doi.org/10.26398/IJAS.0030-003

MULTILEVEL MODELS FOR PERSONAL NETWORKS: METHODS AND APPLICATIONS

Raffaele Vacca[1]

Department of Sociology and Criminology & Law, University of Florida, Gainesville, Florida, USA

***Abstract**. This article reviews the statistical formulation and substantive applications of multilevel models for personal network data. A personal network is a social network sampled around a focal individual, the ego. Network nodes are the ego and his or her social contacts, the alters; network edges are ties between ego and alters, and ties among alters as reported by ego, usually indicating acquaintance or various forms of interaction. Personal network datasets exhibit a classical multilevel structure, with alters or ego-alter ties (level 1) hierarchically nested within egos or ego-networks (level 2). Hierarchical linear and generalized linear models have been used in the social and the health sciences to analyze these data and explain the variation of outcomes observed on alters or ego-alter ties. The paper presents these models and the assumptions and hypotheses they imply; outlines their main research applications; and illustrates their use by analyzing real-world data on personal networks and social support among Sri Lankan immigrants in Milan, Italy.*

***Keywords**: Social networks, Social support, Ego-networks, Hierarchical Linear Models, Generalized Linear Mixed Models*

1. INTRODUCTION

Network analysis in the social sciences has historically comprised two distinct research traditions, one based on sociocentric or whole networks, and the other focusing on egocentric or personal networks (Marsden, 1990; Marin and Wellman, 2011). The sociocentric and egocentric approaches define and sample social networks in fundamentally different ways. In the sociocentric view, a social network is a set of actors and ties that exist within a given social boundary, such as that of an organization, a school, a village, or any meaningfully defined community or social group (Wasserman and Faust, 1994). Different criteria may be adopted to delimit the boundaries of the group, including formal membership in organizations, participation in specific events, or social relationships with "seed" individuals in snowball or link-tracing designs (Marsden, 1990; Wasserman and Faust, 1994). Primary sociocentric data are typically collected from the group members themselves;

[1] Corresponding author: Raffaele Vacca, email: r.vacca@ufl.edu

thus, if the data are obtained through standard surveys, the survey is administered to each network actor.

In the egocentric view, a social network is also a set of actors and ties, but the network boundary is based on the existence of a relationship with a central individual: the network actors include a focal person ("ego") and that person's direct contacts ("alters"), according to a given definition of social relationship or interaction; and the ties are those that the ego reports with the alters and among different alters (Marsden, 1990; Wellman, 2007; Crossley et al., 2015; Perry et al., 2018; McCarty et al., Forthcoming). In other words, an ego-network, or personal network, is a social network sampled around an ego, on which all data are collected from the ego. In the collection of primary ego-network data, the egos are the survey respondents, while the alters are not directly observed.

The typical personal network dataset includes several egos. These represent a sample of individuals extracted from a population of interest, such as high school students, high-tech company employees, or immigrants from a certain ethnic group. In quantitative studies based on primary data, standard sampling designs from survey research are usually adopted to recruit the egos. The resulting datasets include dozens, hundreds or thousands of egocentric networks (Figure 1). Depending on the goals of the study, the egos may be included or excluded from their respective ego-networks before conducting different types of analysis (McCarty and Wutich, 2005).

Personal network data comprise variables about egos, alters, ego-alter ties, alter-alter ties, and entire ego-networks. These datasets exhibit a clear multilevel structure, with the lower-level units of alters or ego-alter ties clustered within the higher-level units of egos or ego-networks. Such data structures can be accommodated by multilevel models, a widely popular class of statistical models in the social sciences (Gelman and Hill, 2006; Goldstein, 2010; Snijders and Bosker, 2012). The analysis of personal network data with multilevel models, first proposed in the 1990s (Snijders et al., 1995; van Duijn et al., 1999), has generated some of the most diverse and insightful applied social network research in the last twenty years. Most of this work has adopted *hierarchical* multilevel models for nested data structures, in which each lower-level unit is associated to one and only one higher-level unit (Gelman and Hill, 2006, p. 2; Goldstein, 2010, p. 1).[2]

[2] *Non-hierarchical* multilevel models, which can accommodate multilevel data that do not display a strictly hierarchical nesting of units, include cross-classified, multiple-membership, and Multiple Membership Multiple Classification (MMMC) models (Rasbash and Browne 2008). These models have not been used in personal network research as extensively, although more recent research has started to explore their application to social network data (van Duijn 2013; Tranmer et al. 2014; Mollenhorst et al. 2016).

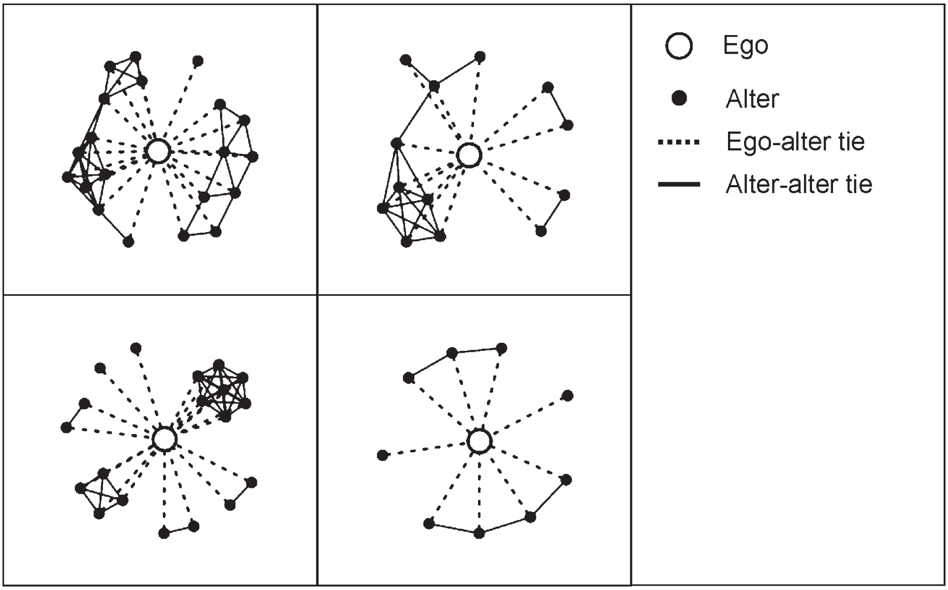


Fig. 1: Four personal networks.

This article introduces the formulation of hierarchical linear and logistic models for personal networks (Section 3), after describing personal network surveys and data in more detail (Section 2). We focus on hierarchical models because this is the class of multilevel models used in the vast majority of existing personal network studies; however, we also briefly mention examples of ego-network data that are more appropriately described by non-hierarchical multilevel models. Section 4 of the article outlines the main applications of hierarchical models for ego-networks in the social and health sciences, with a particular emphasis on the cross-disciplinary study of social support. Finally, we demonstrate the specification, estimation and interpretation of these models with an analysis of data on personal networks and social support collected in 2012 among Sri Lankan immigrants in Milan, Italy (Section 5).

2. PERSONAL NETWORK DATA

Most current literature in social network analysis uses the expressions "ego-network" and "personal network" as synonyms (Marsden, 1990; Wellman, 2007; Crossley et al., 2015). While the terms "egocentric" and "ego-network" emphasize the data collection method, which samples the network data around one ego, the term "personal network" highlights the substantive meaning of these data, which capture some aspect of the social world or "personal community" surrounding the focal individual (Wellman, 2007; Chua et al., 2011). The expression "personal network" is sometimes preferred to indicate, specifically, an *unbounded* ego-network that includes ego's contacts from any kind of relationship and social setting (McCarty et al., Forthcoming). Since the methods discussed in this article are applicable to any type of egocentrically sampled network, we use the expressions "ego-network" and "personal network" interchangeably.

Research questions that motivate personal network studies are typically concerned with one of two types of dependent variable: (*i*) Outcomes observed on the ego (e.g., depression, smoking behavior or socioeconomic attainment), which are explained as a function of both individual attributes of the ego and characteristics of the social world around the ego (e.g., cohesion, diversity or family orientation of the ego's personal network); or (*ii*) Outcomes observed on the relationship between an ego and an alter (e.g., provision of social support, frequency of contact, or level of trust between ego and alter), which are predicted on the basis of characteristics of the ego-alter relationship itself, the ego, the alters, or the broader personal network. The statistical methods presented in this paper address the second type of research questions, in which the dependent variable is an outcome observed on ego-alter ties, and characteristics of ties, individuals (egos and alters), and personal communities are treated as explanatory variables.

2.1 SURVEYS FOR PERSONAL NETWORKS

Primary personal network data are collected with surveys that usually include four main components (Marsden, 1990; Crossley et al., 2015; Perry et al., 2018; McCarty et al., Forthcoming):

i. Questions about non-network attributes of ego.
ii. Question(s) to elicit a list of alters (the "name generators").
iii. Questions about each alter and ego-alter relationship (the "name interpreters").
iv. Questions about the relationships among the alters.

The first component collects individual information about the ego (i.e., the respondent), such as standard sociodemographic data (sex, age, educational level, etc.), information about an ego-level outcome of interest (e.g., depression, smoking behavior, socioeconomic attainment), and other data used to generate relevant explanatory variables (e.g., information about ego's medical history or socioeconomic background).

The second component consists of one or multiple "name generator" questions (Campbell and Lee, 1991), which ask the respondent to list a number of social contacts that meet certain criteria. These are the inclusion or exclusion criteria that researchers have set for alters to be part of the ego-network, which effectively delimit the network boundary. For example, personal network surveys may ask respondents to name the alters "to whom they feel closest" (Wellman, 1979), or with whom they "discuss important matters" (Burt, 1984); the alters who provide ego with specific types of social support (e.g., Fischer, 1982; Marin and Hampton, 2007; Herz, 2015); or, more comprehensively, any contact that the respondent knows and has interacted with in the past two years (Lubbers et al., 2007). The number of elicited alters differs across studies, and it can be variable within the sample (e.g., Fischer, 1982; Perry and Pescosolido, 2010) or fixed (e.g., Lubbers et al., 2010; Kennedy et al., 2012); from as few as 5-10 alters (Wellman, 1979; Burt, 1984) to as many as 40-60 alters (McCarty, 2002; Vacca et al., 2017).[3]

The third survey component includes two different types of questions: those about individual attributes of alters and those about characteristics of ego-alter ties. Questions about attributes of alters may ask, for example, about sociodemographic characteristics or specific behaviors (e.g., religious practices, smoking behavior, leisure activities) of each contact of the ego. Typical questions about characteristics of the ego-alter ties are those asking how long ego has known alter; how frequently ego and alter see each other; how close ego feels to alter; how much ego trusts alter; and whether ego obtains from alter certain types of support. This component of the survey yields data about what is commonly referred to as the *composition* of the ego-network, that is, the distribution of attributes of the ego-network nodes.

Finally, the fourth survey component is concerned with the ties among the alters. Ego is asked to evaluate the existence of certain types of ties, such as acquaintance or a specific definition of interaction, in all (or, sometimes, a fraction of) the pairs of alters. Typical questions are whether each alter talks to each other in the absence of ego, or knows each other by name. These questions provide data about the *structure* of the ego-network, that is, the distribution of ties among the ego-network nodes. Thus, typical personal network data include two types of ties: ties between egos and alters, and ties between alters.[4] Unless specified otherwise, the word "ties" in this paper refers to the ties between egos and alters.

