Article

Sociological Methods & Research
2016, Vol. 45(2) 191-230

Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0049124115578030
smr.sagepub.com

Measuring Collective Efficacy: A Multilevel Measurement Model for Nested Data

Ross L. Matsueda[1] and Kevin M. Drakulich[2]

Abstract

This article specifies a multilevel measurement model for survey response when data are nested. The model includes a test–retest model of reliability, a confirmatory factor model of interitem reliability with item-specific bias effects, an individual-level model of the biasing effects due to respondent characteristics, and a neighborhood-level model of construct validity. We apply this model for measuring informal social control within collective efficacy theory. Estimating the model on 3,260 respondents nested within 123 Seattle neighborhoods, we find that measures of informal control show reasonable test–retest and interitem reliability. We find support for the hypothesis that respondents' assessments of whether their neighbors would intervene in specific child deviant acts are related to whether they have observed such acts in the past, which is consistent with a cognitive model of survey response. Finally, we find that, when proper measurement models are not used, the effects of some neighborhood covariates on informal control are biased upward and the effect of informal social control on violence is biased downward.

[1] Department of Sociology, University of Washington, Seattle, WA, USA
[2] School of Criminology and Criminal Justice, Northeastern University, Boston, MA, USA

Corresponding Author:
Ross L. Matsueda, Department of Sociology, University of Washington, Box 353340, Seattle, WA 98195, USA.
Email: matsueda@uw.edu

Keywords

structural equation model, measurement model, nested data, ordinal indicators, multilevel model, collective efficacy, informal social control

A critical issue in the assessment of social science theories concerns measurement: Can key concepts be operationalized clearly? Can empirical measures of concepts be constructed? Do such measures perform up to standards empirically? Unless a theory meets accepted scientific standards of measurement, the theory cannot be subjected to empirical test, cannot be verified, and therefore, is unlikely to be accepted by the scientific community. Recent attention to measurement in the social sciences covers a range of substantive and theoretical topics, including behavioral expectations (e.g., Manski 2004), social capital (e.g., Paxton 2002), democracy (Coppedge et al. 2011), employment careers (Manzoni et al. (2010), time use data from diaries (Kan and Pudney 2008), and family activities (Broege et al. 2007). In criminology, the measurement of delinquent behavior (e.g., Osgood, McMorris, and Potenza 2002; Piquero, MacIntosh, and Hickman 2002) and self-control (e.g., Piquero, MacIntosh, and Hickman 2000) has been examined using models for survey response. While measurement has been discussed in detail for individual-level psychometric models of test scores and survey responses, it is no less an issue for models of macro-ecological units, such as neighborhoods, communities, and schools. Indeed, Raudenbush and Sampson (1999) have advocated for "ecometrics," the scientific study of ecological measures taken from surveys and observational data. Thus, when data are nested, distinct measurement issues arise at multiple levels of analysis.

We specify a multilevel measurement model for nested survey data in which survey items are nested within individuals, which are in turn nested within neighborhoods. This allows us to examine unreliability due to random measurement error at three levels: test–retest unreliability for repeated measures of one indicator, interitem unreliability for multiple indicators, and between-neighborhood unreliability. Using a cognitive theory of survey response as a framework, we also test for systematic measurement error at the level of the individual and neighborhood. Our model begins with a two-level factor model for ordinal indicators estimated within a structural equation model (SEM) framework, using a threshold model to link ordinal measures to latent continuous indicators, computing scale-appropriate polychoric correlations, and using weighted least squares (WLS) to estimate measurement parameters. Note that this model is equivalent to an item response theory (IRT) model for ordered variables (see Takane and de Leeuw 1987).[1] We then specify a neighborhood-level response model that treats respondents as informants and controls for bias due to compositional differences in informant characteristics by neighborhood. Finally, we examine the construct validity of our measures by incorporating the measurement model into a substantive model. We apply our model to the measurement of neighborhood informal social control, the key concept in collective efficacy theory, an important recent theory of neighborhood crime (Sampson 2012; Sampson, Raudenbush, and Earls 1997). We show how different assumptions about the measurement process can produce differences in substantive coefficients of interest.

Collective Efficacy and the Measurement of Informal Social Control

Collective Efficacy, Social Capital, and Informal Social Control

Social capital theory is one of the most important theoretical developments in the social sciences over the past 25 years. Empirical research on the theory was invigorated by the work of Sampson and colleagues (1997, 1999) on neighborhood social capital. Building on the tradition of Shaw and McKay's [1942] (1969) social disorganization theory, they specified a theoretical mechanism of informal neighborhood social control, linked the mechanism to neighborhood structure through the concept of social capital, and specified operational indicators of the theory's key concepts. Specifically, they merged Coleman's (1990) dimensions of social capital with Bandura's (1986, 1997) concept of "collective efficacy" and in the process redefined collective efficacy to explain neighborhood social control. The innovations made by Sampson and colleagues were to apply the concept of collective efficacy to describe neighborhood *purposive action*—rather than *mere beliefs in a group's efficacy*, as specified by Bandura (1997)—tie it to Coleman's (1990) concept of social capital, and obtain operational indicators from previous surveys of neighborhoods. For Coleman, social capital, including intergenerational closure and reciprocated exchange, is a *structural* resource that facilitates purposive action. Informal social control, based on shared expectations for action, is the task-specific *agentic* direction of purposive action that translates structural resources (social capital) into goal-directed activity—in this case, maintaining neighborhood safety. Sampson et al. (1997:918) define collective efficacy as a property of neighborhoods—namely, "the capacity of residents to control group level processes and visible signs of disorder" which helps reduce "opportunities for interpersonal crime in a neighborhood." The key to collective efficacy, then, is informal (or child-centered) social control, which is the active component that translates neighborhood social capital into safe environments.

Measurement of Informal Social Control

Informal (child-centered) social control is an unobservable theoretical construct requiring operational indicators. In their empirical studies using survey data from the Project on Human Development in Chicago Neighborhoods (PHDCN), Sampson et al. (Sampson et al. 1997; Sampson, Morenoff, and Earls 1999) measured informal control by asking respondents whether their neighbors could be counted on to "do something" if children were hanging out on a street corner, spray-painting graffiti, or showing disrespect to an adult. Taylor (1996) developed these measures previously to operationalize informal control.[2]

Sampson (1997:227) examined the measurement properties of the "collective regulation of adolescent misbehavior" using 80 neighborhood clusters of the PHDCN. He found an interitem α reliability of .79 across three measures. Because the important variation in informal control is across neighborhoods, he estimated a neighborhood-level reliability of .81, which reveals the proportion of total variance in informal control attributed to cross-neighborhood variation. Finally, he examined the construct validity of the measure, finding child social control strongly associated with neighborhood demographics, disadvantage, ethnicity/immigration, and residential stability, as well as predictive of aggregate problem behavior in the neighborhood. In his critique of Gibson et al.'s (2002) study of social integration and collective efficacy in three smaller cities, Taylor (2002) discussed the history of the concept of informal social control, raised questions about the discriminant validity of the concept, and proposed using a structural equation approach to measurement.

Strictly speaking, the PHDCN measures of informal control refer to residents' subjective *beliefs* about whether neighbors would intervene when delinquent activities arise, rather than the key concept, *objective activation* of networks to reduce crime. Therefore, Sampson et al. (1997) devised the following modeling strategy. First, they assumed that collective efficacy is an objective property of neighborhoods and not a subjective belief of residents. Second, they treated survey respondents as "informants" for their neighborhoods and treated informants' assessment of informal control as a fallible indicator of actual neighborhood informal control. Third, they recognized that informant assessments might differ systematically by informant characteristics, requiring an adjustment for the neighborhood composition of informants. For example, if on average, high-income homeowners tend to report higher informal control, then neighborhoods with more wealthy homeowners will score relatively higher on informal control then those with fewer wealthy homeowners—even when true informal control is identical for the neighborhoods.

Therefore, to overcome differences in assessments across neighborhoods due to differences in neighborhoods' composition of informants, responses need to be weighted. In the abovementioned example, the responses of wealthy homeowners would be adjusted downward and those of impoverished renters adjusted upward. In this way, measures of residents' *perceptions* or assessments of whether neighbors will intervene when problems arise can be used as measures of *objective* informal social control. Sampson et al. (1997) used a regression approach for weighting responses, regressing an index of informal control on informant characteristics to adjust each neighborhood's mean score for compositional differences in informants.

