural equation models) to name a few. Here, we focus on two of the most common statistical approaches for longitudinal individual data and, at the same time, recognize the variety of statistical approaches to analyzing longitudinal personality data.

In Table 3, we briefly mention the vast array of statistical models that have been used to study changes over time. In this table, the models are described in terms of the following criteria: whether rank-order changes or absolute changes are studied, whether change is differentially modeled between measurement occasions (time-sequential change, stochastic) or modeled with an overarching change trajectory (deterministic), whether nonlinear forms of change are allowed, whether between-person differences in change are modeled, whether models of changes in latent variables are possible, and the unit of analysis. Additionally, the table includes typical research questions, information regarding sample size and measurement occasion requirements, as well as a recommended reading on the topic and, when available, an R package to estimate the model. We note that Jackson and Allemand (2014) recently reviewed a series of longitudinal models (latent growth model, latent change model, and growth mixture model) for studying within-person change and between-person differences in change in the context of personality development. In our review, we add to this discussion and describe the autoregressive cross-lag and growth models.

**Autoregressive Cross-Lag Models.** Autoregressive cross-lag models (ACLM) can be used to examine *rank-order changes* and time-lagged associations with multivariate longitudinal data. For example, we could examine whether agreeableness affects subsequent changes in popularity. The bivariate ACLM can be written as

$$x_{nt} = g_{0t} + g_{1t} \cdot x_{nt-1} + g_{2t} \cdot y_{nt-1} + u_{nt} \quad (2)$$

$$y_{nt} = b_{0t} + b_{1t} \cdot y_{nt-1} + b_{2t} \cdot x_{nt-1} + e_{nt}$$

where  $x_{nt}$  and  $y_{nt}$  are the two repeatedly measured variables for individual  $n$  at time  $t$  (e.g., agreeableness and popularity at the second measurement),  $g_{0t}$  and  $b_{0t}$  are time-dependent intercepts,  $g_{1t}$  and  $b_{1t}$  are *autoregressive* parameters describing stability of between-person differences over time within each variable,  $g_{2t}$  and  $b_{2t}$  are *cross-lag* parameters describing the association between prior scores on one variable (measured at time  $t-1$ ) and subsequent scores (measured at time  $t$ ) in the second variable,  $x_{nt-1}$  and  $y_{nt-1}$  are the repeatedly measured variables at time  $t-1$  (e.g., agreeableness and popularity at the first measurement), and  $e_{nt}$  and  $u_{nt}$  are time-dependent residual scores. We note that in the ACLM, the time-lag is typically held constant, which enables tests for the equality of autoregressive and cross-lag parameters across time (i.e., drop  $t$  subscript to these parameters).

The ACLM is depicted as a path diagram in Figure 2 and the coefficient  $b_{23}$  denotes the effect of agreeableness measured at the second occasion on popularity measured at the third occasion. However, the autoregressive and cross-lag effects are *simultaneously* estimated, which allows for the cross-lag effects to be interpreted as effects on the amount of change between two consecutive occasions. Returning to our example of agreeableness and popularity over time, a positive cross-lag parameter from agreeableness to popularity would indicate that students who are more agreeable at time 1 showed more positive changes in popularity from time 1 to time 2.

The ACLM incorporates time implicitly, such that parameter estimates are dependent on the time-lag or the length of time between assessments, rather than explicitly (such as in the growth model). Oud and colleagues (Oud & Jansen, 2000; Oud & Delsing, 2010; Voelkle, Oud, Davidov, & Schmidt, 2012) have discussed continuous time approaches to analyzing panel data, where associations between continuous changes in the variables over time and their states are modeled. Thus, the parameter estimates are not dependent on the time-lag between assessments.

**Growth Models.** Growth models (Laird & Ware, 1982; McArdle & Epstein, 1987; Meredith & Tisak, 1984, 1990) are a class of longitudinal models used to study *within-person changes* over time and *between-person differences* in those within-person changes. In our example, we could use growth models to examine whether freshmen's popularity changes over time, whether freshmen differ in their rate of change and if so, whether those differences can be explained by their agreeableness.

The linear growth model can be specified as

$$y_{nt} = b_{1n} + b_{2n} \cdot time_{nt} + e_{nt}, \quad (3)$$

where  $y_{nt}$  is the repeatedly measured variable  $y$  for person  $n$  at time  $t$  (e.g., the popularity of freshman 1 at the three occasions),  $time_{nt}$  is the timing metric organizing observations,  $b_{1n}$  is the intercept for person  $n$  representing the true level of the variable when  $time_{nt} = 0$  (e.g., the popularity of the freshman at the first measurement), and  $b_{2n}$  is the linear slope for person  $n$  representing the true rate of change. Finally,  $e_{nt}$  is the residual term for person  $n$  at time  $t$ . It represents individual variability over time not captured by the linear growth model.

Equation 3 is the within-person equation representing linear change for a particular person (e.g., the linear change in popularity for freshman  $n$ ). One way to conceptualize growth models is to imagine that they provide an intercept and a slope for each subject, resulting in a distribution of  $N$  intercepts and  $N$  slopes (e.g., a distribution of intercepts and slopes for the  $N$  freshmen). Hence, in contrast to ACLMs, growth models track the absolute change and not the ranking of persons over time. The mean of the intercept,  $b_{1n}$ , and the slope,  $b_{2n}$ , are called the fixed-effects parameters of the model and are used to describe the mean trajectory over time. The variability of the intercepts and the slopes across individuals are the random-effects parameters and describe the magnitude of *between-person differences* in the intercept and slope. That is, the variance of the intercept represents the magnitude of true variability in popularity when  $time_{nt} = 0$ , and the variance of the linear slope gives the between-person differences in the rate of popularity change. It should be noted that we treated

the initial assessment as  $time_{nt} = 0$ . Alternative definitions of the time variable are possible, of course, such as defining the last measurement occasion as  $time_{nt} = 0$  (see Singer & Willett, 2003, for a detailed discussion).

Growth models can be extended to include time-varying and time-invariant predictors. For example, the intercept and slope can be outcomes in regression-like equations, such that

$$b_{1n} = \beta_{11} + \beta_{12} \cdot X_n + d_{1n} \quad (4)$$

$$b_{2n} = \beta_{21} + \beta_{22} \cdot X_n + d_{2n}.$$

The parameters  $\beta_{12}$  and  $\beta_{22}$  are regression parameters describing the association between the time-invariant covariate,  $X_n$ , and the intercept and slope. If we are examining changes in popularity and agreeableness is included as a time-invariant predictor,  $\beta_{12}$  would denote the effect of agreeableness on popularity at the first occasion and  $\beta_{22}$  is the effect of agreeableness on popularity change.

Growth models can be constructed within the *structural equation modeling (SEM)*; see Little, 2013; Nestler, in press; 2014a) and *multilevel modeling (MLM)*; see Singer & Willett, 2003; Verbeke & Molenbergh, 2009) frameworks. There are notable advantages of each framework when studying within-person change (Stoel & Galindo-Garre, 2011). The first advantage of the MLM framework is the flexibility of easily modeling within-person change when the timing metric varies at the individual-level as opposed to more or less fixed time-points that all individuals share. We note that certain SEM programs (e.g., *Mplus*, *OpenMx*) now allow for individually-varying time metrics. The second advantage is the ability to directly fit nonlinear models in addition to linear models. Nonlinear models have one or more random coefficients (or latent variables) that appear within an exponent or are raised to an exponent. These random coefficients are said to enter the model in a nonlinear fashion (as opposed to a linear/additive fashion). SEM is a linear modeling framework and thus cannot fit these models directly; however, see Browne and du Toit (1991) and Grimm, Ram, and Hamagami (2011) for approximate approaches.

The first advantage of the SEM framework is the evaluation of model fit. That is, the data structure for a growth model fit within the SEM framework (i.e., separate variable for each repeated measure) allows for the calculation of a null model as well as a saturated model, which are necessary to calculate absolute fit indices. Absolute fit indices enable the examination of the degree to which the model is able to reproduce the observed data. In addition, relative fit indices can be computed in the SEM framework. They can be employed to compare alternative growth models (e.g., with regard to the latent structure of the indicators over time; see Little, 2013) which is another advantage of the SEM framework. Second and again due to the structure of the data, the SEM framework allows for more exploratory forms of change, where the function of time is partially determined by the data (see Grimm, Steele, Ram, & Nesselrode, 2013). Third, the SEM framework also allows for the inclusion of a measurement model and thus, the examination of change in a multiple indicated latent variable (see McArdle, 1988). Fourth, the SEM framework enables the inclusion of covariates and predictors that are also multiple indicated latent variables. Fifth, the SEM framework provides greater flexibility for modeling residual structures (i.e. the structure of variances and covariances of  $u_{tn}$ ; see Grimm & Widaman, 2010). Lastly, SEM programs typically report standardized parameters in addition to unstandardized parameters, which aid interpretability of certain model parameters.

In this section, we discussed in detail the growth model and the ACLM. The latter model as well as the growth model in the framework of SEMs can be estimated using the R package *lavaan* (Rosseel, 2012). The R package *lme4* can be used to fit MLMs (Bates, Maechler, Bolker, & Walker, 2014). Furthermore, as we noted, power for these models are best studied using Monte Carlo simulations. Additionally, effect size for these models must be conceptualized in multiple ways. For the ACLM, autoregressive paths are typically strong ( $\beta > 0.50$ ) for constructs where between-person differences are relatively stable. Cross-lag parameters are typically much smaller, but can be meaningful with standardized beta weights

equal to 0.2, which is typically seen as a small to moderate effect. For growth models, we can try to gauge the magnitude of overall changes, which can be based on the magnitude of between-person differences at study conception (i.e., the intercept) or the magnitude of between-person differences in change. However, magnitude of effect sizes depends on the context of the study (a small change in academic achievement may be considered a large change in behavior given the different measurement of these constructs). More importantly, researchers are often interested in examining effects on the between-person differences in change. Here, the standards for effect sizes by Cohen (1988) have been meaningfully applied to growth models (see Snijders & Bosker, 2012).

Beyond these two models, however, many other models exist (see Table 3). Before closing, we consider it important to mention another extension of longitudinal models that may be of interest. If between-person differences in within-person changes are of interest in multiple constructs, *the parallel process model* (McArdle, 1988) may be applied. Here, the associations between the between-person differences in the rates of change (i.e., the slopes) are considered. Additional alternatives for studying multivariable development include the latent change score model (Jackson & Allemand, 2014; McArdle, 2001), the latent differential model, and the time-varying covariate model (see Ferrer & McArdle, 2003; Grimm, 2007, for comparisons of different approaches).

### **Examining the processes that mediate the personality and social consequences relation**

Mediation models describe the way in which a variable (e.g., agreeableness) affects another variable (e.g., popularity) through its influence on an intermediate or mediating variable (e.g., quantity of interactions). Typically, mediation models are tested with cross-sectional data (see MacKinnon, 2008). However, most mediation hypotheses presume that the assumed causal relationships among the variables take time to unfold (see Selig & Preacher, 2009) and cross-sectional mediation approaches yield biased estimates of such longitudinal mediation (Maxwell & Cole, 2007). Thus, longitudinal data are preferred in mediation

analyses. However, we note that this does not free a researcher to specify a theoretical valid mediation model (Morgan & Winship, 2015). In this vein, we discuss some exemplary longitudinal mediation models (see MacKinnon, 2008, or Selig & Preacher, 2009, for a detailed treatment of longitudinal mediation models) in this section. These models can, for example, be used to study whether the effect of agreeableness on popularity is mediated by the quantity of interactions.

Prior to discussing mediation-related extensions of the ACLM and growth models, we briefly describe a longitudinal mediation where the input (X), mediator (M), and outcome (Y) are each measured once, but in sequence (i.e., X is measured prior to M, which is measured prior to Y). This series of equations can be written as

$$m_n = b_0 + b_1 \cdot x_n + e_n \quad (5)$$

$$y_n = g_0 + g_1 \cdot m_n + g_2 \cdot x_n + u_n$$

In the first equation the input variable,  $x_n$ , is a predictor of the mediator,  $m_n$ , and in the second equation, the input variable and mediator are predictors of the outcome variable,  $y_n$ . With this series of equations, there is an indirect association, namely  $x_n \rightarrow m_n \rightarrow y_n$  with an effect equal to  $b_1 \cdot g_1$ . If  $m_n$  mediates the association between  $x_n$  and  $y_n$ , then the indirect effect will be significant. Typically, the standard error of the indirect effect is estimated using bootstrapping methods (see Preacher & Hayes, 2008; MacKinnon, 2008). We note that it is also possible that personality similarity (e.g., amount of fit in self-reported agreeableness and peer perceived agreeableness) does not directly predict the outcome, but rather influences a mediating variable (e.g., interaction quantity) which in turn affects the outcome (e.g., number of friends). Such a model can be tested by using the set of five variables that are specified in a RSA to predict the mediator and the outcome, respectively (for an application, see for example Edwards & Cable, 2009).

Longitudinal mediation models can be distinguished by the number of variables assessed longitudinally and distinguished in terms of whether they focus on between-person

differences or whether they include within-person variation as a part of the mediation model. Hence, again a number of different longitudinal mediation models can be examined. We discuss three extensions of the models discussed. The first is an extension of the ACLM. Here, between-person changes in the variables are the focus of the analysis. This model, for instance, can be considered when personality, social consequences, and the mediator are all assessed longitudinally (e.g., agreeableness, popularity and quantity of interactions were measures at three time-points). The other two models are extensions of the latent growth model. The first permits the inclusion of within-person variation and can be considered, for example, when personality and the consequence variable are measured at a single time-point (e.g., agreeableness and popularity) and the mediator is assessed longitudinally (e.g., number of interactions). The second models within-person changes in all variables and assesses mediation among intercepts and slopes.

