# Book Review: Newman: Kenny, D. A., Kashy, D. A., & Cook, W. L. (2006). Dyadic Data Analysis. New York: Guilford

Daniel A. Newman

*Organizational Research Methods* 2010 13: 603 originally published online 5 September 2008
DOI: 10.1177/1094428108324689

Organizational Research Methods
13(3) 603-607

http://orm.sagepub.com
SAGE

# Book Review: Newman

Kenny, D. A., Kashy, D. A., & Cook, W. L. (2006).
*Dyadic Data Analysis*. New York: Guilford.

**Reviewed by:** Daniel A. Newman, *University of Illinois at Urbana-Champaign*
DOI: 10.1177/1094428108324689

Kenny, Kashy, and Cook are helping usher in a paradigm shift for the field of psychology—toward the *interpersonal paradigm*—and this particular book summarizes major strides on the road to appropriate data analysis within this rising movement. Although the book's title rightly implies a focus on dyadic data, the authors go to great lengths describing the fundamental statistical techniques that underlie many of the proposed analytic strategies—including surprisingly thorough reviews of multilevel modeling (chapter 4), structural equation modeling (SEM; chapter 5), and introductory social network analysis (chapter 11).

The authors present three major designs for dyadic studies: (a) the standard dyadic design (involving samples of mutually exclusive couples, typically found in research on dating behavior and marital satisfaction), (b) the social relations model (SRM) design (involving groups of three or more individuals who are all interrelated, found in studies of person perception and of family systems), and (c) the one-with-many design (involving nested designs with multiple partners per focal actor). For each of these designs, dyad or unit members can be either *distinguishable* (e.g., husband–wife couples, with members distinguished by gender; families of mother, father, older sibling, and younger sibling, with members distinguished by family role) or can be *indistinguishable* (e.g., homosexual couples, who cannot be distinguished by gender; workgroups with undifferentiated member roles). Distinguishability is a key feature of dyadic designs, and separate analytic strategies are presented throughout the book for distinguishable versus indistinguishable designs.

## The Standard Dyadic Design

Chapters 1 through 7 are rich with detail about how to conduct analyses on the standard dyadic design. As such, these chapters should be required reading for anyone wanting to study couples (spouses, roommates, supervisor–subordinate pairs). The issue of *nonindependence* in the outcome variables is highlighted throughout. A classic example of nonindependence is the finding that two spouses' reports of marital satisfaction are correlated, on account of their membership in the same marriage. The authors explain how to estimate nonindependence for distinguishable designs (using the partial Pearson correlation for interval-level outcomes and Cohen's kappa for categorical outcomes) and for indistinguishable designs (using intraclass correlation; ICC). Page 45 gives an excellent discussion of the effects of ignoring outcome variable nonindependence (both positive and negative nonindependence) on tests of statistical significance, both when the independent variable (IV) is *between dyads* (e.g., length of marriage, dyad communication quality) and when the IV is *within dyads* (e.g., gender, giver vs. receiver role). The authors clarify when statistical conclusions will be either too liberal or too conservative as a result of ignoring nonindependence, both mirroring and extending the work of Bliese and Hanges (2004).