---

[3] In addition to name generators, alternative but less common survey instruments have been proposed for research that focuses on the social capital embedded in personal networks (Portes 1998). These are known as the position generator (Lin and Dumin 1986), which asks respondents to indicate if they know people in specific occupational positions; and the resource generator (Van Der Gaag and Snijders 2005), which asks respondents if they know people who are able to provide specific resources or types of support. These instruments normally do not require respondents to name specific social contacts.

[4] While alter-alter ties are usually an integral part of personal network studies in the social sciences (Marsden 1990; Crossley et al. 2015; Perry et al. 2018; McCarty et al. Forthcoming), some ego-network research designs only collect ego-alter ties. This has been called the "minimal" egocentric network design (Krivitsky and Morris 2017), and can be adopted to reduce respondent burden and data collection costs, as well as for ethical considerations, in order to limit the risk of obtaining potentially identifying information about the alters (who are not research participants).

2.2 THE MULTILEVEL STRUCTURE OF PERSONAL NETWORK DATA

Personal network data include information observed at least on four different levels: egos; alters; ego-alter ties; alter-alter ties. This creates a multilevel structure in which alters or ego-alter ties (the lower level or "level 1") are clustered within egos (the higher level or "level 2"). Clustered data of this type are frequently encountered in the social sciences. Common examples include data on students sampled from a set of schools, patients treated by doctors, residents grouped within cities, and measurements from different time points clustered within respondents in repeated measures designs. These are all cases of two-level hierarchical data structure, with level-1 units (students, patients, residents, measurements, or ties) nested in level-2 units, also called "groups" or "clusters" (schools, doctors, cities, respondents, or egos). Figure 2 compares the hierarchical structure of typical school data and personal network data using unit diagrams and classification diagrams (Goldstein, 2010). The structure of the ego-network data is strictly hierarchical if the personal networks do not overlap, that is, no alter belongs to multiple ego-networks, and no ego appears as an alter in another ego-network (Snijders et al., 1995; van Duijn et al., 1999). The rest of this paper focuses on hierarchical data and models, although Section 3.4 mentions an extension to non-hierarchical models for overlapping personal networks.

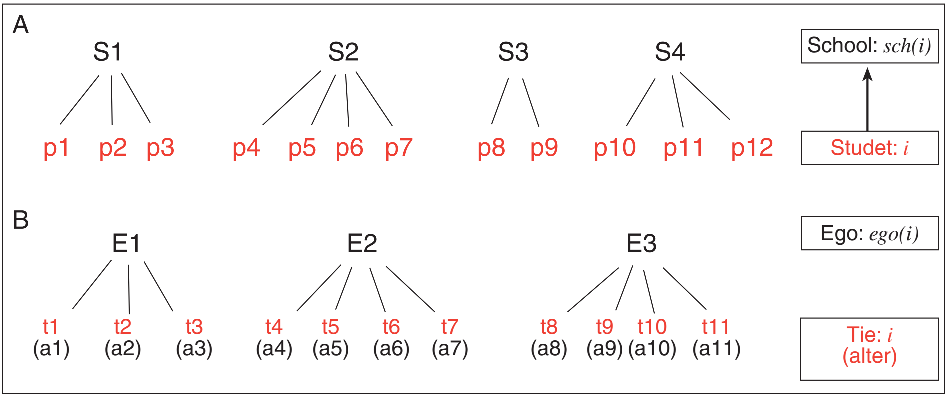


Fig. 2: Unit diagrams and classification diagrams for hierarchical school data (A) and personal network data (B).

Alter-alter ties might also be viewed as lower-level units clustered within ego-networks (e.g., Louch, 2000; Mollenhorst et al., 2011). This approach, however, is not frequent and might be problematic because, unlike ego-alter ties, alter-alter ties in the same ego-network are likely to follow additional, more complex patterns of interdependence (Mollenhorst et al., 2016). More common in multilevel studies of personal networks is to use alter-alter ties to create structural variables at the ego-network level (e.g., personal network density) or at the alter level (e.g., alter's degree centrality) to be included as predictors in models for ego-alter tie variation (e.g., Wellman and Frank, 2001; Lubbers et al., 2010; Martí et al., 2017).

Variables in personal network datasets may be measured on ego-alter ties, alters, egos, or networks (Table 1). While conceptually different, the level of alters and the level of ego-alter ties are technically the same in typical ego-network datasets. This is the case because, if the personal networks do not overlap, a one-to-one correspondence exists between alters and ties: since each alter belongs to one ego-network only (i.e., is only connected to one ego), there is one alter for each ego-alter tie, and one ego-alter tie for each alter. As a result, any variable that varies across alters also varies across ties, and vice versa. Similarly, because each network corresponds to one and only one ego, ego-level variables (e.g., ego's age) and network-level variables (e.g., ego-network density) technically pertain to the same level of analysis.

Much personal network research aggregates the original, level-1 tie data into summary level-2 measures observed on egos, and conducts all analyses on the resulting single-level dataset. This research typically aims to examine the association between summary personal network characteristics and outcomes observed for each respondent. Ego-level measures are calculated to describe the characteristics of each ego's personal community, including its composition (e.g., proportion of family members in the network), its structure (e.g., the density of the network), or a combination of the two (e.g., the density of ties among family members in the network).

An increasingly large proportion of personal network research, however, avoids aggregating all data into level-2 ego summary variables, and maintains the analysis at the finer level of alters and ties. There are at least two good reasons to conduct tie-level disaggregate analyses. First, like all data aggregation, summarization of tie data into ego-level data implies some loss of information, with ego-level averages and proportions potentially hiding patterns of variation across alters or ties within the same egos. Second, similar to studies of summary measures on cities, counties or other spatial aggregates, ego-level aggregate data expose analysts to the risk of ecological fallacy (Robinson, 1950). In personal network research, the ecological fallacy occurs when observations about egos (e.g., "respondents with older social contacts on average tend to have more alters who provide financial support") are used to draw conclusions about ties or alters (e.g., "older social contacts are more likely to provide financial support").

Tab. 1: Example of tie-level personal network dataset.

AlterID	EgoID	Alter Sex	Alter Age	Alter Support	Ego Sex	Ego Age	Proportion Female Alters	Avg Alter Age	Network Density
101	1	M	23	1	F	32	.6	36.4	.6
102	1	F	34	1	F	32	.6	36.4	.6
103	1	F	56	0	F	32	.6	36.4	.6
104	1	F	43	1	F	32	.6	36.4	.6
105	1	M	26	0	F	32	.6	36.4	.6
201	2	M	27	0	M	39	.2	40.8	1.0
202	2	M	55	0	M	39	.2	40.8	1.0
203	2	M	39	0	M	39	.2	40.8	1.0
204	2	F	38	1	M	39	.2	40.8	1.0
205	2	M	45	0	M	39	.2	40.8	1.0
301	3	F	23	1	M	41	.8	43.4	.4
302	3	F	38	1	M	41	.8	43.4	.4
303	3	M	45	0	M	41	.8	43.4	.4
304	3	F	53	1	M	41	.8	43.4	.4
305	3	F	58	1	M	41	.8	43.4	.4

3. HIERARCHICAL MODELS FOR PERSONAL NETWORKS

Hierarchical models allow personal network analysts to examine the variation of a dependent variable observed on ego-alter ties (e.g., frequency of contact or provision of support from alter to ego) as a function of characteristics of the ties themselves, the respondents who nominate the ties (egos), the nominees (alters), and the broader social contexts in which the ties are embedded (ego-networks). The level-1 unit of analysis is the ego-alter tie, and the dependent variable is a tie value. More rarely, alters are the level-1 units, and the dependent variable is a characteristic of the alter (e.g., de Miguel Luken and Tranmer, 2010).

Hierarchical data structures contradict the fundamental assumption, made in standard statistical models, that observations of the dependent variable are identically and independently distributed. The clustering of units in higher-level groups entails dependence (therefore correlation) between values of the dependent variable observed on cases that belong to the same cluster. Alters or ties that are associated with the same ego tend to be correlated, that is, more similar to each other compared with alters or ties attached to different egos, for example in terms of smoking behavior, frequency of contact, or level of support provided. Such similarity may occur for a host of reasons, which might be unobserved, unobservable or unknown, and therefore cannot be incorporated through explanatory variables in a statistical model. Some of these reasons might be, for example, the tendency to homophily in social networks, whereby egos tend to associate with similar alters (McPherson et al., 2001); unobserved individual conditions that cause certain egos to maintain more frequent contact overall with their alters compared with other egos; psychological characteristics that lead certain egos to perceive more support from their contacts and others to perceive less; and so forth.

Hierarchical models account for the dependence among ties from the same ego by positing that one or more model coefficients are random variables that take a different value for each ego. In the following, we exemplify these models by delineating a simple hierarchical version of linear and logistic models for personal network data. Using the classification notation (Browne et al., 2001; Goldstein, 2010, p. 249), we index level-1 units (ties) as i, and level-2 units (egos) as $ego(i)$. The term *ego(i)* is a *classification function* that returns the ego corresponding to tie i. In the following models, x_i indicates an explanatory variable measured at the level of alters or ego-alter ties (level 1); and $w_{ego(i)}$ indicates an explanatory variable observed at the level of egos or ego-networks (level 2). As usual, y_i indicates the dependent variable, which is always a level-1 characteristic of ties or alters.

3.1. HIERARCHICAL LINEAR MODELS

In a hierarchical linear model for personal networks, y_i might represent, for example, a continuous measure for the strength of the tie between ego and alter, which summarizes different indexes for emotional closeness and trust. The model might posit that the ego-alter tie strength depends on whether alter is a member of ego's family (x_i), which is an explanatory variable observed for the ties ($x_i = 1$ if alter is ego's family, $x_i = 0$ otherwise); and on ego's age ($w_{ego(i)}$). For example, we might hypothesize that ties with family members are systematically stronger; and that older egos, in general, tend to report stronger ties with their (family or non-family) contacts. Continuous explanatory variables in hierarchical models are usually centered around their overall mean to aid interpretation of the random coefficients and their variances. Thus, in the following examples, $w_{ego(i)}$ refers to the *centered* age of ego. While, for simplicity, the models in this section only include one x_i and one $w_{ego(i)}$

predictor, they can be extended in the usual way to multiple alter-level ($x_{1i}, x_{2i}, \ldots, x_{ki}$) and ego-level ($w_{1,ego(i)}$, $w_{2,ego(i)}$, …, $w_{q,ego(i)}$) explanatory variables.