**Extensions of the ACLM.** With three repeatedly measured constructs, the bivariate ACLM is extended to include a mediator (see Figure 3). Thus, the model can be written as

$$\begin{aligned}x_{nt} &= g_{0t} + g_{1t} \cdot x_{nt-1} + g_{2t} \cdot y_{nt-1} + g_{3t} \cdot m_{nt-1} + u_{nt} \\m_{nt} &= d_{0t} + d_{1t} \cdot m_{nt-1} + d_{2t} \cdot x_{nt-1} + d_{3t} \cdot y_{nt-1} + s_{nt} \\y_{nt} &= b_{0t} + b_{1t} \cdot y_{nt-1} + b_{2t} \cdot x_{nt-1} + b_{3t} \cdot m_{nt-1} + e_{nt}\end{aligned}\tag{6}$$

where  $x_{nt}$ ,  $m_{nt}$ , and  $y_{nt}$  are the repeatedly measured variables for person  $n$  at time  $t$  (e.g., agreeableness, quantity of interactions, and popularity). Each variable has an autoregressive effect ( $g_{1t}$ ,  $d_{1t}$ , and  $b_{1t}$ ) and cross-lag effects from the other repeatedly measured variables (e.g.,  $b_{2t}$  and  $b_{3t}$  for  $y_{nt}$ ). These cross-lag effects can be used to test for longitudinal mediation. Specifically, we can test whether the effect of  $x_{nt-2}$  to  $y_{nt}$  (e.g.,  $x_1$  to  $y_3$  in Figure 3) may be mediated by  $m_{nt-1}$  (e.g.,  $m_2$  in Figure 3). In terms of our example, we can thus test whether the effect of agreeableness, measured at the first occasion, on popularity assessed at the third occasion is mediated by the quantity of interactions assessed at the second time-point. The indirect effect is given by the product of the respective cross-lag parameters, that is

$d_{2t-1} \cdot b_{3t}$  (in Figure 3:  $d_{22} \cdot b_{33}$ ; see MacKinnon, 2008; Selig & Preacher, 2009). Similarly, one can also test whether  $m_{nt-1}$  mediates the effect from  $y_{nt-2}$  to  $x_{nt}$ . The relevant indirect effect is  $d_{3t-1} \cdot g_{3t}$  (in Figure 3:  $d_{32} \cdot g_{33}$ ; see Blandon, Calkins, Grimm, Keane, & O'Brien, 2010; Ferrer & McArdle, 2003, for applications and discussions)

**Extensions of the Growth Model.** The mediator extension of the ACLM focuses on between-person changes in the variables; however, there might be circumstances where a researcher wants to test whether within-person variation or change mediates the input-outcome relation (see Figure 4). In terms of our example, a researcher may want to examine whether the effect of agreeableness to popularity is mediated by changes in the quantity of interactions (e.g., one has more interactions with agreeable individuals with time and this is related to their popularity). In the situation where the mediator is measured over time, but the outcome and the originating variable are measured at a single occasion, one can extend the linear growth model discussed above, and can specify the *longitudinal growth mediation model* as

$$m_{nt} = b_{1n} + b_{2n} \cdot time_n + e_{nt} \quad (7)$$

$$b_{1n} = \beta_{11} + \beta_{12} \cdot X_n + d_{1n}$$

$$b_{2n} = \beta_{21} + \beta_{22} \cdot X_n + d_{2n}$$

and

$$Y_n = g_0 + g_2 \cdot b_{1n} + g_3 \cdot b_{2n} + g_4 \cdot X_n + u_n$$

where a linear growth model is specified for  $m_{nt}$  (e.g., quantity of interactions) with  $X_n$  (e.g., agreeableness measured at the first occasion) as a predictor of the intercept ( $b_{1n}$ ) and slope ( $b_{2n}$ ). Additionally, there is a regression model with the consequence variable as the outcome and the intercept, slope, and  $X_n$  as the predictors. In such a model, two indirect effects can be computed, one through the intercept,  $g_2 \cdot \beta_{12}$ , and one through the slope,  $g_3 \cdot \beta_{22}$ . For our example, the effect from agreeableness to popularity may be mediated by intercept for interaction quantity or its rate of change.

If there is a situation where the input, mediator, and output are all assessed longitudinally and there is an interest in within-person changes, growth models can be specified for the input, mediator, and outcome. A mediation analyses can then be performed among the intercepts and slopes. For example, the slope of the input could be used as the predictor of the slope of the mediator, and the slope of the input and the slope of the mediator could be used as predictors of the slope of the outcome. One limitation of this approach is that the time sequence of the input, mediator, and the outcome is blurred if all variables are assessed at the same time-points across time. This blurring is due to slopes encompassing all time- points and not being time-specific.

Before closing this section, we note that there are number of different methods to express the effect size for a mediation effect (see MacKinnon, 2008, for an overview). One can, for example, compute the proportion of the total effect that is mediated, one can standardize the indirect effect, or one can compute an index of explained variance. However, Preacher and Kelley (2011) show that all these methods are problematic, and they suggest two new methods to obtain the effect size for mediation effects. The formulas for both methods are implemented in the R package MBESS (Kelly & Lai, 2012); a practical problem with these measures, however, is that they are currently limited to a mediation model that only contains one mediator.

### **Analyzing Dyadic Data**

Having described models to analyze individual data, we now turn to the analysis of dyadic data. As before, we first describe an example together with some illustrative research questions. We then show how the social consequences of personality can be examined in a dyadic context. Thereafter, we present models to analyze longitudinal dyadic data and approaches to examine mediation in dyadic contexts.

### **Typical Example and Illustrative Research Questions**

In our fictitious study, data on the relationship satisfaction of 100 heterosexual couples were collected during three consecutive measurement occasions. Couple members also reported their agreeableness on each occasion. Finally, in an attempt to examine mediators underlying the potential effect of agreeableness on relationship satisfaction, each participant provided an estimate of the amount of stress experienced from the first to the second and from the second to the third measurement occasion. Table 4 gives a summary of the example. The simulated dataset and the R codes can be downloaded from <http://osf.io/dny58/>.

One question that can be investigated with this data set pertains to whether a partner's agreeableness predicts the other partner's relationship satisfaction. Also, we can evaluate whether the similarity of the couple partners' personality predicts their relationship satisfaction. Furthermore, given the longitudinal design of the study, one may investigate whether there are any longitudinal changes in the couples' satisfaction, whether these changes occur for both couple members and whether they can be explained by agreeableness. Finally, one can examine whether the effect of agreeableness on changes in relationship satisfaction is mediated by the amount of stress experienced by each of the couple members. Before we introduce the approaches that can be used to tackle the suggested questions, we explain an important feature of dyads, whether they are distinguishable, and we make some general remarks concerning data preparation.

### **Distinguishability of Dyads and Data Preparation**

An important feature of dyads is their *distinguishability*. Dyad members are distinguishable when they can easily be identified from one another on the basis of a theoretically meaningful variable. Heterosexual couples are distinguishable, for instance, as they can be separated by gender. Other examples of distinguishable dyads include parent-child dyads or boss-employee dyads. Dyad members are indistinguishable when no theoretically meaningful variable exists that can be used to order the dyad members. Examples are homosexual dyads, twins, or dyads of coworkers. Distinguishability is critical

for the selection of the statistical approach to analyze the dyadic data as a technique might be appropriate for distinguishable dyads but *not* for indistinguishable ones. For example, Pearson correlations can be used to assess the dependence between the measures of distinguishable dyads (e.g., a correlation between the wife's relationship satisfaction and the husband's relationship satisfaction). For indistinguishable dyads, however, a Pearson correlation is not appropriate because one can arbitrarily decide which dyad member's score is treated as the  $x$  and which as the  $y$  variable. In this case, it is appropriate to use the intraclass correlation as the magnitude of that coefficient does not depend upon the arrangement of the data (Kenny et al., 2006).

With respect to data preparation, the same considerations regarding missing values and statistical power for individual data analysis apply to dyadic data analysis. Furthermore, dyadic data can be analyzed with MLMs or SEMs. In the MLM framework, the data is structured in the *long* format (see Table 5), that is, the members of the dyads are stacked above each other. When the dyads are distinguishable, a variable indicates the group to which each individual belongs (e.g., 'wife' or 'husband'). In case of SEMs, by contrast, the dyadic data is typically structured in the *wide* format. Here, each row is one *dyad* and each measured variable has two columns, one for each member of the dyad. The wide data format makes clear that the unit of analysis in case of SEMs is the dyad and not the individual. Hence, degrees of freedom usually refer to the number of dyads.

For easier interpretation of the model coefficients, often centering of the predictor variables is advisable. In the case of individual data this is done by subtracting the grand mean of a variable from the individual scores in this variable. For dyadic data, one has to compute the grand mean of the variable of the dyad members *combined*. In our example, one should compute the grand mean across the agreeableness scores of the wives and the husbands, and should use the resulting mean for centering (see Kenny et al., 2006).

**Examining the Effects of Personality and Personality Similarity on Social Consequences**

We now turn to the statistical approaches to study the effects of personality or personality similarity on social consequences in a dyadic context. The general strategy, to put it metaphorically, is to compute the individual level model twice, one for each member of the dyad (cf. Kenny et al., 2006). What makes the analysis complicated is that one has to control for the non-independence of the scores within a dyad and to add partner effects that quantify the cross-partner influences. In the following, we first briefly describe the *Actor-Partner-Interdependence Model* (APIM, see Campbell & Kashy, 2002; Kenny et al., 2006, for introductions). Thereafter, we present a dyadic RSA to examine the effects of personality similarity in a dyadic context.

**APIM.** Imagine that you want to examine whether the husband's agreeableness predicts his relationship satisfaction, and whether the wife's agreeableness predicts the husband's relationship satisfaction. Figure 5 gives a graphical illustration of the APIM, a frequently used framework to examine such questions. The model contains two dyad members (e.g., the husband, *h*, and the wife, *w*) and two variables (e.g., agreeableness, *x*, and relationship satisfaction, *y*). Also, paths for actor effects ('*a*' in Figure 5) and paths for partner effects ('*p*' in Figure 5) are specified. Actor effects describe the influence of a dyad member's predictor (e.g., agreeableness) on his *own* social consequences (relationship satisfaction); they thus indicate *intrapersonal* effects. The partner effects, by contrast, are *interpersonal*, as they denote the influence that the dyad member's predictor has on the *other* partner's outcome. In the example, the partner effect describes the effect of the husband's (or the wife's) agreeableness on the wife's (or the husband's) relationship satisfaction. We note that the residual scores in Figure 5 are correlated. This correlation reflects the dependence in the data that is not explained by the modeled predictor variables.

The goal of the statistical analyses is to estimate the actor and partner effects. In case of distinguishable dyads, SEMs can be used for this purpose (Cook & Kenny, 2005; Kenny et al., 2006) as the standard APIM as depicted in Figure 5 is mathematically equivalent to an

autoregressive cross-lag model involving just two time-points. The stability coefficients are the actor effects and the cross-lagged effects are the partner effects. Three things are worth mentioning concerning the estimation of the APIM with SEMs. First, SEMs allow defining constraints on the model parameters. This can be used to check so-called APIM patterns, such as the actor-only pattern (i.e., nonzero actor effects and zero partner effects). As each pattern implies very different processes occurring in the dyadic relationship, incorporating them in the model is an interesting approach to test a dyadic relationship theory (see Kenny & Ledermann, 2010, for a more detailed treatment). Second, when using SEMs, the unit of the analysis is the dyad; the sample size for the analysis (and for power considerations) is thus the number of dyads. Finally, and this is more a cautionary note, most SEM programs standardize the coefficients separately for each dyad member, rendering the comparison of the coefficients across dyad members impossible. Hence, we suggest to standardize the variables manually before submitting them to the SEM analysis, as described above, or to use the unstandardized coefficients from the SEM output (see Kenny et al., 2006).

Estimation of the APIM in the case of indistinguishable dyads is typically carried out using MLM (but see Olson & Kenny, 2006). Compared to SEMs, the MLMs estimate the parameters of the actor and partner effects with a single equation. One thus obtains only *one* actor effect and only *one* partner effect as a result of the indistinguishability of the dyad. Please note that applying MLM to the analysis of dyadic data translates to viewing dyads as level-2 units and dyad members as level-1 units. As there are only two data points per dyad, this entails that the regression slope of the MLM cannot be allowed to vary between dyads. The APIM in this case is, thus, a special case of the well-known random-intercept model. Finally, and this relates to both MLM and SEM, the two approaches usually assume interval scaled outcome variables. The SEM can easily handle binary or categorical outcomes. In case of MLMs, we suggest to use an approach based on generalized estimating equations (Loeys, Cook, De Smet, Wietzker, & Buysse, 2014, give an accessible description).