Kenny and colleagues effectively recommend a three-step process for handling nonindependence: (a) conduct a power analysis to determine how many dyads are needed to demonstrate statistically significant nonindependence, (b) test for nonindependence in the DV while *controlling for the IVs* and using a cutoff of $p < .20$, then (c) proceed with the analysis, using the result of the empirical test of nonindependence to determine whether to analyze the DV at the individual versus group level of analysis. When contrasted to contemporary organizational research methods used for multilevel designs (cf. Bliese, 2000), Kenny et al.'s strategy is interesting. If the Kenny et al. strategy for assessing nonindependence were followed in research on organizational climate; for instance, it would be akin to calculating ICCs for the climate scale *only after variance attributable to the antecedents of climate had been removed* (cf. Bliese, 2000; George & James, 1993). Furthermore, the Kenny et al. approach would imply that, in circumstances where ICCs are too small to justify aggregating individual-level climate perceptions to the group level, one could then proceed with the analysis at the individual level (because nonindependence had not been established). Thus, there is a potentially important set of distinctions between how organizational researchers view nonindependence (i.e., as a *hurdle* to be cleared before *single variables—e.g., safety climate—conceptualized a priori at the group level* can be aggregated to represent the group) versus how social psychologists view nonindependence (i.e., as an *litmus test* of whether *residualized variables* should rightly be assigned to the group level vs. individual level, *on empirical grounds*). For organizational researchers, the a priori conceptual definition of the group-level construct is paramount, and in the absence of empirically demonstrated nonindependence (e.g., both ICC $> .05$ and $p < .05$) the study cannot go forward. For Kenny and colleagues, by contrast, the empirical test of nonindependence is paramount, and the conceptual definition and hypothesis tests then follow at the individual or group level of analysis, as determined by an empirical criterion ($p < .20$). It might be useful for organizational and social psychologists to debate the relative merits of these two strategies for multilevel analysis in future dialogues.

Chapters 1 through 5 provide all the basic instructions and examples one would need to conduct analyses of the effects of between-dyads and within-dyads IVs on outcomes measured in dyads. These chapters cover data layout; testing and interpreting nonindependence; applying ANOVA or regression, multilevel modeling, and SEM for these tests (including when to use each analysis, along with SAS, SPSS, and HLM syntax for each); and real-data examples that the reader can use to practice the techniques. In writing these how-to sections, the authors took great care to provide enough detail for the reader to really understand the choices being made. As examples of the level of detail, they explain the distinction between maximum likelihood (ML) and restricted ML estimation, the reason why multilevel modeling with dyads requires that slopes be constrained equal or fixed across all pairs (mentioning that this constraint does not bias the test of intercepts-as-outcomes), how to analyze data with negative nonindependence (where nonindependence cannot be specified as a variance), how to adjust SPSS outputs so they give the same $p$ values as other software, and how to impose invariance constraints across dyad members (including factor loadings, variances, errors, and intercepts) when conducting dyadic SEM.

When it comes to the descriptions of SEM given in the book, I had a few (mostly minor) issues. First, the authors recommend that researchers conducting confirmatory factor analysis (CFA) use correlated errors between pairs of indicators if those indicators share a lot of variance; my preference is to specify hierarchical subfactors in this case. Second, the authors recommend post hoc respecifications of SEM models following the discovery of poor model fit or large modification indices (pp. 106, 233), yet they do not adequately clarify that such modified models must be fit to an entirely new data set to avoid capitalizing on chance (MacCallum, Roznowski, & Necowitz, 1992). Third and finally, chi-square difference tests are regularly proposed for comparing the fits of nested models. Although the authors note that chi-square tests are sensitive to sample size, the extreme sensitivity of chi-square tests (and chi-square difference tests) to sample size makes them relatively uninformative in this instance. Using a practical fit index for model comparisons (e.g., CFI) might be preferable.

Related to this, chapter 6 describes an empirical "omnibus test of distinguishability," designed to determine whether dyad members are distinguishable. This statistical test (p. 129) is proposed for determining whether dyad members in theoretically distinct roles (e.g., husbands and wives) are empirically distinguishable, according to the data at hand. (This test is emphasized, as the authors even state in chapter 15 that failure to empirically test for distinguishability is "Sin #1" of the "Seven Deadly Sins of Dyadic Data Analysis" [p. 422].) The distinguishability test is based on a chi-square difference test, so results of this test are largely a function of sample size. The upshot of this is that, when the researcher has fewer dyads to analyze, chi-square will be small, and the researcher will falsely conclude that dyad members (husbands and wives) are not distinguishable, even when in truth they are. Drawing an incorrect inference of indistinguishability (e.g., a Type II error) is not ideal because many of the analytic procedures proposed in the book (including dyadic CFA and SEM) are much better suited to (and more robust for) distinguishable dyads, compared with indistinguishable dyads. And speaking practically, distinguishable dyads are much easier to analyze, too. In brief, I believe that either (a) universal application of the empirical test for distinguishability is premature, or (b) the omnibus test for distinguishability should be based on change in CFI (e.g., change in CFI > .01) rather than change in chi-square, or (c) a formal power analysis of the test for distinguishability must reveal adequate power before this test becomes appropriate. With inadequate power, my default assumption would be that husbands and wives are distinguishable. As such, I currently believe that within-dyad effects (e.g., husbands are more satisfied than wives), when statistically significant, can be interpreted even in the absence of a significant omnibus test of distinguishability.