Sampson et al. (1997) found that, as expected, informal social control measures were biased by informant characteristics, including upwardly biased by homeowners, stable residents, older respondents, and higher socioeconomic status (SES) residents. Finally, they found support for the construct validity of collective efficacy: Disadvantage, immigration, and residential mobility were negatively associated with informal social control as well as with homicide and violent victimization. This approach to measuring informal social control with conventional survey items raises an important question: How do respondents know whether their neighbors will intervene when delinquency occurs in the neighborhood? To consider this question, we need a general framework to guide our understanding and modeling of survey responses.[3]

A Framework for Modeling Survey Responses: Cognitive Theories

A simple but powerful response model would focus on information processing: Respondents likely learn about their neighbors' behavior through communication and direct observation and use this information to respond to survey questions. Along these lines, survey researchers have recently moved beyond stimulus–response models of the survey response process to multistage cognitive theories (see Jobe and Hermann 1996 for a review). Tourangeau, Rips, and Rasinski (2000) developed a comprehensive cognitive model of survey response, in which respondents perform four tasks: (1) comprehend the question, (2) retrieve relevant information from memory, (3) make judgments about that information, and (4) select an appropriate answer. Respondents sometimes engage in shortcuts, rather than performing every cognitive task, but in general, most will engage in some version of each task.

Our research design and data do not allow us to test this cognitive model in a deductive way. Nevertheless, the model is useful as a general framework for thinking about the process of responding to questions of informal control as a series of cognitive stages. First, respondents must assign meaning to the question, "If some children were spray-painting graffiti on a local building, how likely is it that your neighbors would do something about it?" and identify the information sought. This question seeks information on two hypothetical events: (1) children spray-painting graffiti on a local building and (2) neighbors doing something about it. This information is concrete and straightforward, unlikely to produce misunderstandings.

Second, respondents must retrieve the information from long-term memory by adopting a strategy of retrieval, identifying cues that would trigger recall, and collecting disparate memories into a whole. Memory theorists have found that retrieval from long-term memory entails activation of a cue, which then spreads along lines such as temporal, taxonomic, or part whole (e.g., Anderson 1983). Here, respondents likely try to remember specific occasions in which children were spray-painting graffiti in the street (cue activation) and recall whether neighbors intervened or not (cue spread). Respondents may also learn about the actions of neighbors by talking with their neighbors, particularly about crime and other problems in the neighborhood, although such third-party retrieval is often less reliable (Tourangeau et al. 2000). Retrieval would consist of recalling the nature and veracity of these conversations.

Third, respondents must make a judgment about whether their retrieval is complete or incomplete; if incomplete, they may try to retrieve more information or fill in the gaps. Often if retrieval is difficult, incomplete, or sketchy, respondents will assume the events were rare or never took place (Tourangeau et al. 2000), a heuristic device termed "availability" by Tversky and Kahneman (1973). For example, if the respondent is uncertain of whether neighbors would intervene, they might assume they would rarely or never do so. Moreover, if the respondent does not recall ever witnessing children spray-painting graffiti, they might conclude that their neighbors successfully intervened in the past, a response effect that may result in overstating informal control. Furthermore, lacking other specific forms of information about neighbors doing something about graffiti, respondents may infer from a low crime rate or kempt residences that neighbors will intervene. On this point, St. Jean (2007), using qualitative interviews, found that affluent residents simply made assumptions about their neighbors, based on their socioeconomic characteristics. For example, if new neighbors appeared professional and kept their yard and lawn well-manicured, affluent residents would assume not only that the neighbors are trustworthy, but also that they would intervene when problems arose in the neighborhood (presumably to protect their new investments in the neighborhood).

Fourth, respondents must select a response category, including mapping their response to Likert-type scales, such as "very likely," "likely," "unlikely," and "very unlikely," and possibly editing their initial response (Tourangeau et al. 2000). These categories are clearly ordered; however, the distance between categories is not known. For this reason, we treat them as ordered categories and estimate the distances between categories empirically with a threshold model.

This cognitive model provides a way of conceptualizing survey responses and also allows us to specify hypotheses about the substantive sources of responses to informal social control and the potential biasing sources of response error. Such response effects may vary by characteristics of the respondent. For example, residents who are younger and relatively new to the neighborhood may have fewer memories of neighbors intervening in child deviance and therefore infer that such behavior is rare. In contrast, homeowners and parents may be sensitive to child deviance, may be quick to intervene themselves, and consequently may easily recall other instances of neighbors intervening.

A Multilevel Measurement Model for Nested Data

We estimate a four-level hierarchical linear model (HLM) of informal social control. We begin with a second-order confirmatory factor measurement model for ordinal indicators that estimates the test–retest reliability of one measure of informal control. Our confirmatory factor model of interitem reliability tests hypotheses about the sources of respondents' answers to informal control items and relaxes the assumption of τ-equivalence made by previous studies. Our third- and fourth-level models specify an individual-level model of bias in reports of neighborhood informal control due to respondent characteristics (see Raudenbush and Sampson 1999; Sampson et al. 1997), and a neighborhood-level model of neighborhood informal control, which allows us to examine construct validity.

In general, we use more refined estimates of test–retest and interitem reliability than the usual Cronbach's (1951) α, which has been subject to recent criticism. α gives biased estimates of reliability when the assumption of essential τ-equivalence (equal measurement slopes) among items does not hold, when items are ordinal rather than continuous, and when correlated measurement errors are present (e.g., Raykov 2001). Typically, the result is an underestimate of reliability, leading researchers to treat α as a lower bound. Note, however, that correlated errors can lead to overestimates of reliability (Bentler 2009). We follow recommendations of critics of α and use a model-based approach to obtain reliability coefficients for individual indicators (e.g., Bollen 1989; Jöreskog 1971) and scales (e.g., Green and Yang 2009).[4] This allows us to examine three forms of unreliability: (1) unreliability arising from test–retest discrepancies, (2) unreliability arising from internal inconsistency across items, and (3) unreliability arising from inter-coder unreliability across informants within neighborhoods. Because our indicators of informal social control are measured on ordinal scales, we use a structural equation approach for estimation of measurement models with ordinal indicators (e.g., Bentler 2009). The remainder of this section describes our four-level model and presents key measurement hypotheses.

Threshold Model for Ordinal Indicators

We begin with a threshold model that relates observed indicator Y_{hijk} (with four ordinal categories) for the hth repeated measure of the ith indicator of the jth person in the kth neighborhood to a latent continuous variable, Y^*:

$$Y_{hijk} = 1(\text{very likely}) \text{ if } {Y_{hijk}}^* \leq \alpha_1,$$

$$Y_{hijk} = 2(\text{likely}) \text{ if } \alpha_1 < {Y_{hijk}}^* \leq \alpha_2,$$

$$Y_{hijk} = 3(\text{unlikely}) \text{ if } \alpha_2 < {Y_{hijk}}^* \leq \alpha_3,$$

$$Y_{hijk} = 4(\text{very unlikely}) \text{ if } \alpha_3 < {Y_{hijk}}^*.$$

To estimate the thresholds, we must specify a distribution of Y^*. If we assume that Y^* is normally distributed, we can estimate the thresholds as follows:

$$\alpha_m = \Phi^{-1}\left(\sum_{p=1}^{m} n_p/N\right), \quad m = 1, 2, 3, \tag{1}$$

where Φ^{-1} is the inverse of the standard normal distribution function and n_p is the sample size of the pth category, where $N = n_1 + n_2 + n_3 + n_4$. Consider a second ordinal variable X, also with four categories, and related to a continuous normally distributed variable X^* with thresholds:

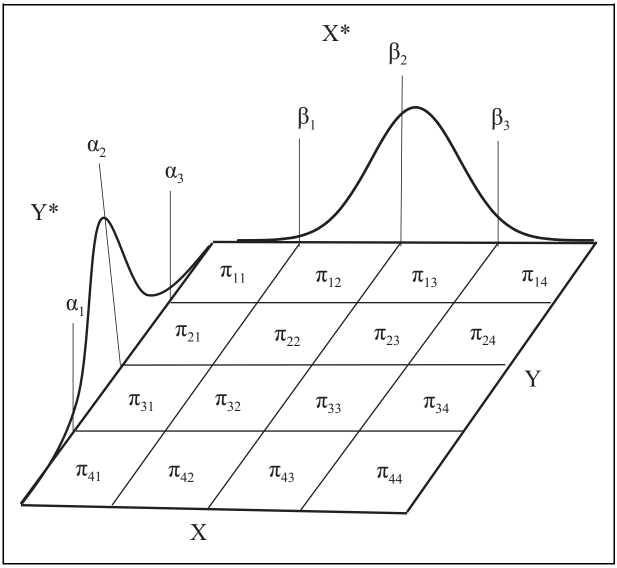

Figure 1. Cross-classified ordinal variables generated by two bivariate normal marginal distributions via thresholds.