**Dyadic Response Surface Analysis.** Having outlined how one can analyze the social consequences of personality in a dyadic context, let us now turn to the examination of the effect of personality similarity on an outcome variable. In the current literature, there is a variety of ways to measure couple similarity, including discrepancy scores and profile correlations (for an overview, see Dyrenforth et al., 2010; Kenny et al., 2006). As described earlier, RSA has the potential to overcome many limitations of these methods. As far as we know, a fully *dyadic* RSA has not been employed yet. However, its implementation seems relatively straightforward, as an RSA already uses two predictors that can be mapped to the members of the dyad (e.g.,  $x_n$  = husband's predictor,  $y_n$  = wife's predictor). In order to model the outcome of both partners, a dyadic RSA involves *two* second-degree polynomials

$$z_{n1} = b_{01} + b_{11}x_n + b_{21}y_n + b_{31}x_n^2 + b_{41}x_ny_n + b_{51}y_n^2 + e_{n1},$$

and

$$z_{n2} = b_{02} + b_{12}x_n + b_{22}y_n + b_{32}x_n^2 + b_{42}x_ny_n + b_{52}y_n^2 + e_{n2}.$$

(8)

Here,  $z_{n1}$  and  $z_{n2}$  refer to the outcome variable of the first and the second dyad member, respectively,  $x_n$  is the predictor variable of the first dyad member and  $y_n$  is the predictor variable of the second dyad member. In our example,  $z_{n1}$  and  $z_{n2}$  reflect the relationship satisfaction scores of the husband and the wife, respectively,  $x_n$  denotes the husbands' agreeableness scores and  $y_n$  reflects the agreeableness scores of the wives.

Considering the first equation in isolation simply shows the standard polynomial regression that was used in the individual data RSA. The difference is that one of the predictors ( $y_n$ ) is a partner variable. The same set of predictors then is used to predict both members' outcomes. The error terms  $e_{n1}$  and  $e_{n2}$  should be correlated to account for the non-independence of the dyadic data structure. Hence, the dyadic RSA can be seen as an extension of the standard APIM by the squared terms and an interaction term. Again, SEM or MLM can be used to estimate the parameters, and we recommend using the former for distinguishable dyads and the latter for indistinguishable dyads. The resulting coefficients can be used to plot

two surfaces:  $b_{01}$  to  $b_{51}$  for the outcome surface for the first dyad member (e.g., the husbands), and  $b_{02}$  to  $b_{52}$  for the outcome surface of the second dyad member (e.g., the wives). The visualization of two regression surfaces makes evident that the (in)congruence between partners can have different effects on the husband's and the wife's outcome. To test whether the effects are distinguishable between roles or not, one can apply the typical APIM constraints on the regression coefficients, such as equating both partner's actor and/or partner effects.

### **Examining longitudinal changes in social consequences and how personality is related to these changes**

The example data that we have described includes longitudinal assessments of personality (e.g., agreeableness), the consequence variable (e.g., relationship satisfaction), and the mediator (e.g., stress). Similar to the individual data case, there are numerous statistical models that can be employed to analyze this longitudinal dyadic data. In fact, all models that we mentioned earlier can be applied to dyadic data as well. To use the metaphor again, this is accomplished by computing the individual level model twice, one for each member of the dyad. In the following, we will first describe how a *single* longitudinal dyadic variable can be analyzed. In terms of our example, we will explain how we can study whether there are longitudinal changes in the relationship satisfaction of our couples.

**Longitudinal changes in a single dyadic variable.** Two approaches can be used to model a single longitudinal dyadic variable - the autoregressive model and the growth model. The autoregressive model for dyadic data measured at three time-points is displayed in Figure 6. It is basically an ACLM where changes in a *single* variable are modeled, but for the two dyad members. Also, the model can be understood as an APIM for a single variable. The actor effects (denoted as 'a' in Figure 6) are interpreted as stability effects and the partner effects (termed 'p' in Figure 6) represent the longitudinal cross-partner influence. In terms of our example, the actor effects describe the stability of the husband's (wife's) relationship

satisfaction between the assessments. The partner effect, by contrast, indicates the extent to which husband's (wife's) satisfaction predicts the wife's (husband's) relationship satisfaction while controlling for the wife's (husband's) earlier relationship satisfaction.

When the dyads are distinguishable, the dyadic autoregressive model just explained can be most easily estimated using SEM software. We note that all considerations that we mentioned regarding the APIM apply to the autoregressive model as well. In the case of indistinguishable dyads, MLM should be used instead of SEMs. However, there are 'two' outcome variables (one for each dyad member), which makes estimation more complicated. The interested reader is referred to Kenny et al. (2006) for a description of this approach. For more information on the autoregressive path models for dyadic data and their analysis, see Kenny et al. (2006; see also Bolger & Laurenceau, 2013, Hutteman et al., 2013).

Besides the autoregressive models, dyadic growth models can also be used to examine changes in a single dyadic variable. Here, a growth model is computed for each dyad member (see Figure 7, for a graphical illustration), and then the correspondence between the parameters of the growth model is evaluated. We note that as in the case of individual-level data, initial assessments are typically treated as  $time_{it} = 0$  (and we stick to the practice, for instance, in Figure 7), but alternative definitions of the time variable are possible of course. A dyadic growth model studies within-dyad member changes over time and how the differences between the dyad members in this change are related. For instance, one can estimate the degree to which the dyads members' values in a variable (e.g., relationship satisfaction) are similar at the first time-point. This is captured by the intercept-intercept covariance (e.g., the path from  $b_{H1}$  to  $b_{W1}$  in Figure 7). The model also allows examining whether there is an association between the rates of change of the dyad members (e.g., the path from  $b_{H2}$  to  $b_{W2}$  in Figure 7). In our example, the slope-slope covariance denotes whether the rates of change in satisfaction for husbands and wives are similar. Finally, one can also examine whether the other-dyad member's intercept and the slope covary. In terms of our example, one can thus

test whether the wife's (husband's) satisfaction at the first time-point is related to the rate of change of the husbands (the wives).

The dyadic latent growth model can be estimated using SEM and MLM in the case of distinguishable and indistinguishable dyads, respectively. Detailed descriptions of each approach are given in Kashy and Donellan (2008) and Kashy et al. (2008). Note that on a conceptual level, the dyadic latent growth model corresponds to a bivariate growth model.

**Incorporate the influence of personality.** The two approaches described so far served to model the longitudinal changes in one dyadic variable, for example, the social consequence variable. To incorporate the influence of a predictor of change, such as personality, we have to extend the two approaches to account for the dyadic nature of the predictor. In the case of the autoregressive model, this could easily be done by expanding the dyadic autoregressive model for the social consequence variable with a dyadic autoregressive model for the repeatedly measured personality variable. When there are two time-points, this model will contain four actor or stability effects (e.g., the husbands' and the wives' stability in personality and social consequences) and twelve partner or cross-lagged effects. It should be noted, that this model allows one to model rank-order stabilities or between-person changes in the variables. When one is interested in modeling between-dyad member differences, an extension of the growth model can be considered. Alternatively, one can also employ a dyadic latent difference score model.

Concerning the growth model, let us first assume that personality was assessed at just one measurement occasion. The resulting model is depicted in Figure 7. As can be seen, this is an APIM in which the predictor variables are the personality traits of the dyad members and the outcome variables are the slopes and the intercepts, respectively. In terms of our example, the predictors would be the agreeableness scores of the wives and the husbands. These are used to predict the slope and the intercepts resulting from the relationship satisfaction growth model of the wives and the husbands, respectively. In this model, there are four actor effects

reflecting the prediction of a dyad members intercept and slope by its own predictor variable. The interpretation of the resulting coefficients is in accordance with the interpretation of the standard growth model. There are also four partner effects. These reflect the influence of a dyad member's personality on the other member's values in the consequence variable at the first time-point (e.g., the initial relationship satisfaction) and the rate of change in the consequence variable (e.g., the rate of change in relationship satisfaction).

Finally, it is also possible to compute a multivariate growth model, in which the two within-dyad member's changes in personality are related to the two within-dyad member's changes in social consequences. This model allows the simultaneous examination of the longitudinal development of the dyads personality and of the social consequences. However, we do not know, to the best of our knowledge, of any application of this model to empirical data. We note that, again, all three suggested models can be estimated using SEM and MLM (depending on the distinguishability of the dyads).

### **Examining the processes that mediate the personality and social consequences relation**

Ledermann, Macho, and Kenny (2011) proposed a SEM model that assesses mediation in an APIM (called "APIMeM"). This model can be used to model mediation when the input variable (e.g., agreeableness), the mediator (e.g., stress), and the outcome (e.g., satisfaction) have been measured at three consecutive time-points. Basically, the APIMeM computes a mediation model for individual data for each member of the dyad.

One complexity of the APIMeM lies in the fact that each of the four direct effects of the APIM (two actor effects, two partner effects) can be mediated via two pathways. For example, the direct husband actor effect  $x_h \rightarrow y_h$  can be mediated via the husband mediator ( $x_h \rightarrow m_h \rightarrow y_h$ ) or via the wife mediator ( $x_h \rightarrow m_w \rightarrow y_h$ ). In our example data set, the husband's agreeableness could influence the husband's experienced stress, which in turn influences the husband's relationship satisfaction. This mediation path would be completely

within the male. However, it is also plausible that the direct husband actor effect is mediated via the female's stress: The husband's inagreeableness triggers the female's stress which in turn lowers the husband's relationship satisfaction. Hence, although the APIMeM is only a dyadic extension of the basic three-variable mediation model, the inclusion of all possible partner effects and mediation paths leads to a rather complex model that contains twelve estimated path coefficients. Therefore, Ledermann et al. (2011) suggest a detailed step-by-step procedure on how to simplify the model by adding several constraints. An instructive application of the APIMeM can be found in Erol and Orth (2013).

As far as we know, there are at present no methodological publications available that suggest mediation models for dyadic data in which the input, the mediator, and the outcome have been measured at all occasions. However, earlier we used a metaphor to describe the general strategy in the analysis of dyadic data, namely that the individual level model is computed twice, one for each member of the dyad. This reasoning implies that the longitudinal mediation models that we have suggested to analyze individual data can be applied to the analysis of dyadic data as well. However, what makes the analysis complicated is the addition of partner effects, which leads to a large number of paths to be estimated.

For example, one can use a multivariate growth model to examine mediation in situations where the input, mediator, and output were all assessed longitudinally. A mediation analyses can then be performed for the intercepts and slopes within the same dyad member but also for the intercepts and slopes between the two dyad members. For example, we could use the slope of agreeableness of the husbands to predict their slopes in stress, and the slope of the husbands' agreeableness and their stress slope to predict their slope in relationship satisfaction. This mediation path is completely within the husbands. Similarly, one can check whether the effect of the husband's slope for agreeableness on their satisfaction slope is mediated by the wife's stress slope. We note that the suggested mediation model contains *six* univariate growth models. A limitation of this approach, in terms of model convergence, is

thus its complexity. Another limitation of this specific model is that the slopes are not time-specific. If the three variables are assessed at the same time-point, the time sequence of the input, mediator, and the outcome is blurred. Finally, and for the sake of completeness, SEM and MLM can be used to estimate the APIMeM as well as the suggested extensions (e.g., using the R package *lavaan*)

### **Analyzing Group Data**

In the final part of the article, we discuss how one can analyze group data that stems from a round-robin design. Following the approach in the other two sections, we will first describe an example dataset together with some illustrative research questions. Then, we will describe how the social relations model (SRM, Kenny, 1994) and descriptive social network analysis (SNA, e.g., Wassermann & Faust, 1994) can be used to better understand such data. This is followed by a description of how one can examine the social consequences of personality using group data, how one can analyze longitudinal group data and finally, we will speculate about mediation models for group data.

### **Illustrative Examples and Illustrative Research Questions**

In our fictitious study, 50 freshmen were asked to judge every other freshman in this group in terms of liking and to identify whether or not every other freshman was a friend. The liking ratings were collected at three consecutive measurement occasions (e.g., during the first week of the semester, 2 months later and another 2 months later), the friendship ratings, by contrast, at the second and the third occasion only. Furthermore, on each occasion, the freshmen's agreeableness was assessed by asking them to fill out an agreeableness scale. Finally, for the time-interval between the first and the second and between the second and third measurement occasion, respectively, all freshmen were also asked to indicate whether they had interacted with each of the other freshman. Table 6 gives a summary of the example. The (simulated) dataset together with the R code can be downloaded from <http://osf.io/dny58/>.

In our example, researchers have assessed one (round-robin) group. Alternatively, researchers may have examined (more than one) group of people and have asked group members to rate one another. Adapting the freshmen example, a researcher may have assessed thirty groups with each group containing five freshmen. In each group, freshmen provided liking ratings for each of their group members at three consecutive time-points, filled out an agreeableness scale and also indicated whether they interacted with each of the other group member during the first and the second, and the second and the third assessment. We note that all of the approaches that we will describe can handle such data situations, but that we will use our first example to explain the rationale behind the different approaches.

Again, a number of research questions could be examined with this data. For instance, a first question may be whether agreeable freshmen were being liked more by the other freshmen or whether they were more often identified as friends. Given the longitudinal design of the study, one may also examine whether there were any longitudinal changes in the liking of the freshmen or how often they were nominated as friends. One may also be interested in whether the probability of a friendship nomination was reciprocal (e.g., when Anne is nominated by Ben as a friend at the first measurement occasion, will she nominate Ben as a friend at the second occasion), whether agreeable individuals were more often nominated as friends at later time-points, and whether similarity in agreeableness influences friendship selection. Finally, we may also be interested in whether the effect of agreeableness on being liked at a later time-point or being more often identified as friends was due to having more conversations with such individuals. Before we turn to a description of some of the concrete approaches to tackle such questions, we first make some general remarks concerning the analysis of group data by introducing the SRM and descriptive SNA.