In chapter 7, the authors introduce an exciting model that has come to be known as the actor–partner interdependence model (APIM). Currently, SEM is the most straightforward method for estimating the APIM with distinguishable dyads, whereas multilevel modeling is recommended for estimating the APIM with indistinguishable dyads. An easy way to think of the APIM model involves two exogenous variables (husband's salary and wife's salary) and two endogenous variables (husband's marital satisfaction and wife's marital satisfaction). The influence of wife's salary on wife's marital satisfaction is called an *actor effect*, whereas the influence of husband's salary on wife's marital satisfaction is called a *partner effect*. In the vast majority of individual-level models studied in psychology, partner effects are assumed to be zero (i.e., they are ignored), which results in biased estimates (usually overestimates) of actor effects. A series of constraints can be imposed on the APIM model to test whether (a) actor effects are equal in magnitude to partner effects (e.g., marital satisfaction is driven equally by one's own and one's spouse's salaries—called the couple-oriented model) or (b) actor effects are equal in magnitude, but opposite in sign, to partner effects (e.g., marital satisfaction is driven positively by one's own salary but negatively by one's spouse's salary—called the social comparison model). Discussion of the APIM model is extended to discuss how one can use it to test actor–partner interaction effects or how partner interaction effects can be used to calibrate measures of dyadic relationships (i.e., to find their true zero point, or the level of dyadic closeness at which predicted partner effects become zero). My only comment on this chapter is that future researchers who are interested in testing partner similarity effects (for which the authors recommend using the absolute difference between actor and partner rather than the product of actor and partner scores; p. 166) should read one of the discussions of difference score modeling given by Edwards (e.g., Edwards, 2001).

## SRM

Discussion of social relations designs begins on page 185, with chapter 8. The SRM is an analysis of interpersonal ratings (e.g., liking, attachment) that can be estimated when members of two groups rate each other (block design) or in a round-robin design where every group member rates every other group member. The SRM is basically a two-way random-effects ANOVA, in which the ratings from a round-robin (or block) design are partitioned into actor main effects (e.g., an actor's general tendency to give high or low ratings), partner main effects (e.g., a partner's general tendency to receive high or low ratings), and relationship effects (actor–partner interaction effect, which is often confounded with error). An example of the relationship effect (a dyad-level effect) is the extent to which actor "o" rates partner "p" highly, controlling for o's actor effect and p's partner effect. The purpose of SRM is to calculate the percentage of variance in a relation (e.g., liking, interpersonal disclosure, trait ratings) that is attributable to the actor, the partner, or the actor–partner interaction. It is also possible to estimate the correlation between actor effects and partner effects (to assess *generalized reciprocity*—individuals who rate others highly are also rated highly by others) or between an individual difference variable (e.g., extroversion) and actor or partner effects. SRM analysis requires at least four members per group, unless some very restrictive assumptions are made to permit three-member groups (see p. 251).

Unlike the description of analyses for the standard dyadic design (chapters 1 to 7), the description of SRM analyses does not provide all the detail one would need to conduct the analyses. Instead, several of the key steps are described and key equations are given, but the reader is referred to Kenny's online software to conduct the actual analyses.

Chapter 9 discusses the SRM for groups with distinguishable members (e.g., families, groups with roles) and provides instructions for how to cleverly partition the variance using CFA (where each person's rating of another person is allowed to triple-load on an overall family factor as well as the corresponding actor factor and partner factor). The unique variance in a rating (indicator variable) represents the relationship variance, whereas the variances in the latent actor factors and partner factors represent actor and partner variances, respectively. Covariances between actor and partner factors estimate generalized reciprocity, whereas correlated uniquenesses (e.g., between the two indicators: mother–father rating and father–mother rating) are used to estimate dyadic reciprocity (e.g., the extent to which mother and father mutually like each other, independent of individual characteristics). The authors again recommend a chi-square difference test for model comparisons.