A simple *random-intercept* hierarchical linear model for tie strength can be defined as follows:

$$y_i = \beta_{0,ego(i)} + \beta_1 x_i + e_i \tag{1}$$

$$\beta_{0,ego(i)} = \gamma_{00} + \gamma_{01} w_{ego(i)} + u_{0,ego(i)} \tag{2}$$

$$e_i \sim N(0, \sigma_e^2)$$

$$u_{0,ego(i)} \sim N(0, \sigma_{u0}^2)$$

Equivalently, Equations (1)-(2) can be rewritten as a single equation:

$$y_i = \gamma_{00} + \gamma_{01} w_{ego(i)} + u_{0,ego(i)} + \beta_1 x_i + e_i$$

Parameters in this model are similar to their counterparts in standard single-level linear models, except that the intercept $\beta_{0,ego(i)}$ is not viewed as a fixed parameter, but as a random variable that takes a different value for each *ego(i)*. This value depends on specific observed characteristics of ego ($w_{ego(i)}$), as well as on a *random effect* associated with each ego ($u_{0,ego(i)}$). The random effect can be regarded as a common deviation from the average tie strength in the population, which characterizes all the ties attached to the same *ego(i)*. Among the non-random parameters, γ_{00} captures the average value of the intercept across all egos; γ_{01} measures the association between the ego characteristic $w_{ego(i)}$ (here, ego's age) and the overall level of tie strength for that ego ($\beta_{0,ego(i)}$); and β_1 indicates the (fixed) association between the alter or tie characteristic x_i (here, whether the tie is family) and tie strength, similar to a standard linear model. Thus, the strength of a tie *i* (y_i) is posited to depend on certain characteristics of the tie itself (e.g., whether the tie is family); on specific characteristics (e.g., age) of the ego to which the tie is attached (*ego(i)*); and on a random deviation that is unique to the specific *ego(i)*.

Since all the ties *i* associated with the same *ego(i)* share the same random deviation $u_{0,ego(i)}$, the model accounts for the correlation between the ties nested in the same ego. In other words, $u_{0,ego(i)}$ captures any unobserved characteristic of *ego(i)* that (after accounting for observed characteristics of egos and ties: $w_{ego(i)}$ and x_i) might cause *ego(i)*'s ties to show more similar (i.e., correlated) strength values. Thus, the model accounts for the outcome y_i being potentially affected by both observed characteristics of the egos ($w_{ego(i)}$), and a set of unknown, unobservable or unobserved traits of the egos, incorporated by $u_{0,ego(i)}$. Since they include a mix of fixed (i.e., non-random) coefficients (here, β_1) and random coefficients (here, $\beta_{0,ego(i)}$), hierarchical linear models are also known as random-effects or mixed-effects models (Searle et al. 1992), although this terminology is not used consistently throughout the statistical literature (Gelman and Hill, 2006, p. 245).

The ego random effect $u_{0,ego(i)}$ is assumed to follow a normal distribution with zero mean and a constant variance, σ_{u0}^2, which is one of the model parameters to be estimated; hence, the random intercept also follows a normal distribution: $\beta_{0,ego(i)} \sim N(\gamma_{00} + \gamma_{01}w_{ego(i)}, \sigma_{u0}^2)$. The variance σ_{u0}^2 quantifies the between-ego variability, i.e., the amount of variation in y_i that is due to systematic differences between the egos, beyond the outcome variation that is explained by the predictors included in the model (here, ego's age and whether the tie is family). By contrast, e_i represents the random deviation of each tie i from the average strength of the ties associated with the same *ego(i)*. This level-1 residual error is also assumed to follow a normal distribution with zero mean and a constant variance, σ_e^2, which is estimated as a model parameter. Thus, the variance σ_e^2 quantifies the amount of unexplained outcome variation that is attributable to differences between ties (within the same ego), rather than between the egos.

A standard assumption in multilevel models is that random effects from different levels are independent (Snijders and Bosker, 2012, p. 49). In our example, this implies $cov(e_i, u_{0,ego(i)}) = 0$. This assumption signifies that the unexplained deviation of a level-1 unit (a tie) from its group mean is not systematically correlated with the unexplained deviation of that unit's group (an ego) from the overall population mean. In other words, the level 1 of ties and the level 2 of egos are two independent sources of outcome variation; and the between-tie (level 1) and between-ego (level 2) variabilities can be separately quantified by two model parameters. Indeed, the assumption of cross-level residual independence implies that, in random-intercept models, the total outcome variation is simply the sum of between-ego variance and between-tie variance:

$$var(y_i) = \sigma_{u0}^2 + \sigma_e^2 \qquad (3)$$

The two variance parameters allow analysts to neatly partition the outcome variation between two sources: the level 2, accounting for systematic differences between respondents (e.g., certain respondents feeling regularly stronger ties to their alters compared with others); and the level 1, representing differences between ties even when they are attached to the same respondent.

Given Equation (3), a simple measure for the amount of variation due to level 2 (differences between respondents) is the Variance Partition Coefficient (VPC, Goldstein, 2010, p. 19):

$$VPC = \frac{\sigma_{u0}^2}{\sigma_{u0}^2 + \sigma_e^2}$$

The VPC quantifies both the proportion of tie variation that is due to the egos; and, in random-intercept models, the correlation between ties that are nominated by the same ego, that is, the Intra-Class Correlation between ties (ICC, Goldstein, 2010, p. 19; Snijders and Bosker, 2012, p. 17). Snijders and Bosker (2012, p. 109) discuss further ways to decompose the outcome variance and measure the variance explained by predictors in hierarchical linear models.

We might hypothesize that different egos not only tend to perceive different levels of tie strength to their alters overall; but they also register a different effect of *family ties*, as opposed to non-family relationships, on strength. In other words, we might think that each ego is characterized by a specific value not only of the overall level of tie strength across all alters, as represented by the intercept $\beta_{0,ego(i)}$; but also of the association between family ties (x_i) and tie strength, as represented by the slope β_1. This leads to a *random-slope* model:

$$y_i = \beta_{0,ego(i)} + \beta_{1,ego(i)} x_i + e_i \tag{4}$$

$$\beta_{0,ego(i)} = \gamma_{00} + \gamma_{01} w_{ego(i)} + u_{0,ego(i)} \tag{5}$$

$$\beta_{1,ego(i)} = \gamma_{10} + \gamma_{11} w_{ego(i)} + u_{1,ego(i)} \tag{6}$$

$$e_i \sim N(0, \sigma_e^2) \tag{7}$$

$$\begin{pmatrix} u_{0,ego(i)} \\ u_{1,ego(i)} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{u0}^2 & \\ \sigma_{u01} & \sigma_{u1}^2 \end{pmatrix}\right), \tag{8}$$

where Equations (4)-(6) can be rewritten as follows:

$$y_i = \gamma_{00} + \gamma_{01} w_{ego(i)} + u_{0,ego(i)} + \gamma_{10} x_i + \gamma_{11} w_{ego(i)} x_i + u_{1,ego(i)} x_i + e_i \,.$$

In this model, each *ego(i)* is associated with a specific effect $u_{0,ego(i)}$ on the intercept, and a different effect, $u_{1,ego(i)}$, on the slope of x_i. The ego random effects on the intercept and the slope are normally distributed with zero means and constant variances (σ_{u0}^2 and σ_{u1}^2, respectively), which are parameters to be estimated. Thus, for the random slope, the model assumes $\beta_{1,ego(i)} \sim N(\gamma_{10} + \gamma_{11} w_{ego(i)}, \sigma_{u1}^2)$. Among the non-random parameters added to the model in Equations (1)-(2), γ_{10} represents the average association between x_i (family relationship) and the outcome variable, across all egos; and γ_{11} quantifies the association between $w_{ego(i)}$ (ego's age) and the random slope $\beta_{1,ego(i)}$ (the strength of x_i's effect on the outcome).

In random-slope models, a non-null correlation is hypothesized between ego random effects on the intercept and on the slope (Snijders and Bosker, 2012, p. 76). For example, we might hypothesize that individuals who tend to perceive overall stronger ties with all their alters (higher $u_{0,ego(i)}$) are also people who tend to regard family and non-family ties as more similar in terms of tie strength, which determines a lower association between strength and the "family" predictor x_i (lower $u_{1,ego(i)}$). The model incorporates this possibility by reserving a specific parameter for $\sigma_{u01} \equiv cov(u_{0,ego(i)}, u_{1,ego(i)})$, whose value is estimated from the data.

It should be noted that while, for simplicity, the model in Equations (4)-(8) assumes that the ego-level random intercept and random slope depend on the same ego characteristic ($w_{ego(i)}$), this need not be the case, and different ego-level predictors (say, $w_{1,ego(i)}$ and $w_{2,ego(i)}$) may appear in Equation (5) for the random intercept and in Equation (6) for the random slope.

3.2. HIERARCHICAL LOGISTIC MODELS

Binary variables are extremely common in the social sciences, and the study of personal networks is no exception. Binary outcomes in multilevel data structures, including personal network data, can be analyzed using hierarchical logistic models. Logistic models can be derived as generalized linear models, and hierarchical logistic models are sometimes presented as an instance of "generalized linear mixed models" (GLMM) in the statistical literature (e.g., McCulloch and Searle, 2001).

In personal network studies of social support, y_i usually indicates whether an alter provides some form of support or aid to ego ($y_i = 1$) or not ($y_i = 0$). This could be financial aid, help with family problems, or emotional support. In the following example, the binary outcome y_i is 1 if alter provides a generic form of support to ego, and 0 if alter does not provide support. A hierarchical logistic model assumes $y_i \sim Bernoulli(\pi_i)$, and expresses π_i (here, the probability of an ego-alter tie providing support) as a logistic function of the linear predictor. Equivalently, the model expresses the expected log-odds of support provision ($\log\left[\frac{\pi_i}{1-\pi_i}\right]$) as equal to the linear predictor. The linear predictor may include a tie-level explanatory variable (x_i), as well as a random intercept and slope depending on an ego-level explanatory variable ($w_{ego(i)}$):

$$\pi_i = \frac{\exp\left(\beta_{0,ego(i)} + \beta_{1,ego(i)} x_i\right)}{1 + \exp\left(\beta_{0,ego(i)} + \beta_{1,ego(i)} x_i\right)} \tag{9}$$

$$\beta_{0,ego(i)} = \gamma_{00} + \gamma_{01} w_{ego(i)} + u_{0,ego(i)} \tag{10}$$

$$\beta_{1,ego(i)} = \gamma_{10} + \gamma_{11} w_{ego(i)} + u_{1,ego(i)} \tag{11}$$

$$\begin{pmatrix} u_{0,ego(i)} \\ u_{1,ego(i)} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{u0}^2 & \\ \sigma_{u01} & \sigma_{u1}^2 \end{pmatrix}\right) \tag{12}$$

Like in the linear examples, x_i is a tie-level binary variable that flags family ties, and $w_{ego(i)}$ is ego's age. If y_i indicates the provision of financial aid, this model may represent, for instance, the hypotheses that individuals are more likely to receive financial support from family members, compared with non-related contacts ($\gamma_{10} > 0$), and that older individuals are more likely to receive financial support from any contact, whether related or not ($\gamma_{01} > 0$).

The model in Equations (9)-(12) describes certain characteristics of egos, such as age ($w_{ego(i)}$), as systematically affecting both the overall probability of a tie providing support (through $\beta_{0,ego(i)}$), and the impact of x_i on this probability (through $\beta_{1,ego(i)}$). Thus, the model can represent a situation in which certain egos (for example, older respondents) (*i*) tend to obtain more support overall ($\gamma_{01} > 0$, determining higher $\beta_{0,ego(i)}$ with older age); and (*ii*) tend to register a stronger effect of particular tie characteristics, such as being a family tie, on the likelihood of support ($\gamma_{11} > 0$, implying higher $\beta_{1,ego(i)}$ with older age). Like the hierarchical linear models exemplified above, hierarchical logistic models might posit that *different* ego characteristics ($w_{1,ego(i)}$, $w_{2,ego(i)}$, …, $w_{q,ego(i)}$) affect the random intercept and the random slope. Obviously, also these models may include only a random intercept, and no random slope.