### **Social Relations Model and Social Network Analysis**

The data collection approach described in our example is an instance of a round-robin design (Kenny, 1994). Generally, a round-robin design can be used to obtain social relations

data and social network data (Kenny et al., 2006; Wassermann & Faust, 1994). It is a design where each person rates every other person and where each person is rated by every other person. In our example, each freshman rated every other freshman in terms of liking (or identified whether or not each other freshman was a friend) and was rated by each of the other freshmen concerning liking (or concerning friendship). The advantage of round-robin designs is that they allow a researcher to untangle different components that affect interpersonal judgments. Consider, for example, Anne's liking-judgment of Ben. At the individual level, this judgment is influenced by how strong Anne likes all other individuals in general, by how strong Ben is liked by all other individuals, and by Anne's unique liking of Ben.

Conceptually, these effects exist in any interpersonal phenomenon: Every single datum (e.g., Anne's liking judgment) is a mixture of them, and the SRM and SNA allow unraveling them.

**Social Relations Model.** The SRM is a mathematical model that aims to describe the components of interpersonal judgments. In the simplest case, the judgment of an individual concerning another individual (e.g., Anne's liking judgment of Ben) is decomposed into a perceiver effect, a target effect, and a relationship effect. Table 7 presents an exemplary data excerpt. Here, every row keeps one directed dyadic relationship. Concerning the liking judgment, for example, Anne likes Ben (e.g., the liking judgment is 4 on a scale ranging from 1 to 5), but this liking is not reciprocated, as Ben does not like Anne (liking = 1). From the data points shown in Table 7, one can infer that Anne has a high *perceiver effect* as she tends to like others in general (i.e., most of her liking judgments are high). Ben, in contrast, has a low perceiver effect indicating that he tends to dislike others in general. The *target effect* for Anne is low (i.e., most liking judgments concerning Ann are low) denoting that on average Anne is not liked by others. Finally, the eighth row demonstrates a high *relationship effect*: Chris likes Ben very much, and this liking cannot be attributed to the perceiver effect of Chris, because he does not tend to like others on average. Likewise, this particular liking cannot be due to the target effect of Ben (he is not liked by the other freshmen on average).

The SRM allows a researcher to compute the person-level variables (i.e., perceiver and target effects) and the dyadic-level variables (i.e., relationship effect) for each participant and each dyad, respectively. The SRM also provides variance estimates of these effects. In terms of our example, these estimates quantify the amount of variation in liking that can be attributed to perceiver effects, target effects, or relationship effects. Beside these variances, one can also estimate the covariance between the perceiver and target effects. For the liking example, a positive perceiver-target covariance would indicate that freshmen who like others in general are also liked more by others. Moreover, a relationship covariance can be computed. It describes the extent to which the relationship effects of the judgments of the same dyad, for example, Anne's likeability judgment of Ben and Ben's likeability judgment of Anne, are correlated. Finally, if one has assessed multiple groups (e.g., five freshmen in ten groups) one can also estimate a group variance indicating the extent to which groups differ in the measured construct (e.g., liking).

Self-ratings can also be assessed in a round-robin design (e.g., freshmen are asked to indicate how much they like themselves). They are, however, not included into the SRM analysis because it is assumed that self-ratings are conceptually different from other-ratings. However, they can be included in subsequent analyses by relating them to the perceiver effects or the target effects of the participants. For example, one could correlate the self-ratings of liking with the perceiver effects of liking to examine whether freshmen like themselves as they like others (this is called *assumed similarity* by Kenny, 1994).

To summarize, the SRM is a mathematical model allowing researchers to untangle important components that affect interpersonal judgment. For a more detailed treatment of the theoretical foundations of the SRM, the interested reader is referred to Back and Kenny (2010) or Kenny (1994; more detailed). A number of statistical approaches exist concerning the computation of the SRM variances and SRM effects. The R package *TripleR* (Schönbrodt, Back, & Schmukle, 2012, 2014) and the FORTRAN-based *SOREMO* (Kenny, 1998)

implement the so-called ANOVA approach for estimating the SRM parameters. *TripleR* automatically handles missing values and can compute SRM effects when two indicators of a variable have been assessed (e.g., two liking measures).<sup>2</sup> *SOREMO*, by contrast, does not allow for missing data but can be also used in the case of an arbitrary number of round-robin variables. Finally, both programs can handle data stemming from multiple groups.

The ANOVA approach, however, is not without problems (see Kenny et al., 2006), and this has led to the development of maximum likelihood (e.g., Snijders & Kenny, 1999), restricted maximum likelihood (Nestler, 2014b), and Bayesian estimation approaches (e.g., Lüdtke, Robitzsch, Kenny, & Trautwein, 2013). All of these approaches are based on embedding the SRM into the framework of mixed models (Verbeke & Molenbergh, 2009). Specifically, the SRM is a mixed model with correlated crossed random factors (Judd, Westfall, & Kenny, 2012) that has a very specific residual correlation structure. We note that MLMs are also special cases of the mixed model. Although similar, MLMs and the SRM are not the same, which makes standard multilevel estimation inappropriate for the SRM (but see Snijders & Kenny, 1999; Kenny & Livi, 2009). At the time of writing this article, the two maximum likelihood estimators and the Bayesian approach have not been implemented in an R package (although it is intended to include them into *TripleR*). The interested reader should consult the *TripleR* homepage (<http://www.persoc.net/Toolbox/TripleR>) and David Kenny's homepage (<http://davidakenny.net/>) for updates, further references and solutions.

**Descriptive Social Network Analysis.** Another way to analyze round-robin data is to use social network analysis (SNA; Kenny et al., 2006; Wassermann & Faust, 1994). One way to conduct SNA is to measure the structure of the network by computing a number of network statistics, such as the network density. Alternatively, a researcher can use statistical models to analyze social network data. Exponential random graph models (ERGMs; Robins & Lusher, 2013) or stochastic actor-based models (often called SIENA; Snijders, 2005) are two of the most well-known statistical approaches. Here, we describe descriptive SNA. We will turn to

the statistical models later when we explain how the social consequences of personality can be examined with cross-sectional and longitudinal network data.

A social network is a structure composed of a set of persons and a set of ties between the persons (Knoke & Yang, 2008). The ties represent a specific relation between the members: If Anne has nominated Ben as a friend at a measurement occasion then there would be a tie between Anne and Ben in the friendship network for this occasion. A friendship network for this example is presented in Figure 8. Here, the nodes represent persons (e.g., the freshmen) and the arrows represent ties (e.g., a friendship nomination). A double-headed arrow indicates mutual nominations (e.g., both freshmen made a nomination).

Social networks can be undirected or directed. In undirected networks, two nodes are either connected or they are not (e.g., Anne and Ben stand in the same room, Anne and Ben are married). In graphical representation of such networks, ties would therefore not be represented by arrows but by lines. In directed social networks, by contrast, relations are represented that can be reciprocated (e.g., a friendship tie from Anne to Ben, which can, or cannot, be reciprocated from Ben to Anne). Furthermore, social network research typically considers binary ties (see Kenny et al., 2006; Wassermann & Faust, 1994), although the strength of a tie can also be measured on an interval scale (see Knoke & Yang, 2008, for treatments of such valued networks). Finally, two kinds of network structures can be investigated (Knoke & Yang, 2008). In the study of *egocentric networks*, a person is asked for all other persons with whom s/he has ties (e.g., her friends) as well as all the ties among these other persons are assessed (e.g., which of her friends are also friends). In a *complete network*, by contrast, there is a fixed set of persons, and each person is asked for their ties to each other person. The friendship nominations of our example are thus an instance of a complete network with binary and directed ties.

The round-robin design matches a complete network structure. Similar to the SRM, SNA allows the computation of *network member variables* (Costantini et al., in press) that

indicate the importance of a person in the network. In a binary directed network one can, for example, compute the number of ties a person receives from all other persons. This measure is called an individual's *indegree* and it is a measure for the popularity of that person. One can also count the number of ties given by that person to the other persons. This is the *outdegree* of an individual and refers to the person's expansiveness.

SNA can also be used to obtain *overall network measures*, such as the density of a network. The density equals the proportion of ties in a network that are actually present. In the example graph, there are 5 ties and 4 times 3 possible ties, resulting in a density of  $5/12 = .42$ . The density of a network is often used to explore the cohesion in the overall network. When interpreting density values one should consider the magnitude of the density and the size of the network as they are related: The larger the number of actors in the network, the more difficult is it to obtain high density values. The density also depends on the construct the tie refers to: friendship networks, for example, tend to be sparse (densities range between .004 und .12; see van Zalk, Branje, Kerr, & Stattin, 2010).

Finally, one can also analyze *triads* and *dyads* in the network. A triad consists of three persons and the ties between them. A central goal in the analysis of triads is to determine how many of them are transitive. When Anne identifies Ben as a friend, and Ben identifies Chris as friend, this triad would be transitive when Anne identifies Chris as a friend too. For the *dyads* in a network, one can count the number of *mutual* dyads (e.g., Anne sends a tie to Ben *and* Ben sends a tie to Anne, one dyad in the Figure), of *asymmetric* dyads (Anne sends a tie to Ben *or* Ben sends a tie to Anne, but not both; three in the example graph), and the number of dyads without any ties (eight in the example graph).

With this brief introduction of some network measures, we have only scratched the surface of descriptive SNA. There are a number of introductory texts for the descriptive analysis of social networks that the interested reader is referred to, including for example Knoke and Yang (2008), Wasserman and Faust (1994; more advanced), or Costantini et al. (in

press). There are several R packages (e.g., *igraph*, Csardi & Nepusz, 2006; *statnet*, Handcock, Hunter, Butts, Goodreau, & Morris, 2008) and commercial software programs (*UCINET*, Borgatti, Everett, & Freeman, 2002) available for the computation of the descriptive network measures. Finally, when one has assessed multiple groups, various methods exist that could be employed to combine the within-group measures (e.g., the mean number of indegrees per group) to get an overall descriptive measure. One can compute, for example, a simple aggregate of the within-group measures or a weighted aggregate (e.g., the weights are the standard errors of the within-group measures). We refer the reader to Snijders and Bosker (2012) for a brief explanation of various combination methods.

Before closing, the interested reader might wonder about the relation between the SRM and descriptive SNA given that SRM target effects (perceiver effects) are similar to SNA indegrees (outdegrees). In fact, in our experience, the respective SRM and SNA measures are highly correlated. However, the standard SRM, first, assumes an interval level of measurement while most SNA models assume binary data. Second, the standard SRM controls for the missing self-rating when person-level effects are computed. This is not done in descriptive SNA when indegrees or outdegrees are computed. Please notice that this latter point is not true in the ERGM-and the SIENA-framework because both account for missing values during estimation. Finally, SNA allows a researcher to analyze triads and other higher-level network structures (e.g., structural holes in a network). This is not possible in the SRM-framework (but see Bond, Horn, & Kenny, 1998; Hoff, 2005).

### **Examining the social consequences of personality**

Let us now describe how the social consequences of personality can be examined using group data. How can we check, for instance, whether agreeableness goes along with higher target effects in liking or higher indegrees in friendship nominations? The easiest way to answer such questions is to compute the person-level effects with the help of the SRM or SNA and to use the statistical approaches that we presented in the section about individual

data. For the effect of agreeableness on popularity in liking, one can use *TripleR* to compute the target effects and then test in a regression whether agreeableness predicts these effects. Similarly, one of the SNA packages can be employed to compute the indegrees of each freshman in the case of popularity for friendship nominations. These values can then be regressed on agreeableness.

As far as we know, this two-step approach (e.g., compute the SRM effects, use these effects in the second-stage analysis) is always used in the case of the SRM and very often used in the case of SNA, at least in the few circumstances in which SNA has been employed in personality psychology (e.g., Clifton & Kuper, 2011). However, it should be noted that from a statistical point of view, the two-step approach can be problematic because using the first-step estimates in the second-step analyses can limit the accuracy of the results (Verbeke & Molenberghs, 2009). This decrease in accuracy (and the resulting loss in power) occurs as the second-step analyses do not consider the error variability that is introduced by estimating the SRM (SNA) effects (Lüdtke et al., 2013; Nestler, 2014b).

For the SRM, a solution to this problem is to employ one of the newer estimation approaches (maximum likelihood, restricted maximum likelihood, Bayes; see Lüdtke et al., 2013; Nestler, 2014b; Snijders & Kenny, 1999) that embed the SRM into the mixed model. All of these approaches allow a researcher to include predictors of the SRM effects into the model, and hence to test their influence with the raw group data. We note that the inclusion of personality into the SRM mixed model is similar to including group-level predictors for the intercept in MLM (but this is just an analogy; MLMs and the SRM are not the same).

With regard to SNA, established one-step approaches exist that allow testing the effects of personality on social consequences with the help of the raw network data. In our example, these approaches thus allow to examine the effect of agreeableness on the emergence of the observed friendship network. We describe the exponential random graph model (ERGM) here.

**ERGMs.** Descriptive social network analysis allows the researcher to understand a social network by computing the values of certain network configurations. For each person in the network, we may compute the indegree and the outdegree, and for the dyads, we may compute the number of mutual dyads. Having obtained these values, a natural next step is to examine which of the network configurations are important for the emergence of the observed social network. In the case of the friendship network, for example, one may wonder whether the tendency to reciprocate friendship ties or to close friendships transitively can explain the observed pattern of ties.