## The One-With-Many Design

Chapter 10 covers the one-with-many design, in which a focal person (leader, mother, customer service agent) is linked with several partners (followers, children, customers). Data for this design can be collected from either source (focal person, partners) or from both sources; but analyses must necessarily account for the fact that partners are nested within focal persons. The effects of predictor variables (both focal person-level and partner-level predictors) can be assessed via multilevel modeling (with indistinguishable partners) or via SEM (with distinguishable partners, with each partner role yielding a separate indicator variable). The specifics for conducting these analyses are clearly presented, using a level of detail not available in other published sources.

## Social Networks

Dyadic data analysis represents nothing more than a series of special cases of social network analysis. Kenny and colleagues note this implicitly when observing (a) a one-with-many design is essentially a set of egocentric networks (but usually ignoring ties between partners), (b) data from a round-robin design resembles a sociomatrix (network matrix), and (c) an SRM is similar to the social network $p_1$ model (Holland & Leinhardt, 1981; although SRM requires interval-level data, permits actor–partner correlations, and is a random-effects model). For a more recent summary of social network exponential random graph models (i.e., the $p^*$ model), I recommend Robins, Pattison, Kalish, and Lusher (2007).

Eventually, it appears that social network $p^*$ models and SRM will subsume one another, with the development of more general models for the emergence of interpersonal structures (e.g., to test dyadic reciprocity, triadic transitivity, etc.). This future integration is insightfully foreshadowed by Kenny et al. on page 313, where they derive SRM actor and partner effects in terms of actor indegree, outdegree, and network density. In the future, models that can robustly integrate features of network structure with individual difference variables will be the sine qua non of the interpersonal paradigm, although for the present time researchers will have to take satisfaction in the constrained models that are currently available.

## Conclusion

This book covers a lot of material (more than 450 pages) and does so very well. I had the sense that perhaps it could have been presented as two books: one on the standard dyadic design (chapters 1 through 7, 13, and 14) and another on SRM (chapters 8, 9, and 11). As such, readers are basically getting two books for the price of one!!

Writing in the clear, meticulous, and impactful style for which Kenny has become well-known, Kenny, Kashy, and Cook lead the reader step-by-step through the various data structures, analyses, assumptions, and current limitations of their techniques. Contrary to many modern psychology books, they demonstrate that including equations within the text can often make things *easier* to understand. The book succeeds in providing something of a one-stop-shop for dyadic data analysis, and the use of clear explanations and examples of every step in the analyses makes this work indispensible for the dyadic researcher. Furthermore, this is not only a "how-to" book, it is also a "why-to" book that carefully gives the rationale for and assumptions of each analytic decision. If you want to know the state-of-the-art methods for analyzing dyadic data (particularly for studying couples and SRM), then this book is a must-have.

### References

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

Bliese, P. D., & Hanges, P. J. (2004). Being both too liberal and too conservative: The perils of treating grouped data as though they were independent. *Organizational Research Methods*, *7*, 400-417.

Edwards, J. R. (2001). Ten difference score myths. *Organizational Research Methods*, *4*, 265-287.

George, J. M., & James, L. R. (1993). Personality, affect, and behavior in groups revisited: Comment on aggregation, levels of analysis, and a recent application of within and between analysis. *Journal of Applied Psychology*, *78*, 798-804.

Holland, P. W., & Leinhardt, S. (1981). An exponential family of probability distributions for directed graphs. *Journal of the American Statistical Association*, *76*, 33-50.

MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992). Model modifications in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin*, *111*, 490-504.

Robins, G., Pattison, P., Kalish, Y., & Lusher, D. (2007). An introduction to exponential random graph (p*) models for social networks. *Social Networks*, *29*, 173-191.