Beyond the variation in support likelihood that is explained by characteristics of ties and egos, hierarchical logistic models incorporate the hypothesis that each *ego(i)* is associated with two random effects on all her ties to alters:

i. A random effect on the intercept ($u_{0,ego(i)}$), which affects the overall probability of support for all ties nominated by *ego(i)*, in the same direction.
ii. A random effect on the slope ($u_{1,ego(i)}$), which affects the association between "being family" and the probability of support, for all ties nominated by *ego(i)*, also in the same direction.

Importantly, hierarchical logistic models (like their single-level version) do not include level-1 residual errors (e_i), because they predict the *probability* of a certain outcome value, rather than the outcome value itself. In addition, the variance of the Bernoulli distributed binary outcome depends on its expected value ($var(y_i) = \pi_i(1 - \pi_i)$), changing the meaning and interpretation of the level-1 residual variance (Snijders and Bosker, 2012, p. 291).

Level-1 residual errors do appear in the linear latent ("threshold") model from which the logistic model can be derived. In this derivation, a continuous latent variable y_i^*, with arbitrary scale, is assumed to determine the observed binary outcome ($y_i = 1$ when $y_i^* > 0$), and its residual errors are assumed to follow a standard logistic distribution with variance fixed to $\sigma_e^2 = \pi^2/3$ (Long, 1997, p. 42; Snijders and Bosker, 2012, p. 303). This *fixed* level-1 residual variance, σ_e^2, does not obviously have the same interpretation as the *estimated* level-1 variance in hierarchical linear models: in logistic models, $\pi^2/3$ is simply an arbitrary fixed scale that is attributed to y_i^* (an unobservable construct) to guarantee model identification, with no substantive meaning. However, in *hierarchical* logistic models, the fixed level-1 variance does have the potentially confusing effect of causing parameter estimates to artificially increase (in absolute value) when fixed or random effects are added (Snijders and Bosker, 2012, p. 307). Intuitively, this happens because the ratios between the fixed parameter estimates and the unexplained variation of y_i^* are meaningful and constrained to remain approximately constant in the estimated models. At the same time, because the (arbitrary) total variance of y_i^* equals the fixed level-1 variance ($\pi^2/3$) plus any level-2 variance, the addition of level-2 random effects increases y_i^*'s total variance. As a result, the fixed effect estimates also tend to increase to preserve the fixed ratio with the unexplained variation of y_i^*. This makes effect estimates from different hierarchical logistic models, with a different set of fixed or random effects, harder to compare than their linear counterparts. An example of this pattern is provided in Section 5.3.

Taking the fixed level-1 variance into consideration, the level-2 Variance Partition Coefficient for a random-intercept logistic model (with fixed slope) becomes:

$$VPC = \frac{\sigma_{u0}^2}{\sigma_{u0}^2 + \pi^2/3}$$

which measures the amount of variation in the continuous propensity variable of the latent model (cf. Snijders and Bosker, 2012, p. 303) that is due to differences between level-2 groups (here, egos); and can be interpreted as the intra-class correlation of the latent variable, that is, the correlation between the latent support propensity of two ties attached to the same ego (Guo and Zhao, 2000; Snijders and Bosker, 2012, p. 306). Goldstein (2010, p. 127) reviews more sophisticated versions of the VPC that are available for hierarchical logistic models, which can be interpreted in terms of the observed binary dependent variable, rather than the continuous latent variable.

3.3. VARIABLES AND HYPOTHESES

Maintaining the data at the original, disaggregated level of alters and ties, hierarchical models are extremely flexible and suitable to test a host of different research hypotheses about social relationships and interactions. These models can examine the association between tie-level outcomes and at least six types of predictors, including variables observed on ties, ego-alter dyads, alters, egos, and ego-networks (Figure 3), as well as cross-level interactions between characteristics of alters or ego-alter ties and characteristics of egos or networks. Dyad-level variables often include similarity measures to test hypotheses of homophily (McPherson et al., 2001), such as the age difference in years between ego and alter, or whether ego and alter are of the same sex. Table 2 exemplifies specific hypotheses about a binary dependent variable (for example, social support) that can be tested using different types of predictors in hierarchical models for personal networks.

A distinct advantage of hierarchical modeling of personal network data is that both the individual and the contextual effect of the same predictor on ego-alter relationships can be evaluated. This is done by including both a tie or alter attribute (x_i: e.g., family tie or alter age), and the network-level summarization of the same attribute ($\bar{x}_{ego(i)}$: e.g., proportion of family ties or the average age of alters in the ego-network) as explanatory variables in the model. For example, if the predictor x_i is "family tie", we can test both whether family relationships are

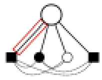
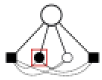
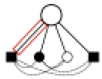
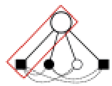
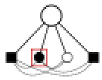
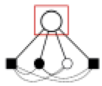
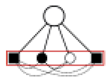
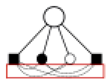
	Unit of observation	Hierarchical level	Example variable	
Dependent variables				
Characteristic of tie	Tie	Level 1	Tie provides financial support (binary)	
Characteristic of alter	Alter	Level 1	Alter’s smoking frequency (categorical)	
Explanatory variables				
Characteristic of tie	Tie	Level 1	Frequency of contact between ego and alter (categorical)	
Characteristic of ego-alter dyad	Dyad	Level 1	Ego and alter are of same sex (binary)	
Characteristic of alter	Alter	Level 1	Alter’s age (continuous)	
Characteristic of ego	Ego	Level 2	Ego’s marital status (Categorical)	
Characteristic of network composition	Network	Level 2	Proportion alters who are ego’s family (continuous)	
Characteristic of network structure	Network	Level 2	Personal network density (continuous)	

Fig. 3: Variables and levels in hierarchical models for personal networks. Large white nodes are egos; small nodes are alters (different colors and shapes represent different alter attributes).

Tab. 2: Examples of explanatory variables and corresponding testable hypotheses in hierarchical models for social support (y_i=1 if the tie provides support, y_i=0 otherwise).

Explanatory variable	**Unit of observation**	**Hierarchical level**	**Example variable**	**Example of testable hypothesis**
Characteristic of tie	Tie	Level 1	Emotional closeness between ego and alter (categorical)	Alters whom ego sees as emotionally closer are more likely to provide support to ego
Characteristic of ego-alter dyad	Dyad	Level 1	Ego and alter are of same sex (binary)	Alters who are of the same sex as ego are more likely to provide support to ego
Characteristic of alter	Alter	Level 1	Alter's age (continuous)	Older alters are more likely to provide support to ego
Characteristic of ego	Ego	Level 2	Ego's marital status (categorical)	Married egos tend to enjoy more social support overall
Characteristic of network composition	Network	Level 2	Proportion alters who are ego's family (continuous)	All alters (family and non-family) are more likely to provide support to ego when the network is more family-dominated
Characteristic of network structure	Network	Level 2	Personal network density (continuous)	All alters are more likely to provide support to ego when they are part of more cohesive (dense) networks
Interaction: Characteristic of tie × Characteristic of network composition	Tie and Network	Level 1 × Level 2	Tie is family (binary) × Proportion family alters (continuous)	Family ties are more likely to be supportive when there is a higher proportion of family members in the network
Interaction: Characteristic of alter × Characteristic of network structure	Alter and Network	Level 1 × Level 2	Alter's age (binary) × Personal network density (continuous)	Older alters are more likely to provide support in more cohesive (dense) networks
Interaction: Characteristic of alter Characteristic of ego	Alter and Ego	Level 1 × Level 2	Alter's sex (binary) × Ego's age (continuous)	Women are more likely to provide support when ego is older

more likely to provide support (individual effect of x_i), and whether social contexts that are more strongly dominated by family relationships encourage *all* contacts to provide more support (contextual effect of $\bar{x}_{ego(i)}$). Further, by including the cross-level interaction of the two covariates ($x_i \times \bar{x}_{ego(i)}$), we can test whether a single family tie is even more likely to provide support in more family-dominated networks; that is, whether the coefficient associated with x_i increases in correspondence with higher values of $\bar{x}_{ego(i)}$. In general, cross-level interactions can test whether the impact of individual alter-level or tie-level predictors changes depending on different contextual characteristics of the ego's social environment.

While multilevel models are not the only statistical method that can deal with clustered data (for an overview of other possible approaches, see Goldstein, 2010, p. 25), they represent the only framework that allows analysts to make inferences about the amount of variation that occurs between groups (by estimating level-2 variances and covariances, such as σ_{u0}^2, σ_{u1}^2 and σ_{u01} in our examples), and about the impact of group characteristics on the outcome (by including group-level variables as predictors) (Snijders and Bosker, 2012, p. 8). "Groups" in personal network analysis are egos, that is, precisely the individuals who are sampled from the main population of interest in the research. Therefore, the focus on groups implied in multilevel models is clearly essential in personal network studies, and is another major reason for the attractiveness of this approach in the field.

3.4. EXTENSIONS

Hierarchical models for personal networks can be adapted to study the *change* of personal relationships over time, including the change following specific life events such as marriage or retirement. Research on this topic relies on longitudinal data in which the personal networks of the same egos are collected at different points in time, normally using the same name generator. These data can be analyzed with hierarchical models in which changes in the alter-ego tie values are nested within egos. Examples are provided by Van Duijn et al. (1999), Terhell et al. (2007), and Mollenhorst et al. (2014). It should be noted that the hierarchical models proposed in this line of research differ from typical hierarchical models for longitudinal data, in that they do not conceive of multiple measurements at different time points as the level-1 units, with the object of measurement (here, the alter-ego tie) being the level-2 unit. Lubbers and colleagues (2010) review and discuss further approaches to the multilevel study of longitudinal personal networks.

In certain personal network sampling designs the fundamental assumptions of independence among the egos and no overlap among the personal networks are not tenable. For example, this is the case when the egos are themselves clustered in higher-level primary groups such as families or schools, and therefore they are likely to know each other or to have acquaintances in common; or when the egos are recruited through some form of link-tracing sampling. Vacca and colleagues (Under review) present an extension of the multilevel approach for personal networks to the case of overlapping egocentric data. The most important difference between such extension and what is discussed in this paper is that, when ego-networks overlap, ego-alter ties are clustered not only within egos, but also within alters – because the same alter may have *multiple* ties to multiple egos. Thus, egos and alters can be viewed as two parallel higher-level partitions of ties, which are not hierarchically nested into each other. This type of data structure can be accommodated by non-hierarchical, cross-classified multilevel models (Goldstein, 1994; Rasbash and Browne, 2008). An analogous type of multilevel model, with crossed random effects for "senders" and "receivers" of ties, is the p_2 model proposed for sociocentric network data by van Duijn and colleagues (2004).