A statistical approach (Robins & Lusher, 2013; Wasserman & Faust, 1994) to examine whether a particular network configuration (i.e., the number of mutual dyads) is important for the emergence of the network, may consist of simulating a large sample of random networks in which all other network configurations are held constant (e.g., same density and same number of transitive triads). For each random network, the value of the particular network configuration is computed (e.g., the number of mutual dyads), and the average of these values can then be compared with the actually observed value of the particular network configuration. When the actual value is significantly greater than the average, there are more particular network configurations than can be expected by chance, indicating that the underlying social process is important for the emergence of the observed social network. If, for example, the observed network has ten mutual dyads, whereas the average of the randomly generated networks with the same density and number of transitive triads, has only three such dyads, then reciprocity seems to be an important feature of the observed network.

However, a critical issue concerning this approach is that the average value depends on the presumptions that one has made prior to the simulation (e.g., fixing both the density and the number of transitive triads). If one had made other assumptions (e.g., only the density is held constant), other random networks with different values of the particular network configuration might have occurred. Furthermore, the approach allows examining the effect of

only one network configuration. This is problematic, as network configurations are dependent upon another. The higher the number of transitive triples, for instance, the more probable are mutual dyads in a social network.

ERGMs try to solve both issues and at the same time stick to the logic of randomly generating networks. In ERGMs, the probability of the observed network  $y$  is modeled with the help of an exponential function:

$$P(Y = y) \propto \exp[b_1 z_1(y) + b_2 z_2(y) + \dots + b_p z_p(y)]. \quad (9)$$

Here, each  $z_i(y)$  refers to a network configuration that is weighted by a parameter  $b_i$ ,  $i = 1, \dots, p$ . Specifically, the  $z$ s are the values of the network configurations in the observed network that one believes are important for the occurrence of the observed network. In the case of the friendship network, for example,  $z_1$  may reflect the number of edges,  $z_2$  may be the number of mutual dyads, and  $z_3$  may be the number of transitive triads. Besides modeling such structural network effects, ERGMs also allow one to test for exogenous variables, such as personality, by including network configurations weighted by the values of this variable. For example, researchers interested in the effect of agreeableness on the indegrees of freshmen, can include the freshmen indegrees weighted by their agreeableness values. Such a network configuration is called an actor attribute effect as it reflects the effect of an attribute (e.g., agreeableness) on the actor (e.g., a freshman). Finally, we note that one can also include dyadic attribute effects (such as the number of interactions of each dyad).

The goal of fitting an ERGM to the observed network is to estimate the  $b_i$  parameter of each included network configuration. These parameters, just as in a regression, describe the importance of the particular network configuration for the emergence of the observed network: The higher  $b_i$ , the higher is the prevalence of network configuration  $z_i$  in the network. If the parameter for the number of mutual dyads, for example, is positive, then there are many mutual dyads (and hence a tendency to reciprocate) in the network. Furthermore, a positive value for an actor attribute effect would indicate that persons with higher values in

the attribute (e.g., agreeableness) receive more ties (e.g., are more often identified as friends). Importantly, and also similar to regression, the value of a parameter estimate denotes the influence of the respective network configuration on the emergence of the network while holding all other network configurations constant.

The foregoing explanations outlined a network-level interpretation of the parameters  $b_i$ . However, Equation 9 can be transformed to a tie-level model (see Wasserman & Pattison, 1996, for the transformations), suggesting a tie-based interpretation of the parameters  $b_i$ . The transformed model is given by

$$P(X_{ij} = 1 | X_{ij}^c) = (\text{logit}[b_1\Delta(z_1(x)) + b_2\Delta(z_2(x)) + \dots + b_p\Delta(z_p(x))])^{-1} \quad (10)$$

Here,  $X_{ij}$  represents the tie between dyad  $ij$ ,  $X_{ij}^c$  denotes the ties of all other dyads in the network,  $b_i$  are the parameters, and  $\Delta(z_i(x))$  is the *change* in the value of a network configuration when one alters the tie of the dyad  $ij$  from 0 to 1. Equation 10 gives the probability of the tie between  $i$  and  $j$  conditional on all other ties in the network. The parameter  $b_i$  indicates how much the probability of a tie is *increased* by a particular network configuration. A positive parameter for the number of mutual dyads thus indicates for example, that when there is a tie between  $j$  and  $i$ , then there is a higher probability of the tie between  $i$  and  $j$  occurring (i.e., the friendship tie is more likely to be reciprocated).

Equation 10 might be familiar to the reader as it is similar to the logit formulation of the logistic regression model.<sup>3</sup> In fact, this similarity has led early network researchers to estimate the ERGM parameters with standard logistic regression techniques (called *pseudolikelihood* estimation in the ERGM literature). Here, each tie of the network is treated as a case of the outcome variable and the change statistics become the independent variables. This estimation approach, however, is problematic since the standard logistic regression model presumes cases to be independent; an assumption typically not met in network research. Alternative estimation approaches have therefore been developed that are based on

Monte Carlo techniques (*Monte Carlo maximum likelihood estimation*). All of these approaches simulate a large sample of random networks. Specifically, they try to find the parameters  $b_i$  so that the average of each modeled network configuration in the simulated random networks equals the value of the respective network configuration in the observed social network.

In ERGMs, researchers can include network configurations deemed important for the formation of the observed network. After they have fitted an EGRM to the network data, they can test the goodness of their model by assessing how well the fitted model captures features of the observed network that were *not* explicitly modeled. This process of model checking is called GOF (goodness of fit) in the social network literature (Koskinen & Snijders, 2013). For example, when the ERGM for the friendship network includes the number of ties, the number of mutual dyads, and the number of transitive triplets, then GOF would check whether this model can reproduce the observed distribution of indegrees.

To sum up, ERGMs are a powerful tool to examine which variables (e.g., network configurations, actor attributes or dyadic attributes) may help to explain the observed network. Interestingly, we know no empirical psychological research that employed ERGMs. The interested reader is referred to Carrington, Scott, and Wasserman (2005) for an advanced treatment of ERGMs. For a non-technical introduction together with empirical illustrations, we recommend Lusher, Koskinen, and Robins (2013). The R package *ergm* (Hunter, Handcock, Butts, Goodreau, Morris, 2008) can be used for the estimation of ERGMs and GOF testing. Finally, when the network data stems from multiple groups, the parameters can be combined (by using one of the combination methods explained, for example, in Snijders & Bosker, 2012) or by treating group as a random factor and computing a multilevel version of the ERGM (see Wang, Robins, Pattison, & Lazega, 2013).

**Examining longitudinal changes in social consequences and how personality is related to these changes**

The last section described how one can examine the effects of personality on social consequences in case of *cross-sectional* group data. When a researcher has collected longitudinal group data, the easiest approach is to compute the respective SRM effects or SNA measures (such as the indegree) and use the longitudinal models described earlier. This two-step approach can be statistically problematic because of the reasons mentioned above (Step 2 depends on accurate and equally reliable estimates in Step 1). However, for the SRM, this claim is troublesome because we know of no approaches that allow researchers to model linear or nonlinear growth of the SRM effects with raw data. We believe that the mixed-model approach to SRMs can be modified for such situations by including a time variable that predicts the repeated round-robin judgments of a single dyad. The slope of the time variable can be allowed to contain a perceiver effect (i.e., reflecting changes in the perceiver effect), a target effect (i.e., reflecting changes in the target effect), and a relationship effect (i.e., the unique liking changes). We note that this approach allows the inclusion of variables that do not vary over time (e.g., such as agreeableness at the first time-point) to model between-person differences in within-person changes of the SRM effects. To draw the analogy to MLM again, the suggested approach translates to having the repeated measures of a single dyad at level 1, the dyad members at level 2 (e.g., perceiver or target), and variables to explain between-person variability in the effect of time at level 3.

**Longitudinal Network Models.** There is a better initial situation for longitudinal social network data as there are a number of approaches for the longitudinal modeling of networks (see Snijders, 2005, for an overview). Exponential random graph models, for example, have recently been extended to cover network panel data (e.g., Krivitsky & Handcock, 2014). However, in this section we concentrate on the *stochastic actor-based model* introduced by Snijders and colleagues (see Snijders, Steglich, & van de Bunt, 2010, for a non-technical introduction, and Snijders, 2005, for a more technical treatment), because it is currently the most often employed approach to analyze longitudinal social network data. The

stochastic actor-based model is implemented in the R package *RSIENA* (Ripley, Boitmanis, & Snijders, 2013) and it is often simply termed *SIENA* in the network literature. Typically, *SIENA* is applied to binary directed complete networks that have to be measured at two occasions at least. For applications to other networks, interested readers are referred to the *SIENA* homepage (<http://www.stats.ox.ac.uk/~snijders/siena/>).

To give a brief introduction to *SIENA*, we assume that two measurements of a binary and directed network exist (e.g., the data of the friendship networks that were assessed at the second and the third measurement occasion). *SIENA* models these longitudinal social network data by simulating a stochastic process taking place between the two measurement occasions. The network measured at the first time-point is taken as the starting value of the process (i.e., it is not modeled) and the network measured at the second time-point ideally should be the end point of the process. Whether the latter is true, however, depends on the adequacy of the model that was specified by the researcher.

As in the case of ERGMs, investigating the evolution of the network presumes that researchers include network configurations deemed important for this evolution. Again, these network configurations can be purely structural effects, actor attributes, or dyadic attributes. For the friendship network, for example, one may include an outdegree effect and a reciprocity effect (i.e., network effects), the effect of agreeableness on the outdegrees and the indegrees (i.e., actor attributes), and the number of interactions for each dyad (i.e., a dyadic attribute). *SIENA* then estimates the influence of the specified effects for the evolution of the network by stochastically simulating *actor-based network changes between* the two measurements of the network.

The basis of this simulation is the initial network and initial values of the model parameters  $b_i$  that describe the impact of the assumed effects (e.g., reciprocity). The first step of the stochastic simulation then consists of randomly selecting an actor who gets the opportunity to make a change to the network. In *SIENA*, this step is modeled by a rate

function  $\rho_i$ , denoting the change rate of actor  $i$ . This *rate function* is used to compute actor-specific waiting times  $\tau_i$ :

$$\tau_i \propto \exp(-\rho_i). \quad (11)$$

Equation 11 shows that the higher  $\rho_i$ , the smaller is the waiting time of actor  $i$  and the higher is the probability that the actor is selected to change a tie. In typical applications of SIENA,  $\rho_i$  is assumed to be constant for all actors (Snijders, 2005, Steglich, Snijders, & West, 2006; actors are selected with equal probability). However, it is also possible to examine the effect of actor attributes on the probability of change (see Snijders et al., 2010, for more on this).

Having randomly selected an actor, say Anne in our friendship network example, the simulation then determines the probability of the actor to change a tie to one of the other actors in the network. Anne, for instance, can either make no change to the network or she can change it for example by sending a tie to Ben, withdraw the tie to Chris, and so on. The probability of each of these changes can be modeled by

$$P(X_{i \rightarrow j}) \propto \exp(f_i(x_{i \rightarrow j})), \quad (12)$$

where  $X_{i \rightarrow j}$  denotes the network that emerges when actor  $i$  changes the tie to actor  $j$ , and  $f_i$  is the objective function that depends on the modeled network configurations weighted by the model parameters  $b_i$

$$f_i(x_{i \rightarrow j}) = b_1 z_1(x_{i \rightarrow j}) + b_2 z_2(x_{i \rightarrow j}) + \dots + b_p z_p(x_{i \rightarrow j}). \quad (13)$$

Equation 13 shows that the higher the value of the objective function, the higher is the probability of the respective change, and the chosen change has the highest probability.

Both steps, the selection of an actor and the selection of the tie change that is made define a new network that either equals the initial network (i.e., the actor decides to do nothing) or a network where one of the ties of the initial network has changed. These two micro-steps are repeated for a finite number of steps until a *model-implied* network for the second measurement occasion emerges. Ideally, this model-implied network equals the

observed network at the second time-point. When this is not the case, however, the comparison of the two networks can be used to improve the estimates of the model parameters  $b_i$ . With this new set of parameters, the simulation is started again, and this procedure goes on until a final set of parameter estimates  $b_i$  is found.

Just as in regression, the model parameter estimates in a SIENA model mirror the influence of a network effect while holding all other effects constant. Furthermore, parameter values indicate whether the actors in the network are motivated to attain the state that is represented by a certain configuration. A negative effect of the outdegree, for example, indicates that actors do *not* strive to establish ties to unspecific others (e.g., they do not establish friendship with anyone). A positive reciprocity effect, by contrast, indicates that the actors in the network tend to reciprocate the relations (e.g., friendship). Finally, a positive parameter for an attribute effect denotes a positive effect of the attribute on friendship selection (e.g., agreeable persons select more friends).

Before closing, three things are worth mentioning concerning SIENA. First, in SIENA researchers include the effects that they think can explain the formation of the observed network. As in the case of ERGMs, the suitability of this model can be tested by measuring the goodness of fit in modeling features of the observed network that were *not* included in the model specification (e.g., the distribution of indegrees). Furthermore, an overall test of model fit has been developed to test the adequacy of the proposed model (Schweinberger, 2013). Second, SIENA is a powerful model to examine the *simultaneous* influence of network effects, actor attributes and dyadic attributes on the evolution of a single network (e.g., friendship). However, in many situations not only is the network changing as a function of the aforementioned effects, but the attributes or behaviors (e.g., such as agreeableness) of the actors might also be changing as a result of the network (cf. Snijders et al., 2010). SIENA allows to model such parallel developments (i.e., this is called *co-evolution* of the network and the attribute in the network literature). In the case of our example, this approach would

allow to simultaneously examine changes in agreeableness and changes in friendship nominations. We refer the reader to Snijders, Steglich, and Schweinberger (2007; see also Steglich, Snijders, & Pearson, 2010) for a theoretical introduction and van Zalk, Kerr, Branje, Stattin, and Meeus (2010) for an application. Finally, SIENA can treat network panel data that was obtained for several groups. Again, this can be done by combining the estimates of the different groups or by computing a multilevel version of SIENA. We refer the reader to the SIENA manual for more information.