$$\beta_l = \Phi^{-1}(\sum_{q=1}^{m} n_q/N), \quad l = 1, 2, 3. \tag{2}$$

Figure 1 relates a cross-classification of two ordinal variables, *X** and *Y**, to two underlying variables with Gaussian marginal distributions (*X* and *Y*). Once the thresholds are estimated, the polychoric correlation ρ, the scale-appropriate correlation between two ordinal variables can be estimated with maximum likelihood (ML). If π_{ml} is the probability of an observation falling into cell (m, l), the log likelihood of the sample is given by:

$$\log L(n_{ml}|\pi_{ml}) = c \sum_{m=1}^{r} \sum_{l=1}^{s} n_{ml} \log \pi_{ml}, \tag{3}$$

where $\pi_{ml} = \Phi_2(\alpha_m, \beta_l) - \Phi_2(\alpha_{m-1}, \beta_l) - \Phi_2(\alpha_m, \beta_{l-1}) - \Phi_2(\alpha_{m-1}, \beta_{l-1})$, Φ_2 is the bivariate normal distribution function with population correlation ρ, and c is a constant (Olsson 1979). Differentiating log L with respect to ρ, setting the result to zero, and solving for ρ yields a ML estimate, $\hat{\rho}_{ML}$. The covariance matrix of the estimate is obtained by taking the expected value of the negative of the inverted matrix of second-order partial derivatives of log L with respect to ρ. We can generalize this to the multivariate case of t variables and the $\frac{1}{2}t^2$ polychoric correlations (ρs) can be placed in a matrix R, with population covariance matrix of estimated ρs, $\Sigma_{\hat{\rho}\hat{\rho}}$, which is of order $\frac{1}{2}t^2 \times \frac{1}{2}t^2$. The covariance matrix of $\hat{\rho}_{ML}$ estimated from the sample is $S_{\hat{\rho}\hat{\rho}}$ (Poon and Lee 1987). We use Jöreskog and Sörbom's (2002) *PRELIS 2.80* program to obtain ML estimates of R and $S_{\hat{\rho}\hat{\rho}}$.

We now discuss our multilevel measurement model from the lowest level to the highest: (1) factor model of repeated measures of one measure of informal control, (2) factor model of informal control, (3) individual-level model of informal control adjusting for bias due to informant characteristics, and (4) neighborhood-level models of construct validity of informal control.

Level 1 Measurement Model for Repeated Measures

Our survey obtained repeated measures of one indicator of informal social control, "If children were fighting out in the street, how likely is it that people in your neighborhood would stop it?" The question was asked near the beginning of the interview with other informal social control items and then again at the end of the interview. We can estimate item-specific reliability within a test–retest framework with the following level-1 model:

$$Y_{hijk} = \lambda_h^T T_{ijk} + \delta_{hijk} \quad h = 1, 2 \quad i = 1, \tag{4}$$

where $h = 1, 2$ means there are two repeated measures of one $i = 1$ of the indicators of informal social control, λ_h^T is a metric slope (one of which is normalized to unity), T_{ijk} is a latent variable (a "true score") capturing an individual's assessment of the item after purging random error due to test–retest unreliability, and δ_{hijk} is a random measurement error term assumed normally distributed, heteroscedastic across h, and otherwise, *iid*. Following Bollen (1989), we use an indicator-specific measure of reliability based on the squared standardized slope, which gives the true score variance divided by the observed item variance:[5]

$$P^2_{Y_{h,i=1}T} = \frac{\sigma^2_{T_{i=1}}}{\sigma^2_{Y_{h,i=1}}} \left(\lambda^T_{h,i=1}\right)^2. \tag{5}$$

Equation (5) allows us to test the hypothesis that test–retest reliability on one of our indicators of informal control is reasonably high:

Hypothesis 1: Test–retest reliability is relatively high.

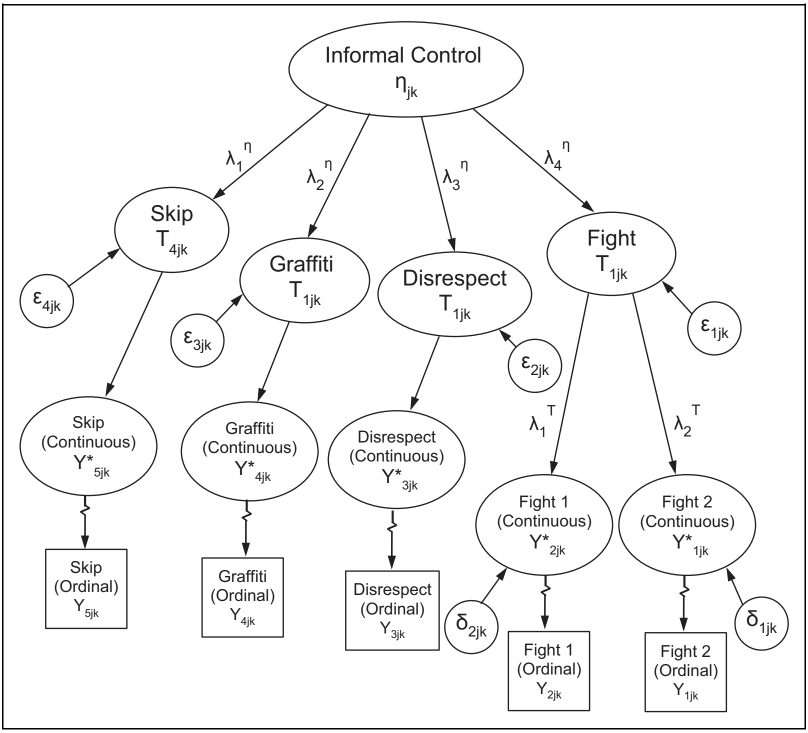

Figure 2. Two-level factor model of test–retest reliability of "fighting in the street" T_{1jk} and a second-order factor model of informal control η_{jk} (Jagged arrows denote nonlinear threshold effects for ordinal indicators).

Level 2 Measurement Model for Indicators of Informal Social Control

With multiple indicators of informal social control, we can examine inter-item reliability of the construct. A simple model is specified as follows:

$$T_{ijk} = \lambda_i^{\eta} \eta_{jk} + \varepsilon_{ijk} \quad i = 1, \ldots, 4 \tag{6}$$

where T_{ijk} is the latent variable from our level-1 model for $i = 1$, whereas for $i = 2,3,4$, it is the observed indicator; λ_i^{η} is a metric slope; η_{jk} is a latent variable (a "true score") capturing informal social control after purging interitem unreliability; and ε_{ijk} is a random measurement error term assumed normally distributed, heteroscedastic across i, and otherwise, *iid* (see Figure 2). Here, we assume that informal control is a multidimensional concept with multiple domains of meaning. We treat our indicators as if they were randomly sampled from an infinite domain of potential measures, and the resulting sampling error is picked by the measurement errors (e.g., Nunnally 1967). Note that for $i = 1$, the level-1 repeated measures allow us to disentangle random measurement error due to test–retest unreliability ($\delta_{h,i=1,jk}$) from random measurement error due to inter-item unreliability ($\varepsilon_{i=1,jk}$; see Figure 2). For $i = 2, 3, 4$, ε_{ijk} pools the two sources of error. In either case, reliability is as follows:

$$P^2_{T_i\eta} = \frac{\sigma^2_\eta}{\sigma^2_{T_i}} (\lambda^\eta_i)^2. \tag{7}$$

This allows us to test hypotheses about interitem reliability:

Hypothesis 2: Interitem reliability of informal social control is relatively high.

We complicate our level-2 model to investigate hypotheses about how informants arrive at their judgments about whether neighbors will intervene when child-related problems arise. In the cognitive process of retrieving memories and judging those memories, when asked about whether neighbors would intervene in a hypothetical problem, informants may first try to recall whether they have observed the problem in their neighborhood. If they have not observed the problem, such as "spray-painting graffiti," they may be more likely to conclude that neighbors *would* intervene, since that may explain the absence of graffiti; if they have observed the problem, they may be more likely to conclude that neighbors *would not* intervene. Thus, responses to informal social control items may reflect actual child deviance in the neighborhood, rather than objective neighborhood intervention. This possibility has far-reaching implications: If measures of informal control are entirely due to actual child deviance, collective efficacy theory would be vulnerable to the same criticisms of tautology that plagued social disorganization theory. To test this hypothesis, for each item, such as "spray-painting graffiti," we asked respondents whether they had observed the child behavior in their neighborhood. We can model this process as follows (see Figure 3):

$$T_{ijk} = \lambda^\eta_i \eta_{jk} + \gamma_i X_{ijk} + \varepsilon_{ijk} \quad i = 1, \ldots, 4, \tag{8}$$

where X_{ijk} is the informant's observation of the specific child deviance corresponding to the ith informal control indicator, and γ_i is a regression coefficient. This allows us to test the following hypothesis:

Hypothesis 3: Responses to informal social control items are a function of respondents' observation of child deviance.