3.5. ESTIMATION AND SOFTWARE

Goldstein (2010) and Snijders and Bosker (2012) provide an overview of the existing frequentist and Bayesian approaches for the estimation of hierarchical linear and generalized linear models. Most commonly, hierarchical *linear* models are fitted with maximum likelihood estimation via the Iterative Generalized Least Squares (IGLS) algorithm. *Restricted* maximum likelihood (REML) can be used to generate unbiased estimates of the random parameters, i.e., the variances and covariances of the random intercepts and slopes. Alternatively, Bayesian Markov Chain Monte Carlo (MCMC) methods have been implemented as well, using Gibbs sampling or Metropolis-Hasting sampling. For hierarchical *logistic* models, maximum likelihood procedures can be implemented via numerical integration. However, since these can be very computationally intensive, quasi-likelihood procedures and MCMC methods have also emerged as popular alternatives for hierarchical logistic and other generalized linear models. Different software programs are currently available to estimate multilevel linear and generalized linear models (see Goldstein, 2010 and Snijders and Bosker, 2012 for an overview). In the R environment (R Core Team 2016), which was used for the analysis presented in this paper, perhaps the most popular packages for multilevel modeling are *nlme* (Pinheiro et al., 2017) and the more recent *lme4* (Bates et al., 2015) for (restricted) maximum likelihood estimation; as well as *MCMCglmm* (Hadfield, 2010) and *blme* (Chung et al., 2013) for MCMC methods.

4. SUBSTANTIVE APPLICATIONS

Personal network methods have been used to investigate a wide variety of substantive topics in the social and health sciences, including social support (e.g., Wellman, 1979; Wellman and Wortley, 1990; Walker et al., 1993), migration and immigrant incorporation (e.g., Lubbers et al., 2007; Vacca et al., 2017), and social determinants of health outcomes (e.g., Perry and Pescosolido, 2012, 2015). In health-related research, in particular, personal networks have been studied to explain needle sharing among intravenous drug users (Latkin et al., 1995; Valente and Vlahov, 2001), risky sex behavior (Bond et al., 1999; Kennedy et al., 2012), contraceptive use (Valente et al., 1997; Kohler et al., 2001), and social support among aging adults (Gierveld and Perlman, 2006). In more recent years, with the diffusion of social media data and the development of computational social science (Lazer et al., 2009), social networking websites such as Facebook and Twitter have emerged as a new and rich source of secondary egocentric data (Takhteyev et al., 2012; Mcauley and Leskovec, 2014; Dunbar et al., 2015).

Multilevel statistical models have been used to analyze personal network data in many different substantive applications in the past twenty years. These include the study of social support and social capital (Wellman and Frank, 2001; de la Haye et al., 2012; Martí et al., 2017), immigrant incorporation (Lubbers et al., 2010; de Miguel Luken and Tranmer, 2010; van Tubergen, 2015), mental health and health-related help seeking (Perry and Pescosolido, 2015; Fulginiti et al., 2016), substance use (Green et al., 2013), sexually transmitted diseases (Kennedy et al., 2013; Hoover et al., 2016; Liu, 2016), transportation and communication (Carrasco and Miller, 2009; Carrasco and Cid-Aguayo, 2012; Matous et al., 2013), urban communities (Völker and Flap, 2007; Mok et al., 2010), civic activity (Beyerlein and Bergstrand, 2016), electoral patterns (Bello and Rolfe, 2014), digital learning in secondary schools (Ünlüsoy et al., 2013), instructional development among university faculty (Waes et al., 2015), and scientific collaboration (Jha and Welch, 2010).

Social support, particularly in its relationship with health outcomes, is one of the substantive topics that have been most extensively studied using multilevel models for personal networks, as well as the subject of the illustrative example in this paper. Social support refers to the amount and diversity of the material and emotional resources that an individual may gain from social relationships (House et al., 1988; Song et al., 2011). Personal networks are a particularly well-suited construct to operationalize this concept. While, in the first studies in this area, various indexes of personal network characteristics were used to create respondent-level measures of social support (see Song et al., 2011 for a review), more recently multilevel models have moved the focus to the level of the single ties between potential support recipients (respondents) and providers (alters). Thus, the multilevel framework has allowed researchers to analyze the factors that facilitate or encourage the exchange of social support in specific relationships between two individuals. Most of this research has used hierarchical generalized linear models for categorical dependent variables, including dichotomous (e.g., supportive versus non-supportive ties) and polytomous outcomes (e.g., non-supportive ties, materially supportive ties, emotionally supportive ties).

Wellman and Frank (2001) provide an early example of this application, using the famous East York personal network data (Wellman, 1979) to test a wide array of hypotheses about social support grounded in sociological theory. This study fits multilevel logistic models to predict two binary dependent variables, namely, whether an alter provides support to ego (*i*) in everyday situations and (*ii*) in emergency situations. The models allow for both intercepts and slopes to include ego random effects, and account for cross-level interactions between characteristics of ego-alter ties and characteristics of ego-networks. Findings show that the supportiveness of social relationships is affected by a number of predictors at all levels, including characteristics of the relationship itself (e.g., emotionally stronger ties provide more support); characteristics of potential support providers (e.g., female alters are more likely to provide emergency support); characteristics of potential support recipients (e.g., female egos are more likely to obtain any type of support); and characteristics of the entire personal network (e.g., any alter is more likely to provide support when the whole network, on average, is more accessible to ego). Certain cross-level interactions between individual and contextual variables also emerge as significant predictors of support exchange. For example, the emotional strength of ties is more strongly associated with everyday support *for male egos*, while this effect is weaker for female egos. In other words, men are more likely to receive everyday support from emotionally closer ties only, while women tend to receive everyday support from all ties regardless of their closeness.

In a similar application, Völker and Flap (2001) analyze data about personal relationships in former East Germany before and after the downfall of the Berlin Wall. They use random-intercept logistic models to examine the factors associated with the development of intimate, “niche” relationships with whom respondents discussed personal and political matters, compared with instrumental, “provision” relationships from whom they obtained specific goods and services. More recent research has applied hierarchical generalized linear models to study support exchange between immigrants and native residents in Spain (de Miguel Luken and Tranmer, 2010); the different roles of local and transnational relationships in providing support to German immigrants in the UK (Herz, 2015); the provision of material and emotional support to homeless youth in the US (de la Haye et al., 2012); the effects of ego-network structural characteristics, such as cohesion and cliquishness, on support provision (Martí et al., 2017); and the individual and contextual factors associated with health-related discussions at the time of first entry into mental health treatment (Perry and Pescosolido, 2015).

5. ILLUSTRATIVE EXAMPLE: SOCIAL SUPPORT AMONG SRI LANKAN IMMIGRANTS IN MILAN, ITALY

5.1. DATA

This section illustrates the models presented in Section 3 using data from a personal network survey conducted in 2012 among Sri Lankan immigrants in Milan, Italy (Vacca et al., 2017). The survey respondents, male Sri Lankan residents in Milan (N=102), were recruited in two ways. About 70% of the sample responded to informational materials, such as leaflets and posters, that was circulated in central places of the city, particularly within Sri Lankan ethnic neighborhoods, including metro and tram stations, Sri Lankan churches and temples, street markets, and Sri Lankan diplomatic buildings. The remaining 30% of the sample was recruited through snowball sampling starting from a dozen of key informants in the Milan Sri Lankan community, including Sri Lankan political organizers and leaders of cultural associations; directors of Sri Lankan schools; leaders of Sri Lankan religious organizations; managers of local Sri Lankan TV networks; employers and employees in Sri Lankan businesses.

Personal networks were elicited through the following name generator:

"Would you list the names of 45 persons whom you know and who know you, with whom you have had some contact in the past two years (face-to-face, by phone, or by the Internet), and whom you could still contact if you needed to?" The requirement of a fixed-size list of 45 alters aimed to obtain an extensive sample of the respondent's total personal network (McCarty et al., 1997), including both strong and weak ties. The resulting networks comprised ego's family members, friends, and acquaintances; personal contacts from different nationalities and living in different countries; and contacts from diverse social settings, including the workplace, the neighborhood, and churches or temples.

Respondents were also asked a set of fixed name interpreters about each alter, including questions about sex, age, nationality, and support provision from each contact. Finally, the ego-network structural data were collected by asking respondents to report on the relationship existing in every pair of alters. Specifically, respondents were asked how likely it was that each pair of alters independently met or talked to each other. For each of 990 pairs of alters (all the undirected pairs between 45 actors), the ego could answer that the two contacts *certainly* met and talked with each other, *maybe* met and talked with each other, or *certainly did not* meet and talk with each other.[5] Here two alters are considered as connected if ego has indicated that they meet and talk with each other either certainly or maybe.

5.2. VARIABLES AND ANALYTICAL STRATEGY

In the examples that follow, we use hierarchical logistic regression to model the probability of ego-alter ties providing (potential) *financial support* to egos. Financial aid is a central aspect of the support provided by social networks, and it has often been the subject of social support research (Walker et al., 1993; Song et al., 2011). The binary dependent variable is observed for each tie between an

[5] This section of the survey was completed using spreadsheet software that presented the respondent with an alter-alter matrix to be filled in. Interviewers filled the matrix with response codes (blank for "Certainly do not know each other", 1 for "Certainly know each other", 2 for "Maybe know each other") based on the respondent's answers. The matrix was automatically produced from previous responses to the name generator and name interpreters. Alters were automatically ordered in the matrix rows and columns based on the type of relationship (close family, extended family, friend, acquaintance) and the context of sociability (family, work, neighborhood, free time, etc.), which had been collected in the previous survey section. The alter-alter matrix, with a meaningful order of alters in rows and columns, made the respondent's task of evaluating alter-alter ties cognitively easier. On average, responses on all alter-alter ties were provided in 30 to 45 minutes.

ego and an alter, and indicates whether the tie is financially supportive for ego ($y_i = 1$) or not ($y_i = 0$). The variable was generated by the following survey question: "Can you ask [alter] for a money loan of at least €300?" (Yes=1, No=0).

The models in this section include the following explanatory variables:

i. Tie/dyad characteristics: The alter is a member of ego's family ("Family relationship"); Ego and alter are in the same age bracket (Age brackets: 18-25, 26-30, 31-35, 36-40, 41-50, 51-60, > 60 years.)
ii. Ego characteristics: Age of ego (in years, divided by 5); Ego is employed (reference category: Yes); Ego's educational level (reference category: Primary school).
iii. Alter characteristics: Sex of alter; Alter's age (age brackets as above, treated as continuous variable); Alter's nationality (reference category: Not Sri Lankan); Alter's degree in the ego-network (divided by 5, i.e., rescaled to units of 5 connections).
iv. Network characteristics: Number of family members in the network (divided by 5); Average degree in the ego-network (divided by 5).

All continuous explanatory variables used in the models are centered around the mean. Average ego-network degree is a measure of overall connectedness or cohesion, which is proportional to network density (Borgatti et al., 2013).

We start by estimating M1, a random-intercept, fixed-slope hierarchical logistic model with no explanatory variable for the probability of a tie being financially supportive, including the two levels of ties and egos. Models of this type, sometimes referred to as "variance components models" (Goldstein, 2010, p. 19), are empty models used to assess the distribution of the outcome variation between the different levels, before taking any predictor into consideration. In our case, these empty models evaluate the relative importance of the ties (level 1) vis-à-vis the egos (level 2) as a source of variation in support provision. Model M1 is represented by Equations (9)-(12) in Section 3.2, once all terms associated with predictors x_i and $w_{ego(i)}$ are removed. Thus, M1 only includes the random intercept $\beta_{0,ego(i)}$, generated as follows:

$$\beta_{0,ego(i)} = \gamma_{00} + u_{0,ego(i)}$$

As a result, the only parameters to be estimated in M1 are γ_{00} and σ^2_{u0}, the random-intercept variance. The single-level version of M1 is a standard logistic empty model with a fixed intercept and no predictor.