To summarize, SIENA offers several advantages over traditional longitudinal data analysis approaches. First, SIENA allows the *simultaneous* examination of network effects, actor attribute effects and dyadic covariate effects. Second, it can be used to study different actor attribute effects in *one* model. In our example, one can, for example, test whether agreeable individuals are more often nominated as friends but also whether they are more motivated to select friends. Finally, SIENA allows a researcher to model the parallel development of the social network and the attribute. Overall, thus, SIENA provides tests of very realistic models of interpersonal developmental processes (Selfhout, Burk et al., 2010).

### **Examining the processes that mediate the personality and social consequences relation**

Having introduced how a researcher can examine the effect of personality on social consequences with cross-sectional and longitudinal group data, the natural next step would be to describe methods to investigate mediation in the case of group data. How can we test, for example, whether the influence of agreeableness on popularity change (in terms of liking or friendship nominations) is due to the number of interactions between the freshmen? Again, the easiest way to answer such questions is to compute the person-level effects (for the input, the mediator, and the outcome) and to use the statistical approaches that we described in the individual data section. Earlier we mentioned, however, that this approach is statistically problematic and we recommended using the raw data to examine one's research questions.

For the SRM this recommendation is, at the time of writing this article, problematic as there are, as far as we know, no methodological publications that suggest how to test for longitudinal mediation with the raw group data. Certainly, the mixed model approach to the SRM can be extended to examine mediation effects. However, due to the structure of the model (i.e., correlated random intercepts), a critical issue is how one can estimate the standard error of the mediation effect. We are confident that future research will extend current mediation models to more complex models such as the SRM, but at present, no such formulas have been carved out.

There are also no methodological publications available, to the best of our knowledge, that deal with mediation in case of statistical social network models. Notwithstanding, the well-known Baron and Kenny (1986) approach to test for mediation can and has been used (see van Zalk, Kerr et al., 2010) in the (longitudinal) social network literature. Here, the first step involves showing that the input variable influences the outcome (e.g., agreeableness goes along with more friendship nominations later in time). In a second step one tests whether the input influences the mediator (e.g., agreeableness and number of interactions). When the mediator, third, significantly predicts the outcome (e.g., number of interaction and changes in friendship nominations) while at the same time, fourth, the effect of the input on the output is significantly reduced, then this is evidence for mediation. We note that the Baron and Kenny approach has been heavily criticized in the mediation literature (MacKinnon, 2008), but that we do not know of any work that has used the nowadays favored product of coefficients approach (i.e., computing the product of the coefficients that result from the second and third step mentioned above).

### **Conclusion**

Given their rich data-sets and a plethora of research questions, personality psychologists are more often faced with complex statistical approaches they can use to tackle these questions. Moreover, there has also been an explosion of research in the methodological

field providing researchers with new sets of analytical tools. This article aimed to provide an overview of both established and novel statistical approaches to examine social consequences of personality. As our exposition has hopefully demonstrated, each of these approaches has its merits and can be used to study different aspects of interesting research questions. We believe that the knowledge of innovative tools has the potential to broaden the view about possible research questions, and that researchers are able to properly analyze their existing research questions. Existing data sets can be explored and harvested with the new tools, and data collection of new studies can be planned with the new analytical approaches in mind.

## References

- Back, M. D., & Kenny, D. A. (2010). The social relations model: How to understand dyadic processes. *Social and Personality Psychology Compass*, *4*, 855–870.
- Back, M. D., Schmukle, S. C., & Egloff, B. (2010). Why are narcissists so charming at first sight? Decoding the narcissism-popularity link at zero acquaintance. *Journal of Personality and Social Psychology*, *98*, 132-145.
- Baron, R. M., & Kenny, D. A. (1986). The Moderator-Mediator variable distinction in social psychological research – Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173–1182.
- Bates, D., Maechler, M., Bolker, B. M., & Walker, S. (2014). *lme4: Linear mixed-effects models using Eigen and S4*. Retrieved from <http://cran.r-project.org/web/packages/lme4>.
- Biesanz, J. C. (2012). Autoregressive longitudinal models. In R.H. Hoyle (Ed.), *Handbook of Structural Equation Modeling* (pp. XX-XX). New York: Guilford Press.
- Blandon, A. Y., Calkins, S. D., Grimm, K. J., Keane, S. P., & O'Brien, M. (2010). Testing a developmental cascade model of emotional and social competence and early peer acceptance. *Development and Psychopathology*, *22*, 737-748.
- Blanton, H., Jaccard, J., Gonzales, P. M., & Christie, C. (2006). Decoding the implicit association test: Implications for criterion prediction. *Journal of Experimental Social Psychology*, *42*, 192–212.
- Bolger, N., & Laurenceau, J-P. (2013). *Intensive longitudinal methods: An introduction to diary and experience sampling research*. New York: Guilford.
- Bollen, K.A., & Curran, P.J. (2006). *Latent curve models: A structural equation approach*. Hoboken, NJ: Wiley.
- Bond, C. F., Horn, E. M., & Kenny, D. A. (1997). A model for triadic relations. *Psychological Methods*, *2*, 79-94.

- Borgatti, S. P., Everett, M. G., & Freeman, L. C. (2002). *UCINET for Windows: Software for social network analysis*. Harvard, MA: Analytic Technologies.
- Browne, M., Du Toit, S. H. C. (1991). Models for learning data. In: L. Collins & J. L. Horn (Eds.). *Best methods for the analysis of change* (pp. 47–68). Washington, DC: American Psychological Association.
- Campbell, L., & Kashy, D. A. (2002). Estimating actor, partner, and interaction effects for dyadic data using PROC MIXED and HLM: A user–friendly guide. *Personal Relationships, 9*, 327–342.
- Carrington, P., Scott, J., & Wasserman, S. (2005). *Models and methods in social network analysis*. New York: Cambridge University Press.
- Clifton, A., & Kuper, L.E. (2011). Self-reported personality variability across the social network is associated with interpersonal dysfunction. *Journal of Personality, 79*, 359-389.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Lawrence Erlbaum Associates: Hillsdale, New Jersey.
- Cook, W. L. (1994). A structural equation model of dyadic relationships within the family system. *Journal of Consulting and Clinical Psychology, 62*, 500–509.
- Cook, W. L. (2005). The SRM approach to family assessment. *European Journal of Psychological Assessment, 21*, 216–225.
- Cook, W. L., & Kenny, D. A. (2005). The actor-partner interdependence model: A model of bidirectional effects in developmental studies. *International Journal of Behavioral Development, 29*, 101-109.
- Costantini, G., Epskamp, S., Borsboom, D., Perugini, M., Möttus, R., Waldorp, L. J., & Cramer, A. O. J. (in press). State of the aRt personality research: A tutorial on network analysis of personality data in R. *Journal of Research in Personality*.

- Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *International Journal of Complex Systems, 1695*.
- Dyrenforth, P. S., Kashy, D. A., Donnellan, M. B, Lucas, R. E. (2010). Predicting relationship and life satisfaction from personality in nationally representative samples from three countries: The relative importance of actor, partner, and similarity effects. *Journal of Personality and Social Psychology, 99*, 690-702.
- Edwards, J. R. (2001). Ten difference score myths. *Organizational Research Methods, 4*, 265–287.
- Edwards, J. R. (2002). Alternatives to difference scores: Polynomial regression analysis and response surface methodology. In F. Drasgow & N. W. Schmitt (Eds.), *Advances in measurement and data analysis* (pp. 350–400). San Francisco: Jossey-Bass.
- Edwards, J. R., & Cable, D. M. (2009). The value of value congruence. *Journal of Applied Psychology, 94*, 654-677.
- Edwards, J. R., & Parry, M. E. (1993). On the use of polynomial regression equations as an alternative to difference scores in organizational research. *Academy of Management Journal, 36*, 1577–1613.
- Enders, C. K. (2010). *Applied missing data analysis*. Guilford Press.
- Erol, R. Y., & Orth, U. (2013). Actor and partner effects of self-esteem on relationship satisfaction and the mediating role of secure attachment between the partners. *Journal of Research in Personality, 47*, 26–35.
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G\*Power 3.1. *Behavior Research Methods, 41*, 1149-1160.
- Ferrer, E., & McArdle, J.J. (2003). Alternative structural models for multivariate longitudinal data analysis. *Structural Equation Modeling, 10*, 493-524.
- Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.

- Gonzalez, R., & Griffin, D. (1997). On the statistics of interdependence: Treating dyadic data with respect. In S. Duck (Ed.), *Handbook of personal relationships: Theory, research, and interventions* (2nd ed., pp. 271–302). New York, NY: Wiley.
- Grimm, K. J. (2007). Multivariate longitudinal methods for studying developmental relationships between depression and academic achievement. *International Journal of Behavioral Development, 31*, 328-339.
- Grimm, K. J., Ram, N., & Hamagami, F. (2011). Nonlinear growth curves in developmental research. *Child Development, 82*, 1357-1371.
- Grimm, K. J., Steele, J. S., Ram, N., & Nesselroade, J. R. (2013). Exploratory latent growth models in the structural equation modeling framework. *Structural Equation Modeling, 20*, 568-591.
- Grimm, K. J., & Widaman, K. F. (2010). Residual structures in latent growth curve modeling. *Structural Equation Modeling, 17*, 424-442.
- Handcock, M. S., Hunter, D. R., Butts, C. T., Goodreau, S. M., & Morris, M. (2003). *statnet: Software tools for the Statistical Modeling of Network Data*. <http://statnetproject.org>
- Hoff, P. D. (2005). Bilinear mixed-effects models for dyadic data. *Journal of the American Statistical Association, 100*, 286–295.
- Hunter, D. R., Handcock, M. S., Butts, C. T., Goodreau, S. M., & Morris, M. (2008). ergm: A package to fit, simulate, and diagnose exponential-family models for networks. *Journal of Statistical Software, 24*, 1-30.
- Hutteman, R., Bleidorn, W., Keresteš, G., Brkovic, I., Butkovic, A., & Denissen, J.J.A. (2013). Reciprocal associations between parenting challenges and parents' personality development in young and middle adulthood. *European Journal of Personality, 28*, 168-179.

- Jackson, J. J., & Allemand, M. (2014). Moving personality development research forward: Applications using structural equation models. *European Journal of Personality, 28*, 300-310.
- Johns, G. (1981). Difference score measures of organizational behavior variables: A critique. *Organizational Behavior and Human Performance, 27*, 443–463.
- Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: a new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology, 103*, 54-69.
- Kashy, D. A., Donnellan, M. B., Burt, S. A., & McGue, M. (2008). Growth Curve Models for Indistinguishable Dyads using Multilevel Modeling and Structural Equation Modeling: The Case of Adolescent Twin's Conflict with their Mothers. *Developmental Psychology, 44*, 316-329.
- Kashy, D. A., & Donnellan, M. B. (2008). Comparing MLM and SEM Approaches to Analyzing Developmental Dyadic Data: Growth Curve Models of Hostility in Families. In N. A. Card, J.P. Selig, & T. D. Little (Eds.), *Modeling Dyadic and Interdependent Data in the Developmental and Behavioral Sciences*. New York, NY: Routledge.
- Kelley, K., & Lai, K. (2012). *MBESS* (Version 3.3.3). Retrieved from <http://www.cran.r-project.org/>
- Kenny, D. A. (1994). *Interpersonal perceptions: A social relations analysis*. New York: Guilford Press.
- Kenny, D. A. (1998). *SOREMO* (computer software and manual). University of Connecticut, CT.
- Kenny, D. A., Kashy, D. A., & Cook, W. L. (2006). *Dyadic data analysis*. New York: Guilford.

- Kenny, D. A., & Ledermann, T. (2010). Detecting, measuring, and testing dyadic patterns in the actor–partner interdependence model. *Journal of Family Psychology, 24*, 359–366.
- Kenny, D. A., & Livi, S. (2009). A componential analysis of leadership using the Social Relations Model. In F. J. Yammarino & F. Dansereau (Eds.), *Multilevel issues in organizational behavior and leadership* (pp. 147-191). Bingley, UK: Emerald
- Koskinen, J., & Snijders, T. A. B. (2013). Simulation, estimation, and goodness of fit. In D. Lusher, J., Koskinen, & G. Robins (Eds.), *Exponential Random Graph Models for Social Networks: Theories, Methods and Applications* (pp. 141-166). New York: Cambridge University Press.
- Knoke, D., & Yang, S. (2008). *Social Network Analysis* (2<sup>nd</sup> Edition). London: Sage Publications.
- Krivitsky, P. A., & Handcock, M. S. (2014). A separable model for dynamic networks. *Journal of the Royal Statistical Society B, 76*, 29-46.
- Küfner, A. C. P., Nestler, S., & Back M. D. (2013). The two pathways to being an (un-)popular narcissist. *Journal of Personality, 81*, 184-195.
- Laird, N. M., & Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics, 38*, 963-974.
- Ledermann, T., Macho, S., & Kenny, D. A. (2011). Assessing mediation in dyadic data using the actor-partner interdependence model. *Structural Equation Modeling, 18*, 595–612.
- Little, T. D. (2013). *Longitudinal structural equation modeling*. Guilford: New York.
- Loeys, T., Cook, W., De Smet, O., Wietzker, A., & Buysse, A. (2014). The actor–partner interdependence model for categorical dyadic data: A user-friendly guide to GEE. *Personal Relationships, 21*, 225–241.
- Lüdtke, O., Robitzsch, A., Kenny, D. A., & Trautwein, U. (2013). A general and flexible approach to estimating the social relations model using Bayesian methods. *Psychological Methods, 18*, 101-119.