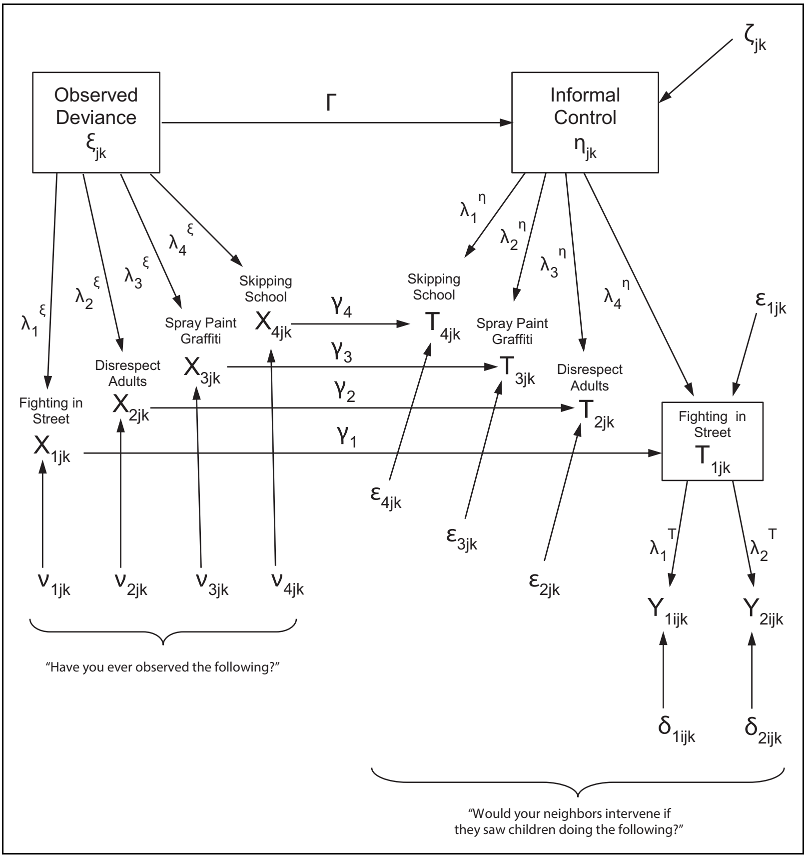

Figure 3. Two-level factor model of test–retest reliability in "fighting in the street" T_{1jk} and a second order factor model of informal control η_{jk}, with direct effects of observed deviance X_{ijk} on indicators of informal control T_{ijk}, and an observed deviance factor. (Nonlinear threshold effects for ordinal indicators omitted for simplicity.)

To parallel our factor model of informal control, we treat child deviance items as indicators of a latent variable capturing child deviance according to a domain sampling model:

$$X_{ijk} = \lambda_i^{\xi}\xi_{jk} + \upsilon_{ijk} \quad i = 1, \ldots, 4, \tag{9}$$

where λ_i^{ξ} is a metric slope, ξ_{jk} is a latent variable of child deviance purged of measurement error, and υ_{ijk} is a random measurement error term assumed normally distributed, heteroscedastic across i, and otherwise *iid.* Then, we can also regress the "true scores" of informal control on child deviance:

$$\eta_{jk} = \Gamma\xi_{jk} + \zeta_{jk}, \tag{10}$$

where Γ is a regression coefficient and ζ_{jk} is a disturbance term assumed normally distributed and *iid* (see Figure 3).

Our level-1 and level-2 models can be written as an SEM in which the two repeated measures comprise a first-order factor model for control of fighting (equation [4]), and then the fighting factor comprises an indicator (along with the three other informal control observable measures) of a second-order factor model (equation [6]). It is straightforward to include the more complicated model specified in equations (7)–(9) into the SEM. The model can be estimated using Browne's (1984) asymptotic distribution-free generalized least squares estimator, which minimizes the following fit function:

$$F = [\hat{\rho} - \sigma(\theta)]' S_{\hat{\rho}\hat{\rho}}^{-1} [\hat{\rho} - \sigma(\theta)], \tag{11}$$

where $\hat{\rho}$ is a vector of ML estimates of the polychoric correlations, $S_{\hat{\rho}\hat{\rho}}$ is the asymptotic covariance matrix of the estimated polychoric correlations, and $\sigma(\theta)$ is a vector of parameters to be estimated. By choosing $S_{\hat{\rho}\hat{\rho}}$ as a weight matrix in the quadratic form of equation (11), we obtain a best asymptotic normal estimator (Browne 1984). Minimizing F with respect to $\sigma(\theta)$ also provides a test statistic $(n - 1)$ F, which is distributed χ^2 in large samples. The asymptotic covariance matrix of the parameter vector $\sigma(\theta)$, from which standard errors are the square roots of the diagonal elements, is as follows:

$$\Sigma_{ACOV} = 1/n\{ [\partial\sigma(\theta)/\partial\theta]' S_{\hat{\rho}\hat{\rho}}^{-1} [\partial\sigma(\theta)/\partial\theta]'\}^{-1}. \tag{12}$$

After obtaining $\hat{\rho}$ and $S_{\hat{\rho}\hat{\rho}}$ from equation (12), we fit these models using Jöreskog and Sörbom's (2001) *LISREL 8.8* program, which uses the fitting function in equation (11).

Latent variable scores can be obtained for individuals from the level-2 factor models. When latent scores, or "covariance-preserving factor scores," are used in a two-step procedure, they provide consistent estimates of regressions among latent scores (Skrondal and Laake (2001).[6] We use latent variable scores derived from Anderson and Rubin (1956) and implemented in LISREL (see Jöreskog 2000; Ten Berge et al. 1999) for child deviance and informal social control after correcting for unreliability due to test–retest discrepancies and inter-item variation. We assume that η_{ij} is a normally distributed random variable and use the latent scores in an HLM to estimate our third and fourth level models.[7]

Level 3 Regression Model for Informant Bias in Neighborhood Informal Social Control

Following Sampson et al. (1997), we use our respondents as informants about social control in the neighborhood and, using regression weights, adjust the neighborhood mean scores for the potential biasing effects of a neighborhood's composition of informants:

$$\eta_{jk} = \pi_{0k} + \sum_{q=1}^{Q} \beta_q X_{qjk} + e_{jk}, \tag{13}$$

where η_{jk} is the latent score from above, π_{0k} is the intercept for the *k*th neighborhood, X_{qjk} are q characteristics of j individuals within k neighborhoods, and e_{jk} *i*s a random disturbance assumed normally distributed and *iid.* This is a conventional two-level HLM model, depicted in Figure 4. It allows the latent true scores η_{jk} to vary systematically, via $\sum_{q=1}^{Q} \beta_q X_{qjk}$, and randomly, via e_{jk}, around their neighborhood-specific means π_{0k}. The latter is assumed a normally distributed random variable capturing neighborhood informal social control after using regression weights to adjust for individual characteristics within neighborhoods. For example, if wealthy older respondents tend to overestimate informal control compared to others, then neighborhoods with higher proportions of wealthy older respondents would have upwardly biased mean informal control. The regression weights adjust for neighborhood composition differences so that the composition of each neighborhood approximates the city as a whole. Thus, we are using the term "biased" in *relative*, not *absolute*, terms. It could be that older respondents overestimate their neighborhoods' informal control relative to other residents because they are privy to the neighborhood "true score." We would still want to adjust for compositional differences in the percentage of older residents in neighborhoods so that each neighborhood would have the same "average" bias (see Raudenbush and Sampson 1999).