Model M2 is subsequently constructed by adding the tie-level, ego-level, alter-level, and network-level explanatory variables listed above. This model reflects common sociological hypotheses about individual and contextual factors affecting the provision of social support, including the hypotheses that financial aid is more likely to be provided by family ties, by alters with similar sociodemographic characteristics (here, in the same age bracket), by co-ethnic social networks, and by contacts who share a higher number of mutual acquaintances with the ego (Walker et al., 1993; Portes and Sensenbrenner, 1993; Wellman and Frank, 2001). Similar to M1, model M2 includes a random intercept but no random slope. Thus, M2 is defined by Equations (9)-(10) and (12), with β_1 being treated as a fixed parameter (rather than the random variable defined by Equation (11)). The single-level version of M2 is a standard logistic model that includes the same predictors as M2 but disregards tie clustering and assumes the intercept to be a fixed parameter.

The final model shown here, M3, is obtained by adding a random slope for the "Family relationship" predictor. Model M3 is described by Equations (9)-(12), with x_i (the predictor with a random slope) representing "Family relationship"; all other predictors in M3 (not included in Equations (9)-(12)) have fixed coefficients.

All hierarchical models are estimated with the *glmer* function from the *lme4* R package, which implements maximum likelihood estimation via numerical integration. The residual deviance of the estimated models is calculated as -2*LogLikelihood, based on the unconditional absolute likelihood. Wald tests for individual predictors are conducted to assess the significance of fixed effects. The significance of random effects is evaluated using Likelihood Ratio Tests (LRT) for the hypotheses that the variance of the random intercept (σ_{u0}^2) and the variance of the random slope of "Family relationship" (σ_{u1}^2) are significantly different from zero (Goldstein 2010:41; Snijders and Bosker 2012:97). The LRTs compare M1 and M2 with their respective single-level versions (tests for random intercept), and M3 with M2 (test for random slope). To assess the variability of estimators, *lme4* provides standard errors for fixed effect parameters and confidence intervals for variance component parameters.[6]

[6] Due to the asymmetry of estimator distributions for variance and covariance parameters, standard errors are considered inadequate measures of estimator variability for variance components, and are not calculated by *lme4* (Bates et al. 2015). However, confidence intervals for these parameters can be obtained in *lme4* using the *confint* function on the model (*merMod*) object.

5.3. RESULTS AND DISCUSSION

Random-intercept models

The 45 ties for each of 102 egos generate a potential sample size of N = 4590 in the alter-level dataset; after removing ties with missing values on the dependent or explanatory variables, this reduces to N=3972.[7] Results for M1 (Table 3) show evidence that the provision of financial support in our sample is indeed clustered by egos, with part of the outcome variation due to systematic differences between respondents, and the ties associated with the same respondent being correlated in their tendency to provide support. The variance of level-2 residuals is estimated as $\hat{\sigma}_{u0}^2 = 0.81$, corresponding to VPC=0.20. This means that approximately 20% of the variation in the tendency of ties to provide support (in the log-odds scale) is attributable to recurrent differences between egos. Equivalently, 0.20 is the estimated correlation between two ties nominated by the same ego in their latent propensity to provide financial support.[8]

A LRT comparing M1 with its single-level version shows that the addition of a random effect on the intercept (i.e., the addition of the σ_{u0}^2 parameter) reduces the deviance by 416.31 units (5570.77 in the single-level model minus 5154.46 in M1). This indicates a strong improvement in model fit, and is obviously a significant decrease when compared with a χ_1^2 distribution. We conclude that the overall likelihood of support in our data does vary significantly across egos, with some respondents systematically reporting more supportive ties, and others regularly reporting less supportive ones.

In the results for M2 a number of hypotheses are supported by the fixed coefficient estimates. Family ties are strongly more likely to provide financial support, with family relationships having 42% higher odds to provide monetary aid compared with non-related ties. While the coefficient interpretation in terms of odds-ratio is equivalent to single-level logistic-models, it should be noted that fixed coefficients in a multilevel model express cluster-specific, within-group effects. In this case groups are egos, meaning that, for example, the family ties of a given ego have 42% higher odds of providing support compared with non-family ties *of the same ego*; in other words, the interpretation is conditioned on $u_{0,ego(i)}$ being held constant, and we are not comparing ties from different egos.

---

7 Of the 618 cases that were removed due to incomplete data, 92% (569 cases) had a missing value on the dependent variable (whether the tie is financially supportive). Of the remaining 49 cases, 42 had a missing value on ego's age, 6 on alter's age, and 1 on both variables.

8 This interpretation assumes a threshold model for the continuous latent variable representing the propensity of an ego-alter tie to provide financial support (Snijders and Bosker 2012:303).

Among the ego-level predictors, age and employment status do not appear to have a significant effect on the likelihood of support in M2, indicating that older and unemployed Sri Lankan immigrants are as likely to receive financial support, on average, as younger and employed respondents. By contrast, an unexpected finding is that higher educational level of ego is significantly and strongly associated with higher likelihood of receiving support – a result that does not change in models adding further ego-level controls, such as marital status and monthly income (results not shown). The finding suggests that Sri Lankan immigrants with higher educational attainment tend to be surrounded by financially more supportive personal communities. While our data do not include educational and socioeconomic information about the alters (i.e., the potential support providers), we can interpret this effect on the basis of two well-established facts in sociological research: educational level is associated with socioeconomic status and wealth (see Orr, 2003; Sirin, 2005 for recent reviews); and people tend to associate with peers with similar education and socioeconomic status (McPherson et al., 2001). Thus, we can speculate that Sri Lankan immigrants with higher educational level are embedded in networks of family and friends (including those left behind in Sri Lanka) who also have higher education, and presumably higher socioeconomic status, being therefore more able and inclined to provide financial aid to our respondents.

The results for alter-level predictors show that Sri Lankan immigrants are more likely to obtain support from their older contacts, with each subsequent age bracket increasing the odds of support provision by approximately 7% for a given ego. As expected, co-ethnic acquaintances are substantially more likely to provide support than non-Sri Lankan ones, with Sri Lankan alters having 68% higher odds to provide financial aid. This is evidence of a clustering of support exchanges within the co-ethnic immigrant community. Our estimates also confirm the "Simmelian hypothesis" that alter's degree in the ego-network is positively associated with the likelihood of providing support (Wellman and Frank, 2001). Five additional common contacts between alter and ego are predicted to increase the odds of support by 43-46% in M2 and M3. Thus, everything else being equal, alters are more likely to provide support when they have a higher number of family and friends in common with ego. In other words, what sociologists call the "structural embeddedness" of social relationships does appear to facilitate stronger bonds and more supportive relationships between individuals in Sri Lankan personal communities.

The addition of predictors in M2 does not change the estimate of the between-ego variance compared with M1 ($\hat{\sigma}_{u0}^2 = 0.81$). A LRT for M2 versus its single-level version shows a reduction of deviance from 5094.34 to 4723.09 (*p-value* < 0.001 on χ_1^2), confirming that there is statistically significant between-ego variance in the likelihood of the ties being supportive.

Tab. 3: Hierarchical logistic models for financial support. Dependent variable: y_i=1 if ego-alter tie provides financial support, $y_i = 0$ otherwise. N = 3972. Standard Errors are reported for fixed effects, Confidence Intervals are reported for variance components (cf. Section 5.2). [a] Variable centered around its mean and divided by 5. [b] Ref category: *Primary*. [c] Variable centered around its mean. [d] Ref category: *Not Sri Lankan*. [e] The LRTs compare, in this order: (1) M1 vs. its single-level version with fixed intercept; (2) M2 vs. its single-level version with same predictors and fixed intercept; (3) M3 vs. M2. * p-value < .05, ** p-value < .01, *** p-value < .001 (Wald tests).

	M1	M2	M3
Fixed effects (SE)			
Intercept	0.05 (0.10)	-0.67 (0.20) ***	-0.67 (0.21) **
Tie/dyad predictors			
Family relationship		0.35 (0.11) **	0.38 (0.14) **
Same age bracket		0.07 (0.08)	0.07 (0.08)
Ego predictors			
Age (years) [a]		-0.01 (0.05)	-0.02 (0.05)
Employed: No		-0.20 (0.25)	-0.28 (0.26)
Education: Secondary [b]		0.53 (0.22) *	0.57 (0.22) **
Education: University [b]		1.01 (0.31) **	1.1 (0.32) ***
Alter predictors			
Sex: Female		-0.12 (0.09)	-0.12 (0.09)
Age (brackets) [c]		0.07 (0.03) **	0.07 (0.03) **
Nationality: Sri Lankan [d]		0.52 (0.11) ***	0.52 (0.11) ***
Ego-network degree [a]		0.36 (0.03) ***	0.38 (0.03) ***
Network predictors			
Number family members [a]		-0.16 (0.11)	-0.19 (0.11)
Avg degree in ego-net [a]		-0.22 (0.11)	-0.23 (0.11) *
Variance components (CI)			
Egos: Intercept (σ_{u0}^2)	0.81 (0.58, 1.14)	0.81 (0.58, 1.15)	0.93 (0.65, 1.35)
Egos: Slope of "Family relationship" (σ_{u1}^2)			0.55 (0.21, 1.11)
Egos: Intercept-Slope covariance (σ_{u01})			-0.28 (-0.72, 0.03)
-2LogLikelihood	5154.46	4723.09	4707.49
LRT: P-value [e]	<.001	<.001	<.001

Adding a random slope

In the random-slope model M3, the association between a tie being a "Family relationship" and its likelihood of providing support is allowed to vary across egos. As a result, the "Family relationship" coefficient becomes a random variable, whose variance is estimated as $\hat{\sigma}^2_{u1} = 0.55$. A LRT comparing M3 with M2 confirms that the random slope significantly improves model fit, suggesting that the difference between family and non-family ties, in their tendency to provide support, does vary systematically across respondents (*p-value* < 0.001 on χ^2_2). In other words, while for some respondents financial support is primarily or exclusively provided by family relationships (high $\beta_{1,ego(i)}$), others report similar levels of support provided by family and non-family ties (low $\beta_{1,ego(i)}$).

The negative estimate for the covariance between the random intercept and the random slope ($\sigma_{u01} \equiv cov(u_{0,ego(i)}, u_{1,ego(i)})$) is of substantive interest, revealing that the respondents who tend to enjoy a higher overall level of financial support (higher $\beta_{0,ego(i)}$) are also those who report a smaller difference between family and non-family ties in terms of supportiveness (lower $\beta_{1,ego(i)}$). This pattern, which is akin to what is sometimes called the "fanning in" of regression lines in hierarchical linear models, suggests that the Sri Lankan family is more central as a support provider when support is a scarcer resource in general. In other words, the Sri Lankan immigrants whose personal communities are overall less able or inclined to provide financial aid, are also those who more strongly (or exclusively) rely on kin relationships to obtain support. By contrast, the immigrants who are embedded in broadly more supportive communities, are able to more evenly distribute their support requests among family and non-family ties, and therefore attribute less importance to family for obtaining financial aid.