- Lusher, D., Koskinen, J., & Robins, G. (2013). *Exponential Random Graph Models for Social Networks. Structural Analysis in the Social Sciences*. New York: Cambridge University Press.
- MacKinnon, D. P. (2008). *Introduction to statistical mediation analysis*. New York: Erlbaum.
- Maxwell, S. E., & Cole, D. A. (2007). Bias in cross-sectional analyses of longitudinal mediation. *Psychological Methods, 12*, 23–44.
- McArdle, J.J. (1988). Latent variable growth within behavior genetic models. *Behavior Genetics, 16*, 163–200.
- McArdle, J. J., (2001). A latent difference score approach to longitudinal dynamic structural analyses. In R. Cudeck, S. du Toit, & D. Sorbom (Eds.), *Structural equation modeling: Present and future* (pp. 342–380). Lincolnwood, IL: Scientific Software International
- McArdle, J.J., & Epstein, D.B. (1987). Latent growth curves within developmental structural equation models. *Child Development, 58*, 110–133.
- Meredith, W., & Tisak, J. (1984). “Tuckerizing” curves. Paper presented at the annual meeting of the Psychometric Society, Santa Barbara, CA.
- Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika, 55*, 107–122.
- Morgan, S. L., & Winship, C. (2015). *Counterfactuals and causal inference: Methods and principles for social research* (2<sup>nd</sup> Ed.). Cambridge: Cambridge University Press.
- Muthén, L. K., & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling, 9*, 599-620.
- Nestler, S. (2014a). How the 2SLS/IV estimator can handle equality constraints in structural equation models: A system-of-equations approach. *British Journal of Mathematical and Statistical Psychology, 67*, 353-369.

- Nestler, S. (2014b). *Restricted maximum likelihood estimation for parameters of the social relations model*. Unpublished manuscript, University of Münster.
- Nestler, S. (in press). Using instrumental variables to estimate the parameters in unconditional and conditional second-order latent growth models. *Structural Equation Modeling*.
- Olsen, J. A., & Kenny, D. A. (2006). Structural equation modeling with interchangeable dyads. *Psychological Methods, 11*, 127–141.
- Oud, J. H. L., & Jansen, R. A. (2000). Continuous time state space modeling of panel data by means of SEM. *Psychometrika, 65*, 199-215.
- Oud, J. H. L., & Delsing, M. J. M. H. (2010). Continuous time modeling of panel data by means of SEM. In K. van Montfort, J. Oud & A. Satorra (Eds.), *Longitudinal research with latent variables* (pp. 201-244). New York: Springer.
- Peter, J. P., Churchill, G. A., & Brown, T. J. (1993). Caution in the use of difference scores in consumer research. *Journal of Consumer Research, 19*, 655–662.
- Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods, 40*, 879-891.
- Preacher, K. J., & Kelley, K. (2011). Effect size measures for mediation models: Quantitative strategies for communicating indirect effects. *Psychological Methods, 16*, 93-115.
- R Core Team (2014). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from <http://www.R-project.org/>.
- Ripley, R., Boitmanis, K., & Snijders, T. A. B. (2013). RSiena: Siena - Simulation Investigation for Empirical Network Analysis. Retrieved from <http://cran.r-project.org/web/packages/RSA>
- Robins, G., & Lusher, D. (2013). A simplified account of ERGM as a statistical model for social networks. In D. Lusher, J. Koskinen, & G. Robins (Eds), *Exponential random*

- graph models for social networks: Theories, methods and applications* (pp. 9-15).  
New York: Cambridge University Press.
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48, 1–36.
- Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7, 147-177.
- Schönbrodt, F. D. (2014a). Testing congruence hypotheses with incommensurable scales. Manuscript submitted for publication.
- Schönbrodt, F. D. (2014b). Response Surface Analysis with the RSA package. Manuscript submitted for publication.
- Schönbrodt, F. D. (2015). *RSA: An R package for response surface analyses*. Retrieved from <http://cran.r-project.org/web/packages/RSA>
- Schönbrodt, F. D., Back, M. D., & Schmukle, S. C. (2012). TripleR: An R package for social relations analyses based on round-robin designs. *Behavior Research Methods*, 44, 455–470.
- Schönbrodt, F. D., Back, M. D., & Schmukle, S. C. (2014). *TripleR: A package for round robin analyses using R (version 1.4)*. Retrieved from <http://cran.r-project.org/web/packages/TripleR>
- Schönbrodt, F. D., Stas, L., & Loeys, T. (2014). *fSRM: Social Relations Analyses with roles*. Retrieved from <http://cran.r-project.org/web/packages/RSA/index.html>
- Selfhout, M. H. W., Burk, W., Branje, S., Denissen, J. J. A., Van Aken, M. A. G., & Meeus, W. (2010). Emerging late adolescent friendship networks and Big Five personality traits: A social network perspective. *Journal of Personality*, 78, 509-538.
- Selig, J. P., & Preacher, K. J. (2009). Mediation models for longitudinal data in developmental research. *Research in Human Development*, 6, 144-164.

- Singer, J., & Willett, J. (1993). It's about time: Using discrete-time survival analysis to study duration and the timing of events. *Journal of Educational Statistics, 18*, 155-195.
- Singer, J., & Willett, J. (2003). *Applied longitudinal data analysis*. Oxford: Oxford University Press.
- Shanock, L. R., Baran, B. E., Gentry, W. A., Pattison, S. C., & Heggstad, E. D. (2010). Polynomial regression with response surface analysis: A powerful approach for examining moderation and overcoming limitations of difference scores. *Journal of Business and Psychology, 25*, 543–554.
- Snijders, T. A. B. (2005). Models for Longitudinal Network Data. In P. Carrington, J. Scott, & S. Wasserman (Eds.), *Models and methods in social network analysis* (pp. 215-247). New York: Cambridge University Press.
- Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2<sup>nd</sup> Ed.). London: Sage Publishers.
- Snijders, T. A. B., & Kenny, D. A. (1999). The social relations model for family data: A multilevel approach. *Personal Relationships, 6*, 471–486.
- Snijders, T.A. B., Steglich, C. E. G., & van de Bunt, G. G. (2010). Introduction to actor-based models for network dynamics. *Social Networks, 32*, 44-60.
- Snijders, T. A. B., Steglich, C. E. G., & Schweinberger, M. (2007). Modeling the co-evolution of networks and behavior. In K. van Montfort, H. Oud, & A. Satorra (Eds.), *Longitudinal models in the behavioral and related sciences* (pp. 41-71). Lawrence Erlbaum.
- Stas, L., Schönbrodt, F., & Loeys, T. (in press). Getting the most out of family data with the R-package fSRM. *Journal of Family Psychology*.
- Stoel, R. D., & Galindo-Garre, F (2011) A comparison of multilevel regression analysis and structural equation modeling of growth curve models. In J. J. Hox, & J. K. Roberts

- (Eds.), *Handbook of advanced multilevel analysis* (pp. 99-111). Taylor & Francis Group.
- Steglich, C. E. G., Snijders, T. A. B., & West, P. (2006). Applying SIENA: An illustrative analysis of the coevolution of adolescents' friendship networks, taste in music, and alcohol consumption. *Methodology*, 2, 48-56.
- Steglich, C. E. G., Snijders, T. A. B., & Pearson, M. (2010). Dynamic networks and behavior: separating selection from influence. *Sociological Methodology*, 40, 329-393.
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using Multivariate Statistics* (6th ed). Boston : Pearson.
- Van Zalk, M. H. W., Branje, S. Kerr, M., & Stattin, H. (2010). It Takes Three: Selection, Influence, and De-selection Processes of Depression in Adolescent Peer Networks. *Developmental Psychology*, 46, 927-938.
- Van Zalk, M. H. W., Kerr, M., Branje, S., & Stattin, H. (2010). Peer Contagion and Adolescent Depression: The Role of Failure Anticipation. *Journal of Clinical Child and Adolescent Psychology*, 39, 837-848.
- Verbeke, G., & Molenberghs, G. (2009). *Linear mixed models for longitudinal data*. Berlin: Springer
- Voelkle, M. C., Oud, J. H. L., Davidov, E., & Schmidt, P. (2012). An SEM approach to continuous time modeling of panel data: Relating authoritarianism and anomia. *Psychological Methods*, 17, 176-192.
- Wang, P., Robins, G., Pattison, P., & Lazega, E. (2013). Exponential random graph models for multilevel networks. *Social Networks*, 35, 96-115.
- Wasserman, S., & Faust, K. (1994). *Social Network Analysis: Methods and Applications*. New York and Cambridge, ENG: Cambridge University Press.
- Wasserman, S., & Pattison, P. (1996). Logit models and logistic regression for social networks: I. An introduction to markov graphs and  $p^*$ . *Psychometrika*, 61, 401-425.

Wrzus, C., & Mehl, M. R. (2015). Lab and/or field? Measuring personality processes and their social consequences. *European Journal of Personality*, THIS ISSUE; INSERT CORRECT PAGE NUMBERS WHEN DETERMINED.

## Footnotes

<sup>1</sup> In order to be consistent with previous literature on RSA, we use the variable label  $z$  for the outcome here.

<sup>2</sup> When the members of one or more round-robin groups have clear roles, such as 'mother', 'father', or 'child' in the case of families, the *SRM with roles* has to be employed (Cook, 1994, 2005). There is also an R package to handle this variant of the SRM (*fSRM*; Schönbrodt, Stas, & Loeys, 2014; Stas, Schönbrodt, & Loeys, 2014).

<sup>3</sup> When one includes only the count of the number of edges in the model, then the ERGM reduces to a logistic regression where the intercept equals the density of the network (see Wasserman & Pattison, 1996).

Table 1

*Summary of the Example Individual Dataset*

Variable	Object	No. of Measurements	No. of data points	Method of Assessment
Personality	Agreeableness	3	3*100	Self- report on Likert scale
		1	100	Peer-report on Likert scale
Consequences	Number of friends	3	3*100	Self- report on Likert scale
Mediators	Number of interactions	3	3*100	Self- report on Likert scale

*Note.* The size of the sample is  $N = 100$ .

Table 2

*Excerpt of a Dataset for the Level of the Individual in Long Format (e.g., for Four Freshmen)*

Person	Time	Agreeableness		No. of friends	No. of interactions
		Self-report	Peer-report		
1	1	4	3	2	10
2	1	4	4	3	9
3	1	5	3	3	10
4	1	3	2	2	9
...					
1	2	4	-	1	9
2	2	5	-	2	10
...					

*Note.* The symbol ‘-’ denotes missing values (in the example, the peer report was assessed at the first time-point, only).

Table 3

*Features of Different Approaches to Analyze Longitudinal Data.*

Method	Type of change that is modeled	Is change between single time steps modeled?	Can one model nonlinear change?	Can one model between-person differences in change?	Are Latent variables possible?	Typical research questions	Required sample size and no. of measurement occasions	R package	Key reference
Autoregressive path and cross-lag models	Rank	Each	No	No	Yes	Is a variable stable between two time-points? If it is unstable: Who gains or loses more than predicted by the overall trend of the group? For cross-lag models: Are there time-lagged associations between the different variables.	$N \geq 200$ $o \geq 2$	lavaan	Biesanz (2012)
Continuous time modeling	Rank	n.a.*	No	Yes	Yes	What is the continuous time model behind the changes observed at the different time-points? What is the stability (cross-lagged effect) at time-point t?	$N \geq 200$ $o \geq 2$	under construction	Voelkle et al. (2012)
Latent difference score models	Mean	Each	No	Yes	Yes	Does a variable change in its absolute level? Can other variables predict which persons changes more or less?	$N \geq 200$ $o \geq 2$	lavaan	McArdle (2001)
Latent Growth Models	Mean	Average	Yes (if $o > 2$ )	Yes	Yes	Does a variable change in its absolute level? Can other variables predict which persons changes more or less?	$N \geq 200$ $o \geq 2$	lavaan	Bollen & Curran (2006)
Longitudinal Multilevel Models	Mean	Average	Yes	Yes	No	Does a variable change in its absolute level? Can other variables predict which persons changes more or less?	$N \geq 35$ $o \geq 2$	lme4	Singer & Willett (2003)
Survival time analysis	Mean	Average	No	Yes	Yes	Is there change in the occurrence of discrete outcome (e.g., events)? Does the occurrence depend on other variables?	$N \geq 50$ $o \geq 2$	the base function glm	Singer & Willett (1993)
Longitudinal network analysis (SIENA)	n.a.	n.a.*	Yes	Yes	No	Do person develop ties to other persons? Does this development depend upon the network structure and features of the persons?	$N \geq 20$ $o \geq 2$	RSiena	Snijders et al. (2010)

*Note.*  $N$  refers to the size of the sample and  $o$  to the number of measurement occasions.

\*The category is not applicable as a continuous time model is estimated.

Table 4

*Summary of the Example Dyad Dataset*

Variable	Object	No. of Measurements	No. of data points	Method of Assessment
Personality	Agreeableness	3	3*2*100	Self- report on Likert scale
Consequences	Relationship satisfaction	3	3*2*100	Self- report on Likert scale
Mediator	Amount of stress	2	2*2*100	Self- report on Likert scale

*Note.* The size of the sample is  $N = 100$  couples. For personality there are thus 3 (number of assessments) \* 2 (two members of the dyad) \* 100 (number of dyads) data points.