From the standpoint of a cognitive theory of survey response, we would expect that respondents who have rarely observed children skipping school, fighting, and spray-painting graffiti may infer that neighbors have intervened in the past, and would in the future. Similarly, if respondents are uncertain that neighbors would intervene, they may be more likely to assume that such behavior is rare. In contrast, respondents who are married, have children, and own homes may be particularly concerned about child safety, may themselves have intervened in the past, and may assume that others will as well. Finally, residents who have been victims of crime may assume that neighbors

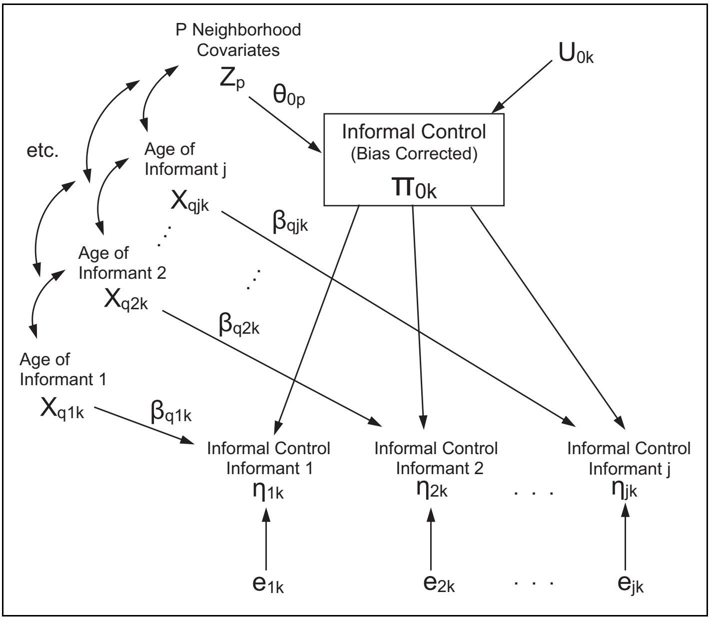

Figure 4. Two-level hierarchical linear model (HLM) model of an individual-level model of bias in informal control η_{jk} (using Age X_{qjk} as an example covariate) and a neighborhood level model of bias-corrected informal control Πok predicted from structural covariates Z_p.

are unlikely to intervene, given that their victimization was not thwarted (e.g., Maccoby, Johnson, and Church 1958).

> **Hypothesis 4:** Individual responses to informal social control are a function of an individual's reports of observed child deviance, age, sex, marital status, number of children, education, income, race-ethnicity, residential mobility, homeownership, previous criminal victimization, and uncertainty in reports of informal control.

We can test the specific hypothesis, derived from St. Jean's (2007) ethnography that affluent residents infer that their neighbors will intervene when problems arise if their neighbors are also of relatively high SES. This implies a cross-level interaction in an HLM model between neighborhood SES, measured by concentrated affluence, and respondent SES, measured by income and education in individual-level models of informal social control:

Hypothesis 5: Individual responses to informal social control are higher on average when respondents have high income and education and the neighborhood is relatively advantaged.

The degree of intersubjective correspondence among respondents is given by the intra-neighborhood correlation ρ^I, which indexes the proportion of variation between neighborhoods versus within neighborhoods: $\rho^I = \tau_\pi^2/(\tau_\pi^2 + \tau_e^2)$, where τ_π is the between-neighborhood variance in responses and τ_e is the within-neighborhood variance (e.g., Raudenbush and Bryk 2002). This correlation estimates the neighborhood reliability for a single respondent per neighborhood. Increasing the number of respondents per neighborhood (n_k) increases reliability—in the same way that increasing the number of test items increases the reliability of a test by reducing error variance—according to the following:

$$\lambda_k = \frac{\tau_\pi^2}{\tau_\pi^2 + \tau_e^2/n_k}. \tag{14}$$

This can be seen to measure aggregate reliability for each neighborhood because the numerator gives the *true* (i.e., true score) variance of the group means and the denominator gives the expected *observed* variance of the group means. We can obtain an average score for the *k* neighborhood reliabilities as:

$$1/K \sum_{k=1}^{K} \lambda_k. \tag{15}$$

Level 4 Validation Model for Construct Validity of Informal Social Control

We examine the construct validity (e.g., Messick 1995) of informal social control at the neighborhood level by modeling the random intercepts π_{0k} from equation (13) according to the following:

$$\pi_{0k} = \Theta_{00} + \sum_{p=1}^{P} \Theta_{0p} Z_p + U_{0k}, \tag{16}$$

where Θ_{00} is the city-wide intercept, Z_p are neighborhood covariates, Θ_{0p} are regression coefficients, and U_{0k} are random disturbances assumed normally distributed and *iid* (see Figure 4). Here we examine whether informal social control is associated with neighborhood theoretical constructs, such as concentrated disadvantage, residential stability, ethnicity/immigration, and population density. Moreover, we will compare models with different measurement assumptions to assess whether differences in models produce different substantive estimates.

Hypothesis 6: Neighborhood informal control is associated with concentrated disadvantage, residential stability, ethnicity/immigration, and population density.

Finally, as an additional test of construct validity, we estimate a model predicting counts of violent crime. We use a two-step procedure to estimate these models. First, we estimate specific three-level measurement models as outlined previously, obtain empirical Bayes's (shrinkage) estimates of model parameters, and, for each model, compute scores for neighborhood informal social control based on the empirical Bayes's residuals.[8] Second, we regress violence on informal social control scores plus neighborhood covariates taken from the census. Because our dependent variable (violent criminal acts), y_i, is an overdispersed count variable, we fit a negative binomial model estimated by ML:

$$E(y_i) = \mu_i = \exp(x_i \beta + \varepsilon_i), \tag{17}$$

where μ_i (and, equivalently, ϵ_i) is a random variable assumed to follow a gamma distribution, x_i is a vector of neighborhood covariates, and β is a vector of logit coefficients (e.g., Cameron and Trivedi 1998; Long 1997). In this model, ε_i captures overdispersion due to cross-sectional heterogeneity. This allows us to test our final hypothesis:

Hypothesis 7: Net of neighborhood structural characteristics, neighborhood informal control is associated with future violent crime.

Again, we will compare measurement models to assess whether more plausible models yield different substantive estimates.

Research Design, Data, and Measures

The Seattle Neighborhoods and Crime Survey

We analyze data from the Seattle Neighborhoods and Crime Study (SNCS), which conducted a survey of 5,755 residents within 123 census tracts in 2002 to 2003 (Matsueda 2010).[9] This project defined neighborhoods in terms of 123 census tracts in the city of Seattle. The SNCS collected data from three separate sampling schemes: (1) a random sample of households within census tracts, (2) a sample of households within each of six street segments of 100 census tracts sampled by Terrance Miethe's (1991) earlier survey of Seattle residents, and (3) a race-ethnic oversample of census blocks with the highest proportion of racial and ethnic minorities. We focus on the first two samples, which are strongly representative of the city of Seattle. The 3,739 cases in these samples were selected from continuously updated white pages and interviewed by telephone by the Social Behavioral Research Institute at California State University, San Marcos.[10] The survey data were merged with 2000 U.S. Census data for each tract and crime data by census tract from the Seattle Police Department. Our analyses are based on 3,260 cases, which allow us to capitalize on the asymptotic properties of WLS and ML estimators.[11]

Measures

Informal social control. We use four measures of informal social control taken from earlier surveys (e.g., Sampson et al. 1999; Taylor 1996). Each item asks how likely it is that neighbors would do something about four hypothetical situations involving youths engaging in delinquent activities (skipping school, spray-painting graffiti, disrespecting adults, and fighting in the street; see Table 1). To estimate test–retest reliability, we asked the question about fighting in the street twice, once at the beginning of the survey, and again toward the end, approximately 30 minutes later. The time interval between tests, which is similar to other studies (e.g., Hindelang, Hirschi, and Weis 1981) is sufficiently short to ensure that the true scores have not changed, and long enough—particularly with over one hundred intervening survey questions for respondents to answer—to minimize memory effects.

Observation of child deviance. For each of the delinquent activities mentioned in the informal control items (e.g., spray-painting graffiti), we also asked respondents whether they had ever observed the activity. This allows us to test the hypothesis that responses to informal control measures are formed in part by whether respondents had previously observed the behavior. We expect that those respondents who have observed the deviant act in their neighborhood are likely to infer that neighbors would not intervene if the deviant act occurred again. Conversely, respondents who have never observed the deviant act are more likely to infer that neighbors would intervene (and perhaps already have). We also include a latent variable score of observed deviance in our individual-level models of response bias in informal social control.

Table 1. Descriptive Statistics for Individual-Level Measures.