Among the network-level characteristics, overall network cohesion (as measured by average degree) is significantly and negatively associated with the likelihood of support in the best-fitting model M3. However, this network-level predictor is the average of a tie-level explanatory variable, namely, individual alter degree. Therefore, the coefficients of the two predictors need to be interpreted in conjunction. The level-1 coefficient measures the positive *within-ego* effect of alter degree: within the same ego-network, alters with 5 additional contacts in common with ego have 46% higher odds of providing support. The *sum* of the level-1 and level-2 coefficients (0.38 – 0.23 = 0.15) measures the *between-ego* effect: egos with a more cohesive network (+5 in average degree), have on average 16% higher odds of receiving support from their ties. Thus, across different ego-networks, higher average degree (i.e., higher cohesion of the personal community) is still positively associated with average support provision. Finally, the negative level-2 coefficient measures the contextual effect of average degree, above and beyond the within-group effect of individual alter degree, on the tie-level outcome for ties from different networks. Thus, while individual alters with higher degree are more likely to be supportive *within the same network*, when we compare different networks, an alter (with the same individual degree) has a lower likelihood of being supportive if he is part of a network with higher average degree (21% lower odds with 5 additional units of ego-network average degree). Substantively, this might reflect the fact that central alters are more likely to provide support in sparser networks, in which they are rare.

Models M2 and M3 lead to substantially similar conclusions in terms of direction and significance of fixed effects. However, a pattern of higher parameter estimates in M3 emerges, which reflects the issue, discussed in Section 3.2, of fixed effect estimates increasing (in absolute value) when random effects are added to a hierarchical logistic model (Snijders and Bosker, 2012, p. 307). In our example, most fixed effect estimates increase by a similar proportional amount (5-9%) in M3 compared with M2, and the intercept variance estimate is higher as well, due to the addition of a random effect for the slope of "Family relationship". This pattern is an artifact of the models' formulation, and should not be attributed substantive meaning. In fact, it should be noted that the *ratio* between parameter estimates is approximately constant in the two models.

6. CONCLUSION

This paper aimed to introduce the reader to personal network data and to the multilevel statistical framework for personal network analysis. The article described the multilevel nature of personal network data and delineated basic hierarchical models that can be used to account for it. We emphasized the unique ability of these models to partition the outcome variation between the different levels of individuals and ties, to quantify the proportion of variability that is attributable to systematic differences between egos (i.e., respondents), and to measure the impact of ego and network characteristics on social relationships. This ability allows personal network analysts to test an exceptionally wide set of research hypotheses, which cannot be investigated with other methods. The increasing number and diversity of their substantive applications, which the article outlined, attest to the high adaptability and flexibility of multilevel regression methods for personal networks. Future work could contribute to this line of research by exploring more sophisticated multilevel models, incorporating more than two levels and non-hierarchical forms of clustering, to account for overlapping personal networks, for the nesting of egos within higher-level groups, and for complex multilevel structures arising in longitudinal personal network designs.

REFERENCES

Bates, D., Mächler, M., Bolker, B. and Walker, S. (2015). Fitting linear mixed-effects models using lme4. In *Journal of Statistical Software*, *67*(1): 1–48. doi:10.18637/jss.v067.i01.

Bello, J. and Rolfe, M. (2014). Is influence mightier than selection? Forging agreement in political discussion networks during a campaign. In *Social Networks*, *36*: 134–146. doi:10.1016/j.socnet.2013.06.001.

Beyerlein, K. and Bergstrand, K. (2016). It takes two: A dyadic model of recruitment to civic activity. In *Social Science Research*, *60*: 163–180. doi:10.1016/j.ssresearch.2016.07.008.

Bond, K.C., Valente, T.W. and Kendall, C. (1999). Social network influences on reproductive health behaviors in urban northern Thailand. In *Social Science & Medicine*, *49*(12): 1599–1614. doi:10.1016/S0277-9536(99)00205-1.

Borgatti, S.P., Everett, M.G. and Johnson, J.C. (2013). *Analyzing Social Networks*. Los Angeles; London: SAGE Publications.

Browne, W.J., Goldstein, H. and Rasbash, J. (2001). Multiple membership multiple classification (MMMC) models. In *Statistical Modelling*, *1*(2): 103–124. doi:10.1177/1471082X0100100202.

Burt, R.S. (1984). Network items and the general social survey. In *Social Networks*, *6*(4): 293–339. doi:10.1016/0378-8733(84)90007-8.

Campbell, K.E. and Lee, B.A. (1991). Name generators in surveys of personal networks. In *Social Networks*, *13*(3): 203–221. doi:10.1016/0378-8733(91)90006-F.

Carrasco, J.A. and Cid-Aguayo, B. (2012). Network capital, social networks, and travel: An empirical illustration from Concepción, Chile. In *Environment and Planning A*, *44*(5): 1066–1084. doi:10.1068/a43222.

Carrasco, J.A. and Miller, E.J. (2009). The social dimension in action: A multilevel, personal networks model of social activity frequency between individuals. In *Transportation Research Part A: Policy and Practice*, *43*(1): 90–104. doi:10.1016/j.tra.2008.06.006.

Chua, V., Madej, J. and Wellman, B. (2011). Personal communities: The world according to me. In J. Scott & P.J. Carrington (Eds.), *The Sage Handbook of Social Network Analysis* (pp. 101–115). SAGE Publications.

Chung, Y., Rabe-Hesketh, S., Dorie, V., Gelman, A. and Liu, J. (2013). A nondegenerate penalized likelihood estimator for variance parameters in multilevel models. In *Psychometrika*, *78*(4): 685–709.

Crossley, N., Bellotti, E., Edwards, G., Everett, M.G., Koskinen, J. and Tranmer, M. (2015). In *Social Network Analysis for Ego-Nets*. London: SAGE Publications.

de la Haye, K., Green, H.D., Kennedy, D.P., Zhou, A., Golinelli, D., Wenzel, S.L. and Tucker, J.S. (2012). Who is supporting homeless youth? Predictors of support in personal networks. In *Journal of Research on Adolescence*, *22*(4): 604–616. doi:10.1111/j.1532-7795.2012.00806.x.

de Miguel Luken, V. and Tranmer, M. (2010). Personal support networks of immigrants to Spain: A multilevel analysis. In *Social Networks*, *32*(4): 253–262. doi:10.1016/j.socnet.2010.03.002.

Dunbar, R.I.M., Arnaboldi, V., Conti, M. and Passarella, A. (2015). The structure of online social networks mirrors those in the offline world. In *Social Networks*, *43*: 39–47. doi:10.1016/j.socnet.2015.04.005.

Fischer, C.S. (1982). What do we mean by 'friend'? An inductive study. In *Social Networks*, *3*(4): 287–306. doi:10.1016/0378-8733(82)90004-1.

Fulginiti, A., Pahwa, R., Frey, L.M., Rice, E. and Brekke, J.S. (2016). What factors influence the decision to share suicidal thoughts? A multilevel social network analysis of disclosure among individuals with serious mental illness. In *Suicide & Life-Threatening Behavior*, *46*(4): 398–412. doi:10.1111/sltb.12224.

Gelman, A. and Hill, J. (2006). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge ; New York: Cambridge University Press.

Gierveld, J. and Perlman, D. (2006). Long-standing nonkin relationships of older adults in the Netherlands and the United States. In *Research on Aging*, *28*(6): 730–748. doi:10.1177/0164027506291873.

Goldstein, H. (1994). Multilevel cross-classified models. In *Sociological Methods & Research*, *22*(3): 364–375. doi:10.1177/0049124194022003005.

Goldstein, H. (2010). *Multilevel Statistical Models* (4 edition.). Chichester, West Sussex: Wiley.

Green, H.D., de la Haye, K., Tucker, J.S. and Golinelli, D. (2013). Shared risk: Who engages in substance use with American homeless youth? In *Addiction*, *108*(9): 1618–1624. doi:10.1111/add.12177.

Guo, G. and Zhao, H. (2000). Multilevel modeling for binary data. In *Annual Review of Sociology*, *26*, 441–462.

Hadfield, J.D. (2010). MCMC methods for multi-response Generalized Linear Mixed Models: The MCMCglmm R package. In *Journal of Statistical Software*, *33*(2): 1-22. doi:10.18637/jss.v033.i02.

Herz, A. (2015). Relational constitution of social support in migrants' transnational personal communities. In *Social Networks*, *40*: 64–74. doi:10.1016/j.socnet.2014.08.001.

Hoover, M.A., Green, H.D., Bogart, L.M., Wagner, G.J., Mutchler, M.G., Galvan, F.H. and McDavitt, B. (2016). Do people know I'm poz? Factors associated with knowledge of serostatus among HIV-positive African Americans' social network members. In *AIDS and Behavior*, *20*(1): 137–146. doi:10.1007/s10461-015-1039-5.

House, J.S., Umberson, D. and Landis, K.R. (1988). Structures and processes of social support. In *Annual Review of Sociology*, *14*(1): 293–318. doi:10.1146/annurev.so.14.080188.001453.

Jha, Y. and Welch, E.W. (2010). Relational mechanisms governing multifaceted collaborative behavior of academic scientists in six fields of science and engineering. In *Research Policy*, *39*(9): 1174–1184. doi:10.1016/j.respol.2010.06.003.

Kennedy, D.P., Tucker, J.S., Green, H.D., Golinelli, D. and Ewing, B. (2012). Unprotected sex of homeless youth: Results from a multilevel dyadic analysis of individual, social network, and relationship factors. In *AIDS and behavior*, *16*(7): 2015–2032. doi:10.1007/s10461-012-0195-0.

Kennedy, D.P., Wenzel, S.L., Brown, R., Tucker, J.S. and Golinelli, D. (2013). Unprotected sex among heterosexually active homeless men: Results from a multi-level dyadic analysis. In *AIDS and Behavior*, *17*(5): 1655–1667. doi:10.1007/s10461-012-0366-z.

Kohler, H.-P., Behrman, J.R. and Watkins, S.C. (2001). The density of social networks and fertility decisions: Evidence from South Nyanza district, Kenya. In *Demography*, *38*(1): 43–58. doi:10.1353/dem.2001.0005.

Krivitsky, P.N. and Morris, M. (2017). Inference for social network models from egocentrically sampled data, with application to understanding persistent racial disparities in HIV prevalence in the US. In *The Annals of Applied Statistics*, *11*(1): 427–455. doi:10.1214/16-AOAS1010.

Latkin, C., Mandell, W., Vlahov, D., Knowlton, A., Oziemkowska, M. and Celentano, D. (1995). Personal network characteristics as antecedents to needle-sharing and shooting gallery attendance. In *Social Networks*, *17*(3–4): 219–228. doi:10.1016/0378-8733(95)00263-N.

Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A.-L., Brewer, D., et al. (2009). Computational social science. In *Science*, *323*(5915): 721–723. doi:10.1126/science.1167742.

Lin, N. and Dumin, M. (1986). Access to occupations through social ties. In *Social Networks*, *8*(4): 365–385. doi:10.1016/0378-8733(86)90003-1.

Liu, H. (2016). Egocentric network and condom use among mid-age female sex workers in China: A multilevel modeling analysis. In *AIDS Patient Care and STDs*, *30*(4): 155–165. doi:10.1089/apc.2015.0349.

Long, J.S. (1997). *Regression Models for Categorical and Limited Dependent Variables*. Thousand Oaks: Sage Publications.

Louch, H. (2000). Personal network integration: Transitivity and homophily in strong-tie relations. In *Social Networks*, *22*(1): 45–64. doi:10.1016/S0378-8733(00)00015-0.