Table 5

*Excerpt of a Dataset for Dyads in Long Format (e.g., for Two Couples)*

Couple	Gender	Time	Agreeableness	Satisfaction	Stress
1	0	1	5	3	-
1	1	1	3	4	-
2	0	1	3	4	-
2	1	1	4	4	-
...					
1	0	2	5	4	3
1	1	2	4	3	3
2	0	2	3	4	2
2	1	2	4	5	2
...					

*Note.* For the gender variable, 0 indicates the husband and 1 the wife. The symbol ‘-’ denotes missing values (in the example, stress was not assessed at the first time-point).

Table 6

*Summary of the Example Group Dataset*

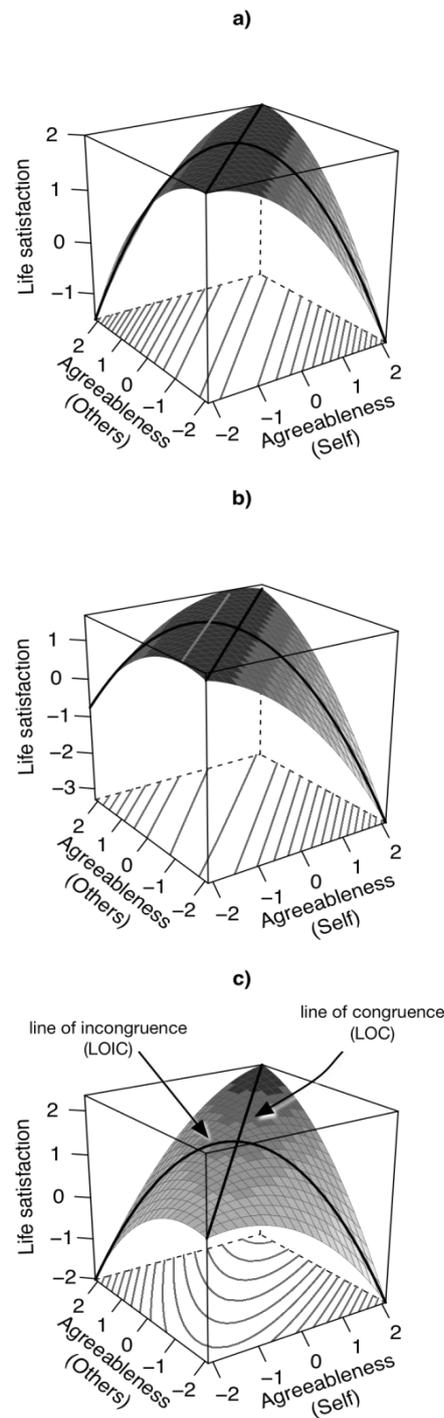
Variable	Object	No. of Measurements	No. of data points	Method of Assessment
Personality	Agreeableness	3	100	Self- report on a Likert scale
Social Consequences	Liking	3	3*100*99	Self- report on Likert scale
	Friendship	3	3*100*99	Self- report as a binary choice
Mediators	Number of Interaction	2	2*100*99	Self-report on binary scale

*Note.* The size of the sample is  $N = 250$ .

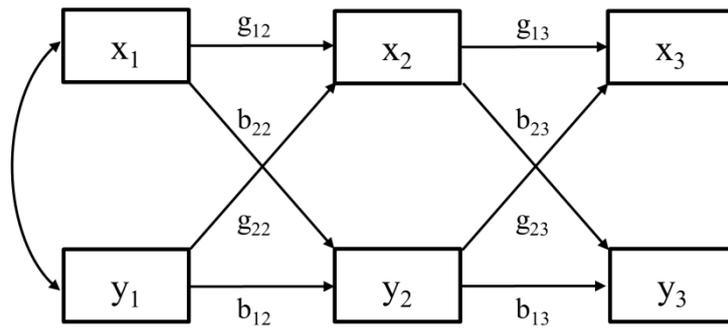
Table 7

*Excerpt of a Dataset for Groups in Long Format*

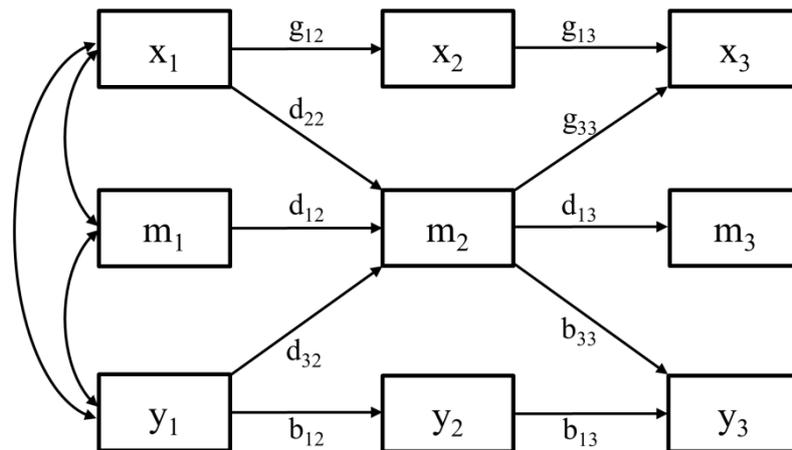
Case	Time	Group	Freshman as:		Liking	Friendship
			Perceiver	Target		
1	1	1	Anne (1)	Ben (2)	4	1
2	1	1	Anne (1)	Chris (3)	4	1
3	1	1	Anne (1)	Don (4)	5	1
4	1	1	Ben (2)	Anne (1)	1	0
5	1	1	Ben (2)	Chris (3)	2	0
6	1	1	Ben (2)	Don (4)	2	0
7	1	1	Chris (3)	Anne (1)	2	1
8	1	1	Chris (3)	Ben (2)	5	1
9	1	1	Chris (3)	Don (4)	2	0
10	1	1	Don (4)	Anne (1)	1	0
11	1	1	Don (4)	Ben (2)	1	0
12	1	1	Don (4)	Chris (3)	1	0
...						
9901	2	1	1	2	2	0
9902	2	1	1	3	4	1
...	...	...				



*Figure 1.* A response surface plot from a polynomial regression. Plot a) depicts a pure similarity hypothesis, where all predictor combinations along the line of congruence are optimal. Plot b) shows a shifted similarity pattern, where predictor combinations are optimal for which other-ratings are higher than self-ratings. Plot c) shows a rising ridge pattern, where similarity on a high predictor level is better than similarity on a low predictor level.

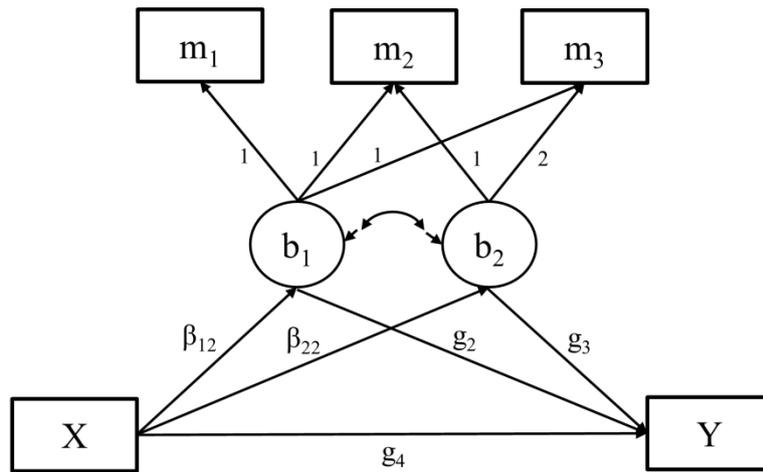


*Figure 2.* An autoregressive cross-lag panel model for two variables,  $x_t$  and  $y_t$ , measured at three time-points. Residual terms and the correlation between the residual terms of one variable across measurement occasions and between the two variables at one measurement occasion are not shown in the Figure.



*Figure 3.* An autoregressive cross-lag panel model including an input variable,  $x_t$ , a mediator,  $m_t$ , and the outcome,  $y_t$ , with each variable measured at three time-points. Residual terms and the correlation between the residual terms of the same variable across measurement occasions and of the three variables at one measurement occasion, respectively, are not shown.

Furthermore, some of the cross-lagged paths are omitted to enhance the comprehensibility of the Figure.



*Figure 4.* Mediation model in which the mediator,  $m_t$ , but neither the input variable,  $X$ , nor the output,  $Y$ , were assessed longitudinally. Residual terms and correlations between residual terms of the manifest variables are not shown.

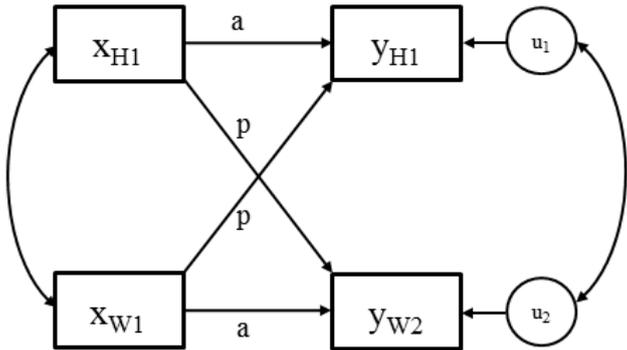
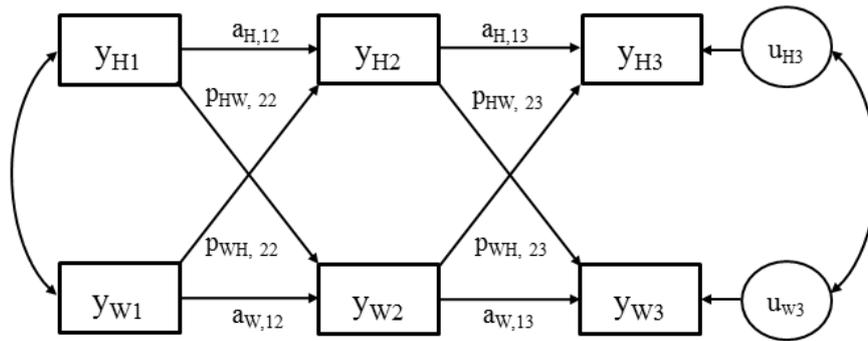
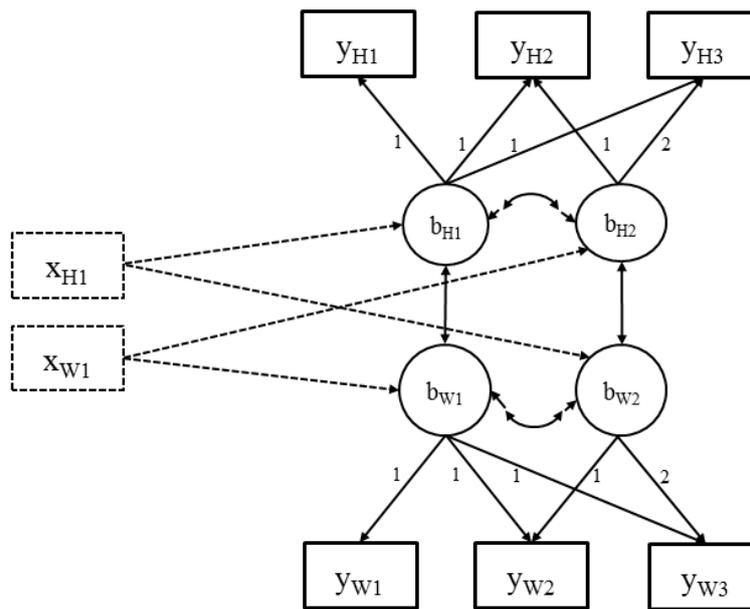


Figure 5. The Actor-Partner Interdependence Model for two variables. The subscript H denotes husbands and the subscript W denotes wives. *a* is the actor effect and *p* is the partner effect (see Kenny et al., 2006).



*Figure 6.* A longitudinal autoregressive model for a single dyadic variable. The subscript H denotes husbands and the subscript W denotes wives. The *as* are the actor effect and *ps* are the partner effect. Residuals for the second assessment are not shown in the Figure as are the correlations between the residuals across and within a time-point.



*Figure 7.* A dyadic growth model for a single dyadic variable with a dyadic personality predictor ( see Kashy & Donellan, 2008). The subscript H denotes husbands and the subscript W denotes wives. The figure does not contain the residual terms of the manifest variables and the correlations between them.

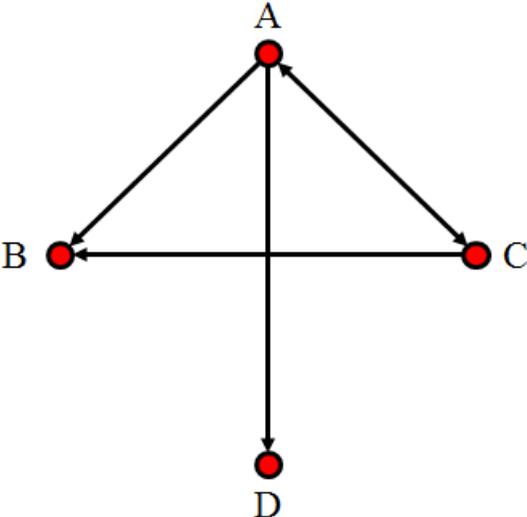


Figure 8. A social network representation of the excerpt of the freshmen data (see Table 7).