	Mean	*SD*
Informal social control		
Neighbors would do something if children hanging out	2.67	.91
Neighbors would do something if kids painting graffiti	3.44	.72
Neighbors would scold the child if disrespecting adults	2.53	.83
Neighbors would stop it if children fighting (test)	3.18	.78
Neighbors would stop it if kids fighting (retest)	3.18	.83
Uncertainty		
Respondent changed their answer to the repeated question	.36	.56
Observation of child deviance		
Observed children hanging out	.16	.37
Observed children painting graffiti	.04	.18
Observed children disrespecting adults	.20	.40
Observed children fighting	.11	.31
Recent victimization		
Violent victimizations in last two years	.07	.25
Property victimizations in last two years	.41	.49
Demographics, socioeconomics, and residential status		
Female	.49	.50
Age (in tens of years)	4.85	1.57
Married/ cohabitating	.55	.50
Number of children living at home	.42	.82
Years of education completed	16.12	2.49
Household income in thousands of dollars (mean-replaced)	69.83	45.40
Income missing flag	.11	.31
Asian	.06	.24
African American	.04	.19
Hispanic	.05	.21
Foreign-born	.12	.32
Number of residential moves in the last five years	.85	1.43
Number of years at current address	11.63	12.19
Homeowner	.68	.47
$N = 3{,}166$		

Note: SD = standard deviation.

Individual characteristics. We are using residents as informants for informal social control in their neighborhoods. As noted earlier, these evaluations are likely to vary as a function of the demographic and biographical characteristics of informants. To correct this relative bias, we control for characteristics of informants that may influence their scores (see Sampson et al. 1997), including sex, age, marital status, number of children, education, income, race-ethnicity, residential mobility, homeowner, and victim of a violent or property crime. The latter allows us to test a version of the hypothesis that being a victim of a crime affects responses to informal social control. To test the hypothesis—derived from Tversky and Kahneman's (1973) availability heuristic—that respondents who are uncertain about informal social control in their neighborhoods are more likely to assume that such control would rarely occur, we computed a difference score from our repeated measure of informal control. By taking the absolute value of the difference, we create an uncertainty score in which high values reflect changes in respondents' answers, which we assume means respondents are less certain of informal control in their neighborhoods.

Neighborhood measures. We can assess construct validity by examining associations between informal social control and other neighborhood-level constructs specified by theory to be strongly correlated. Following previous research (e.g., Land, McCall, and Cohen 1990; Sampson et al. 1997), we created for each census tract five indexes of neighborhood structural characteristics from the U.S. Census: concentrated disadvantage, concentrated affluence, ethnicity/immigration, residential stability, and population density. We also created a three-year average of violent crimes (murders, rapes, robberies, and aggravated assaults) by census tract for 2003 to 2005 from the Seattle Police Department (see Table 2).

Estimation of Model Parameters

Models of Test–Retest Reliability and Interitem Reliability

We estimate our first- and second-level models, which include ordinal indicators, as a single system using Jöreskog and Sörbom's (2001) WLS estimator, which provides consistent and asymptotically efficient estimates in very large samples (e.g., >1,000).[12] Our model, diagrammed in Figure 2, is a simple two-level, three-factor model, in which the first-level models test–retest reliability in fighting in the street, the second level models inter-indicator reliability for measures of informal control and observed deviance. Note that the measurement model for informal control also tests Hypothesis 3, in which the respondent's *observation* of a neighborhood problem (e.g., spray-painting graffiti) affects their *belief* that neighbors will intervene in the problem (see equation 8). This model decomposes the observed correlation between a specific child deviant act (e.g., skipping school) and the corresponding informal control item (e.g., intervene when children are skipping

Table 2. Descriptive Statistics for Neighborhood-Level Measures.

	Mean	*SD*
Concentrated disadvantage		
Average of z-scores	.00	.80
Proportion in poverty	.12	.09
Proportion unemployed	.05	.04
Proportion on public assistance	.03	.03
Proportion single-mother households	.09	.07
Proportion African-American	.08	.10
Concentrated affluence		
Average of z-scores	.00	.92
Proportion households with income > US$100K	.16	.10
Proportion college graduates	.47	.17
Proportion managerial or professional occupations	.48	.13
Ethnicity/Immigration		
Proportion Latino	.05	.04
Asian/foreign-born (average of z-scores; $\alpha = .97$)	.00	.98
Proportion Asian	.12	.12
Proportion foreign born	.16	.11
Residential stability		
Average of z-scores	.00	.97
Proportion homeowners	.50	.23
Proportion in same residence five years ago	.44	.13
Population density		
Tens of thousands of person per square mile by tract	.94	.69
Violent crime rate		
Average yearly violent crimes per 1,000 population 2003–2005	8.21	12.01
Total violent crimes 2003–2005	84.01	82.26
N = 123		

Note: SD = standard deviation.

school) into two components: (1) a direct effect between the two indicators and (2) a direct effect between the two factors, observed child deviance and informal control.

Given the large sample size and large number of overidentifying restrictions, the model fits the data reasonably well: $\chi^2 = 52.93$, $df = 20$, $p < .001$; root mean square error of approximation = .022). Table 3 presents coefficients for the model of equations (4) to (6). The level-1 test–retest factor model reveals high test–retest reliability for our two measures of neighbors intervening if children were fighting in the street (Hypothesis 1). The first item has slightly smaller measurement error variance and consequently a higher reliability (.79) than the retest (.69), perhaps due to respondent fatigue

Table 3. Parameter Estimates of Measurement Models.

Variable		Error	Metric	
Latent	Indicators	Variance	Slope	Reliability
Stop fight	Stop fight (test)	.21	1.00[a]	.79
	Stop fight (retest)	.32	.93	.69
Observe deviance	Observe skip school	.48	1.00[a]	.52
	Observe graffiti	.70	.75	.30
	Observe disrespect	.35	1.11	.66
	Observe fight	.32	1.14	.69
Informal control	Do something skip school	.44	1.00[a]	.49
	Do something graffiti	.45	1.01	.50
	Scold child disrespect adult	.62	.87	.37
	Stop fight	.18	1.07	.72

Note: $N = 3{,}260$.
[a]Indicates a fixed coefficient.
All coefficients are statistically significant at $p < .001$.

or recognition that it was asked earlier. These reliabilities imply a test–retest correlation between the two observable items of $(.79 \times .69)^{1/2} = .74$. The metric slope of the retest is slightly smaller than 1.0, the value fixed for the test, suggesting a slight correlation between measurement error and latent factor. A likelihood ratio test rejects the assumption of τ-equivalence for our two test–retest measures ($\chi^2 = 9.61$, $df = 1$, $p < .001$).

The second-level measurement model captures inter-item reliability. Because we have test–retest items for fighting we can disentangle the measurement error due to test–retest unreliability from that of inter-item unreliability. Thus, the measurement error variance for fighting due to inter-item unreliability is very small and the reliability coefficient is relatively large (.72). For the other three informal social control items, measurement error variances are larger and reliabilities smaller (.4–.5) because they pool inter-item unreliability and test–retest unreliability. The corresponding quantity is simply the squared indirect effect of the informal control factor on the fighting item for the test $(.83 \times .89)^2 = .55$ and retest $(.83 \times .83)^2 = .47$ (see Figure 5), which is comparable to the reliabilities of the other three items. For the fighting item, pooled reliability is approximately evenly split between test–retest and inter-item reliability. Thus, we find support for Hypothesis 2.