Lubbers, M.J., Molina, J.L., Lerner, J., Brandes, U. and Ávila, J. (2010). Longitudinal analysis of personal networks. The case of Argentinean migrants in Spain. In *Social Networks*, *32*(1): 91-104.

Lubbers, M.J., Molina, J.L. and McCarty, C. (2007). Personal networks and ethnic identifications. The case of migrants in Spain. In *International Sociology*, *22*(6): 721–741. doi:10.1177/0268580907082255.

Marin, A. and Hampton, K.N. (2007). Simplifying the personal network name generator. Alternatives to traditional multiple and single name generators. In *Field Methods*, *19*(2): 163–193. doi:10.1177/1525822X06298588.

Marin, A. and Wellman, B. (2011). Social network analysis: An introduction. In J. Scott & P.J. Carrington (Eds.), *The Sage Handbook of Social Network Analysis* (pp. 11–25). SAGE Publications.

Marsden, P.V. (1990). Network data and measurement. In *Annual Review of Sociology*, *16*, 435–463.

Martí, J., Bolíbar, M. and Lozares, C. (2017). Network cohesion and social support. In *Social Networks*, *48*: 192–201. doi:10.1016/j.socnet.2016.08.006.

Matous, P., Todo, Y. and Mojo, D. (2013). Boots are made for walking: Interactions across physical and social space in infrastructure-poor regions. In *Journal of Transport Geography*, *31*: 226–235. doi:10.1016/j.jtrangeo.2013.04.001.

Mcauley, J. and Leskovec, J. (2014). Discovering social circles in ego networks. In *ACM Transactions on Knowledge Discovery from Data*, *8*(1): 1–28, article no.4. doi:10.1145/2556612.

McCarty, C. (2002). Structure in personal networks. In *Journal of Social Structure*, *3*(1): 1–11.

McCarty, C., Bernard, H.R., Killworth, P.D., Shelley, G. A. and Johnsen, E. C. (1997). Eliciting representative samples of personal networks. In *Social Networks*, *19*(4): 303–323. doi:10.1016/S0378-8733(96)00302-4.

McCarty, C., Lubbers, M.J., Vacca, R. and Molina, J.L. (Forthcoming). *Personal Networks.* New York: Guilford Press.

McCarty, C. and Wutich, A. (2005). Conceptual and empirical arguments for including or excluding Ego from structural analyses of personal networks. In *Connections*, *26*(2): 80–86.

McCulloch, C.E. and Searle, S.R. (2001). *Generalized, Linear, and Mixed Models*. New York: John Wiley & Sons.

McPherson, M., Smith-Lovin, L. and Brashears, M.E. (2006). Social isolation in America: Changes in core discussion networks over two decades. In *American Sociological Review*, *71*(3): 353–375. doi:10.1177/000312240607100301.

McPherson, M., Smith-Lovin, L. and Cook, J.M. (2001). Birds of a feather: Homophily in social networks. In *Annual Review of Sociology*, *27*: 415–444.

Mok, D., Wellman, B. and Carrasco, J. (2010). Does distance matter in the age of the Internet? In *Urban Studies*, *47*(13): 2747–2783. doi:10.1177/0042098010377363.

Mollenhorst, G., van Duijn, M.A.J., Rydgren, J. and Edling, C. (2016). Triadic closure in core networks: Disentangling the effects of social distance, national origin similarity and shared contexts. In *International Review of Social Research*, *6*(4): 146–162. doi:10.1515/irsr-2016-0018.

Mollenhorst, G., Völker, B. and Flap, H. (2011). Shared contexts and triadic closure in core discussion networks. In *Social Networks*, *33*(4): 292–302. doi:10.1016/j.socnet.2011.09.001.

Mollenhorst, G., Volker, B. and Flap, H. (2014). Changes in personal relationships: How social contexts affect the emergence and discontinuation of relationships. In *Social Networks*, *37*: 65–80. doi:10.1016/j.socnet.2013.12.003.

Orr, A. J. (2003). Black-white differences in achievement: The importance of wealth. In *Sociology of Education*, *76*(4): 281–304. doi:10.2307/1519867.

Perry, B.L., Borgatti, S. and Pescosolido, B.A. (2018). *Egocentric Network Analysis: Foundation, Methods, and Models*. Cambridge University Press.

Perry, B.L. and Pescosolido, B.A. (2010). Functional specificity in discussion networks: The influence of general and problem-specific networks on health outcomes. In *Social Networks*, *32*(4): 345–357. doi:10.1016/j.socnet.2010.06.005.

Perry, B.L. and Pescosolido, B.A. (2012). Social network dynamics and biographical disruption: The case of "first-timers" with mental illness. In *American Journal of Sociology*, *118*(1): 134–175. doi:10.1086/666377.

Perry, B.L. and Pescosolido, B.A. (2015). Social network activation: The role of health discussion partners in recovery from mental illness. In *Social Science & Medicine*, *125*: 116–128. doi:10.1016/j.socscimed.2013.12.033

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D. and R Core Team. (2017). *nlme: Linear and Nonlinear Mixed Effects Models*. https://CRAN.R-project.org/package=nlme.

Portes, A. (1998). Social capital: Its origins and applications in modern sociology. In *Annual Review of Sociology*, *24*: 1–24.

Portes, A. and Sensenbrenner, J. (1993). Embeddedness and Immigration: notes on the Social Determinants of Economic Action. In *American Journal of Sociology*, *98*(6): 1320–1350.

R Core Team. (2016). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Rasbash, J. and Browne, W.J. (2008). Non-hierarchical multilevel models. In J. de Leeuw & E. Meijer (Eds.), *Handbook of Multilevel Analysis* (pp. 301–334). New York: Springer.

Robinson, W.S. (1950). Ecological correlations and the behavior of individuals. In *American Sociological Review*, *15*(3), 351–357. doi:10.2307/2087176.

Searle, S.R., Casella, G. and McCulloch, C.E. (1992). *Variance Components* (1 edition.). Hoboken, NJ: Wiley-Interscience.

Sirin, S.R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. In *Review of Educational Research*, *75*(3), 417–453. doi:10.3102/00346543075003417.

Snijders, T. and Bosker, R.J. (2012). *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling* (2nd edition.). London; Thousand Oaks, Calif: SAGE Publications.

Snijders, T., Spreen, M. and Zwaagstra, R. (1995). The use of multilevel modeling for analysing personal networks: Networks of cocaine users in an urban area. In *Journal of Quantitative Antropology*, *5*: 85–105.

Song, L., Son, J. and Lin, N. (2011). Social support. In J. Scott & P.J. Carrington (Eds.), *The SAGE Handbook of Social Network Analysis* (pp. 116–128). SAGE Publications.

Takhteyev, Y., Gruzd, A. and Wellman, B. (2012). Geography of Twitter networks. In *Social Networks*, *34*(1): 73–81. doi:10.1016/j.socnet.2011.05.006.

Terhell, E.L., Broese van Groenou, M.I. and van Tilburg, T. (2007). Network contact changes in early and later postseparation years. In *Social Networks*, *29*(1): 11–24. doi:10.1016/j.socnet.2005.11.006.

Tranmer, M., Steel, D. and Browne, W.J. (2014). Multiple-membership multiple-classification models for social network and group dependences. In *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *177*(2): 439–455. doi:10.1111/rssa.12021.

Ünlüsoy, A., de Haan, M., Leander, K. and Völker, B. (2013). Learning potential in youth's online networks: A multilevel approach. In *Computers & Education*, *68*: 522–533. doi:10.1016/j.compedu.2013.06.007.

Vacca, R., Solano, G., Lubbers, M.J., Molina, J.L. and McCarty, C. (2017). A personal network approach to the study of immigrant structural assimilation and transnationalism. In *Social Networks*. doi:10.1016/j.socnet.2016.08.007.

Vacca, R., Stacciarini, J.-M. R. and Tranmer, M. (Under review). Using cross-classified multilevel models to analyze tie variation in overlapping personal networks.

Valente, T.W. and Vlahov, D. (2001). Selective risk taking among needle exchange participants: implications for supplemental interventions. In *American Journal of Public Health*, *91*(3): 406–411.

Valente, T.W., Watkins, S.C., Jato, M.N., Van Der Straten, A. and Tsitsol, L.-P. M. (1997). Social network associations with contraceptive use among Cameroonian women in voluntary associations. In *Social Science & Medicine*, *45*(5): 677–687. doi:10.1016/S0277-9536(96)00385-1.

Van Der Gaag, M. and Snijders, T.A.B. (2005). The Resource Generator: Social capital quantification with concrete items. In *Social Networks*, *27*(1): 1–29. doi:10.1016/j.socnet.2004.10.001.

van Duijn, M.A.J. (2013). Multilevel modeling of social network and relational data. In J.S. Simonoff, M.A. Scott and B.D. Marx (Eds.), *The SAGE Handbook of Multilevel Modeling* (pp. 599–618). Los Angeles: SAGE.

van Duijn, M.A.J., Snijders, T. and Zijlstra, B.J.H. (2004). p2: a random effects model with covariates for directed graphs. In *Statistica Neerlandica*, *58*(2): 234–254. doi:10.1046/j.0039-0402.2003.00258.x.

van Duijn, M.A.J., van Busschbach, J.T. and Snijders, T. (1999). Multilevel analysis of personal networks as dependent variables. In *Social Networks*, *21*(2), 187–210. doi:10.1016/S0378-8733(99)00009-X.

van Tubergen, F. (2015). Ethnic boundaries in core discussion networks: A multilevel social network study of Turks and Moroccans in the Netherlands. In *Journal of Ethnic and Migration Studies*, *41*(1): 101–116. doi:10.1080/1369183X.2014.886955.

Völker, B. and Flap, H. (2001). Weak ties as a liability. The case of East Germany. In *Rationality and Society*, *13*(4): 397–428. doi:10.1177/104346301013004001.

Völker, B. and Flap, H. (2007). Sixteen million neighbors. A multilevel study of the role of neighbors in the personal networks of the Dutch. In *Urban Affairs Review*, *43*(2): 256–284. doi:10.1177/1078087407302001.

Waes, S.V., Bossche, P.V. den, Moolenaar, N.M., Maeyer, S.D. and Petegem, P.V. (2015). Know-who? Linking faculty's networks to stages of instructional development. In *Higher Education*, *70*(5): 807–826. doi:10.1007/s10734-015-9868-8.

Walker, M.E., Wasserman, S. and Wellman, B. (1993). Statistical models for social support networks. In *Sociological Methods & Research*, *22*(1): 71–98. doi:10.1177/0049124193022001004.

Wasserman, S. and Faust, K. (1994). *Social Network Analysis: Methods and Applications*. New York: Cambridge University Press.

Wellman, B. (1979). The Community Question: The intimate networks of East Yorkers. In *American Journal of Sociology*, *84*(5): 1201–1231.

Wellman, B. (2007). The network is personal: Introduction to a special issue of Social Networks. In Social Networks, 29(3): 349–356. doi:10.1016/j.socnet.2007.01.006.

Wellman, B. and Frank, K.A. (2001). Network capital in a multi-level world: Getting support from personal communities. In N. Lin, K.S. Cook and R.S. Burt (Eds.), *Social Capital: Theory and Research*. 233-273: Transaction Publishers.

Wellman, B. and Wortley, S. (1990). Different strokes from different folks: community ties and social support. In *The American Journal of Sociology*, *96*(3): 558–588.