The metric slopes suggest modest correlations between factor and error: Relative to skip school, respondents scoring low on informal control tend to overstate on disrespect and understate on fight, and those scoring high

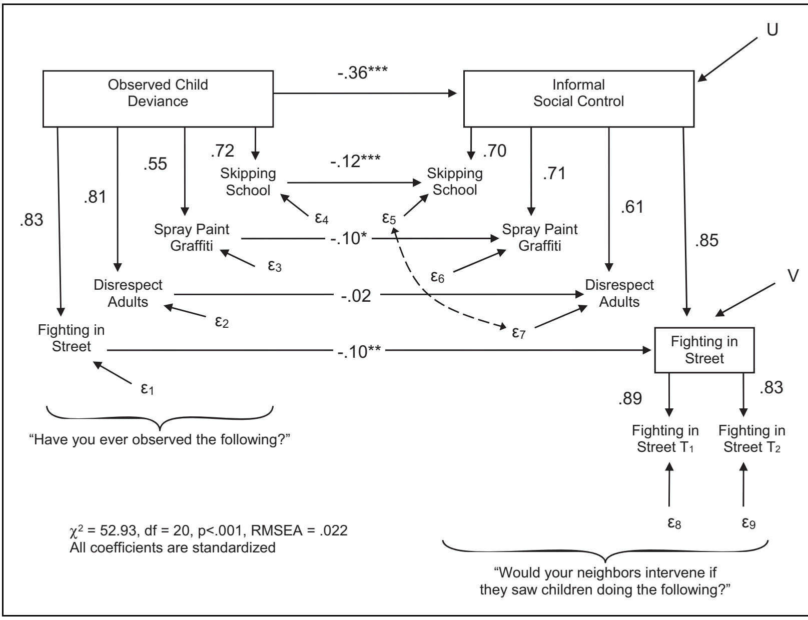

Figure 5. Standardized parameter estimates for a measurement model of neighborhood informal social control and observed child deviance.

on informal control tend to understate on respect and understate on fight.[13] Although these departures from unity are modest, we are able to reject the hypothesis of τ-equivalence for measures of informal control ($\chi^2 = 22.42$, $df = 3$, $p < .001$). We also reject the hypothesis of equal measurement error variances ($\chi^2 = 10.23$, $df = 3$, $p < .001$), and thus parallel measures (equal slopes and variances: $\chi^2 = 32.65$ $df = 6$, $p < .001$). For indicators of observed child deviance, we obtain reasonable reliabilities (approximately .60), with the exception of spray-painting graffiti, which is only .30. The metric slopes reveal that, with the exception of observing graffiti, respondents appeared to use the ordinal scales in similar ways. Finally, for the observed deviance factor, we reject hypotheses of τ-equivalence, equal measurement error variances, and thus, parallel measures.

This model provides support for the hypothesis that informants base their assessments of neighbors intervening in child deviance in part on whether they have personally observed the specific child deviant act. Thus, at the level of the item (specific child act), we find modest support for Hypothesis 3. Informants who have observed the act (such as graffiti) are less likely to believe their neighbors would intervene if the act were to occur—presumably because neighbors already would have—and vice versa. The coefficients are significant but small (one-tenth of a standard deviation) with the exception of disrespecting adults, which is nonsignificant. Nevertheless, even with these direct effects in the model, the reliabilities of informal control items remain reasonably high. Finally, if one could only use one indicator of informal control, we would recommend using the item with the highest reliability and smallest measurement error variance—intervene in kids fighting in the street.

The two factors are substantially negatively correlated (−.36) as hypothesized. Informants who report observing child deviance in the neighborhood score lower on informal social control—presumably because they infer from observing child deviance that neighbors have not and will not intervene. The substantial correlation between factors underscores the importance of including observed child deviance as a bias factor in individual-level models of informal control, to which we now turn.

Individual-Level Models of Informant Characteristics on Informal Control

We estimated our level-3 and level-4 statistical models—on latent variable scores from our level-1 and level-2 models—using the statistical packages R (R Development Core Team 2010) and HLM (Raudenbush and Bryk 2002). Tables 4 and 5 present ML estimates of the parameters. Table 4 presents the individual-level coefficients from the regression of informal control scores on characteristics of respondents.[14] We find that respondents who are older, married or cohabitating, and have more children in the home tend to overstate informal control compared to their younger, single, and childless counterparts. Thus, as hypothesized, such respondents may be more concerned with controlling children, which enables them to recall instances of neighbors intervening. Those with more education slightly understate informal control. Compared to renters, homeowners tend to overstate informal control, as do immigrants compared to the native-born. Also as expected, respondents who are uncertain about the local level of informal control—those who change their responses to the repeated measure of informal control—understate informal control moderately, a finding consistent with Tversky and Kahneman's availability heuristic. We do not find support for the hypothesis, based on St. Jean's (2007) findings, that affluent respondents assume that high-SES neighbors would intervene when child problems arise. The interaction between income and education, on the one hand, and

Table 4. Person-Level Coefficients From Multilevel Models Predicting Informal Social Control.

Person Level	β	*SE*	St. β
Female	.00	.02	—
Age	.04***	.01	.09
Married/cohabitating	.04*	.02	—
# of children in home	.04**	.01	.04
Years of education	−.01	.00	−.02
Income (mean replaced)	.00	.00	.02
Did not report income	.00	.03	—
Asian	−.04	.04	—
African American	.00	.04	—
Hispanic	.04	.04	—
Foreign born	.08**	.03	—
Residential moves	.00	.01	.00
Years in neighborhood	.00	.00	.00
Homeowner	.06*	.02	—
Violent victimization	−.03	.03	−.01
Property victimization	−.01	.02	−.01
Uncertainty	−.10***	.01	−.08
Observed child delinquency	−.47***	.02	−.31
Intercept	2.55***	.03	—
Variance explained (within)	.22		

Note: $N = 3,166$ persons, 123 neighborhoods. This is the person-level portion of model 3 of Table 5 and includes controls for sample type and the neighborhood covariates listed in Table 5.
*$p < .05$. **$p < .01$. ***$p < .001$ (two-tailed).

concentrated affluence, on the other, is small and fails to reach statistical significance (coefficient not shown).

Finally, respondents who have observed acts of child deviance substantially understate informal social control. This supports Hypothesis 3 at the construct level, based on a cognitive response theory, that an important source of respondents' assessments of informal control is the degree of deviance they observe in the neighborhood. This is perhaps a positive, substantively, in that this is a reasonable source of information, but also a negative, methodologically, because ignoring the effect may bias substantive coefficients due to compositional differences in observing deviance across neighborhoods. This coefficient is far and away the largest in the model (−.31 standardized); the next largest is age (.09). The other coefficients are very small in magnitude and the model as a whole explains only 22 percent of

Table 5. Neighborhood-Level Coefficients Predicting Informal Social Control.

	Model 1	Model 2	Model 3
Concentrated disadvantage	−.06*	−.06*	−.02
	(.03)	(.03)	(.02)
	−.19	−.20	−.08
Concentrated affluence	.07**	.05*	.02
	(.03)	(.02)	(.02)
	.27	.20	.08
Proportion Latino	−.17	−.25	−.35
	(.44)	(.42)	(.34)
	−.03	−.04	−.07
Asian/foreign born	−.04*	−.04	−.04*
	(.02)	(.02)	(.02)
	−.18	−.16	−.18
Residential stability	.11***	.05**	.03
	(.02)	(.02)	(.02)
	.46	.21	.16
Population density	−.07**	−.06*	−.02
	(.03)	(.03)	(.02)
	−.20	−.17	−.08
Intercept	2.95***	2.95***	2.55***
	(.04)	(.04)	(.03)
Variance explained (between)	.85	.87	.87

the variance in reports of informal social control. In other words, in this model, the majority of error variability in reports of informal control is random rather than systematic, which provides some evidence of the validity of our individual reports of informal social control. Nevertheless, adjustment for compositional differences in residents across neighborhoods appears warranted to avoid biases in substantive models.

Neighborhood-Level Models of Informal Social Control

Estimates of our level-4 neighborhood model of informal social control appear in Table 5. These estimates provide reduced-form coefficients of a structural model of neighborhood informal control and provide evidence of construct validity of our informal control construct across neighborhoods. Moreover, by comparing the effects of neighborhood structure on informal control across our three measurement models, Table 5 reveals whether estimating more plausible measurement models has an effect on substantive quantities of interest to criminologists.

Model 1 is a simple naive model in which we fail to model test–retest and inter-item unreliability and also fail to control for individual characteristics at level 3. We find, for model 1, that concentrated disadvantage exerts a moderate and significant negative effect on informal control, as found in the prior studies. Concentrated affluence exerts a strong positive effect on informal control (standardized coefficient is .27), while Asian/Immigration exerts a moderate negative effect (−.18). As previous research has shown, residential stability exerts the strongest effect on informal control (.46), which is consistent with a systemic model of neighborhood control. Finally, as expected, population density is significantly negatively associated with informal social control. Thus, as expected, neighborhoods with greater affluence, a sparse population base, a high percentage of homeowners, and a high proportion of long-term residents have effective informal control perhaps because they have both the incentives and the resources to act jointly for the collective good. These results, consistent with those of Chicago, provide support for the construct validity of informal social control in Seattle neighborhoods.

Model 2 is similar to the measurement model of Sampson et al. (1997), which controls for inter-item unreliability at level 2 and controls for individual characteristics at level 3. With one exception, the point estimates are relatively similar to those of model 1 (column 2). The exception is the coefficient for residential stability, which is about 50 percent smaller in model 2. Thus, controlling for the potential biasing influence of individual characteristics on informant reports of informal social control and correcting for the disproportionate distribution of such characteristics across neighborhoods via regression weights—as recommended by Sampson et al. (1997)—leads to correcting for a 54 percent overestimate of the effect of residential stability on informal control.

Model 3, our most plausible model, includes our level-2 measurement model with observed child deviance items loading on an observed deviance factor (see Figure 3), which is included as an individual-level covariate in our level-3 model. By comparing columns 3 with 2, we find severe attenuation in several coefficients for model 3 when compared to model 2. For example, the coefficient for concentrated disadvantage is 60 percent smaller in model 3 and no longer significant (line 1); the coefficient for concentrated affluence is 65 percent smaller and no longer significant (line 2), the coefficient for residential stability is 27 percent smaller and no longer significant (line 5), and the coefficient for population density is over 40 percent smaller and no longer significant (line 6). In this model, only Asian/foreign-born remains statistically distinguishable from zero.

Table 6. Coefficients From a Negative Binomial Regression of Violent Crime (2003–2005) on Neighborhood-Level Covariates.

	Model 1	Model 2	Model 3
Logged population	.43**	.41**	.42**
	(.13)	(.13)	(.13)
	1.54	1.51	1.52
Concentrated disadvantage	.39***	.39***	.44***
	(.09)	(.09)	(.10)
	1.48	1.48	1.56
Concentrated affluence	−.11	−.13	−.18
	(.10)	(.10)	(.10)
	.90	.88	.84
Proportion Latino	2.20	1.98	1.98
	(1.50)	(1.49)	(1.54)
	9.01	7.22	7.22
Asian/foreign born	.06	.08	.08
	(.07)	(.07)	(.08)
	1.06	1.08	1.08
Residential stability	−.35***	−.42***	−.46***
	(.08)	(.07)	(.07)
	.70	.66	.63
Population density	−.08	−.07	−.03
	(.10)	(.09)	(.10)
	.92	.93	.97
Informal social control	−1.45***	−1.86***	−2.24**
	(.41)	(.50)	(.83)
	.23	.15	.11
Intercept	.35	.51	.47
	(1.05)	(1.04)	(1.07)
Dispersion (θ)	3.95	3.99	3.78
	(.53)	(.54)	(.50)

Neighborhood-Level Models of Violence

As another test of construct validity, we regressed neighborhood violence on informal social control plus our neighborhood structural covariates using a negative binomial model. We again compare coefficients from our three measurement models to assess whether differences in measurement specifications make a difference for quantities of interest to criminologists. Results of our three models of violence appear in Table 6. Model 1 is based on the naive measurement model that fails to control for individual covariates. This model reveals that, as expected, violence rates are higher in neighborhoods that are more disadvantaged and experience more population turnover. The coefficient for informal social control is -1.45: As hypothesized, net of other structural covariates, neighborhoods with greater informal social control experience significantly fewer violent crimes.

By comparing the coefficient for informal social control across models, we can determine whether different controls for response errors alter our key result. Model 2, which controls for individual characteristics in our level-3 models, reveals a coefficient for informal social control of -1.86, which is 28 percent larger than the coefficient for model 1. Thus, controlling for individual covariates substantially disattenuates the effect of informal control on violence. A more plausible specification, model 3, treats observed child deviance measures as indicators of a latent observed child deviance factor and includes the factor, along with other individual characteristics, in our level-3 response model. Model 3 reveals a coefficient for informal social control of -2.24, which is still 20 percent larger than the coefficient for model 2 and over 50 percent larger than the coefficient for model 1. Thus, previous models of violence may have *underestimated* the effect of informal control; those effects should be considered to be *conservative* estimates.

Discussion

Nested data, which are increasingly prevalent in the social sciences, present distinct measurement issues at multiple levels of explanation. We have proposed a multilevel measurement model for nested data that combine the psychometric models based on classical test theory at the level of indicators and individuals, and ecometric models of response bias for ecological (neighborhoods) units. Our models generalize to nested data on other important units, including groups, cities, counties, states, and nations.

By explicitly modeling the measurement process of informal social control, we are able to (1) test hypotheses about the social process producing responses to measures, (2) assess the validity and reliability of measures at multiple levels, and (3) estimate the potential bias in estimates of substantive coefficients of interest that results from failure to use the appropriate measurement models. We discuss each in turn. With respect to hypotheses about measurement, we find that informants' responses to whether neighbors will intervene in child deviance are shaped by their previous observations of child deviance. This effect is present both in item-specific models and in individual-level models of the informal control construct. This finding is consistent with a cognitive model of survey response, in which respondents first comprehend the question and second, retrieve two pieces of information from memory: (1) Do children engage in such deviant acts? (2) Do neighbors intervene into such acts of deviance? At times, information on the former may inform the latter. We also find, consistent with Sampson et al. (1997, 1999), that responses to informal control are a function of respondent characteristics as hypothesized, which implies that adjustments for neighborhood composition of such characteristics are necessary to obtain unbiased estimates of neighborhood informal social control.

Our models find that measures of informal social control have reasonable measurement properties, including moderate test–retest reliability for fighting in the street, and reasonable inter-item reliabilities of all indicators. On this score, our results are consistent with those of Sampson and colleagues using the Chicago PHDCN data. The projects differ somewhat in year of survey, sampling design, data collection method, and city demographics, racial composition, and history. The similarity of measurement properties of informal social control across the two studies provides evidence for the invariance of measurement structure of the concept. We can also make recommendations for survey researchers who want to incorporate the best measures of informal social control at the least cost. The observed distributions of skipping school and disrespecting adults are closest to normal, and thresholds are closest to the observed ordinal intervals (see Appendix Figure 1 [which can be found at http://smr.sagepub.com/supplemental/]); therefore, these items may be best to use if informal control is endogenous and if conventional methods, which treat items as continuous and Gaussian, are used. However, when using more appropriate models for ordinal indicators, we find that fighting in the street followed by skipping school and spray-painting graffiti have the smallest measurement error variances, similar relative slopes, and therefore, the highest reliabilities. Note that when correctly modeled, using all four indicators would be better because they are all reasonably reliable, and adding additional well-behaved indicators increases the statistical power to detect relationships among latent constructs (see Matsueda and Bielby 1986). Nevertheless, if cost limits a researcher to a single item and models for ordinal measures are unavailable, we would recommend using the skipping school item, which is closest to normally distributed, has intervals that approximate thresholds, and has high reliability.

Perhaps the most important and striking finding from our models is that measurement models matter: Using the correct measurement model entails correcting for attenuation due to unreliability, which reduces bias in substantive coefficients in which social scientists care. In models of the causes and consequences of informal control, our most plausible measurement model—which corrected for inter-item unreliability, controlled for respondent characteristics in level 3, and controlled for respondents having observed past deviance—finds that previous research overestimated the effects of neighborhood covariates on informal control and underestimated the effects of informal control on violence. Models that fail to control for respondents having observed deviance, such as that of Sampson et al. (1997), tend to overestimate coefficients in models of informal control and underestimate the effect of informal control on violence. Moreover, our naive models that fail to control for inter-item unreliability, respondent characteristics, and observed deviance dramatically overestimate coefficients in models of informal control and dramatically underestimate the effect of informal control on violence. The results of previous research using such naive models should perhaps treat the effects of neighborhood covariates on informal control as upper bounds, and effects of informal control on violence as lower bounds.

Our modeling benefited from the use of a cognitive theory of survey response, allowing us to test hypotheses about the sources of responses to measures of informal control. We have not, however, fully explored the ways in which respondents comprehend questions, retrieve relevant information from memory, make judgments about that information, and select an appropriate answer. Such an exploration would require a different research design. For example, qualitative in-depth interviewing of respondents may reveal the specific cognitive steps respondents take in arriving at a frequency answer to a factual question, such as neighborhood informal control. Experimental studies of memory effects may identify precise pathways of retrieval of information. Such studies of memory effects contrast two models of how frequency estimates are stored and retrieved. A tally model, like Bayesian learning models, suggests that individuals keep a running tally of the frequency of events, and update them as new information is encountered (e.g., Howell 1973) and find that retrieval can be automatic (Hasher and Zacks 1984). A competing model argues that frequency estimates are computed from either memories of individual instances (such as instances of child deviance and whether neighbors intervened) or from overall familiarity with the events (Tourangeau et al. 2000). For example, as noted earlier, Tversky and Kahneman (1974) found that events that could be easily retrieved from memory were deemed more probable by subjects, a heuristic they termed "availability." They also found that subjects often failed to update probabilities in light of new information, particularly when that information was not vivid or salient.

If this line of theorizing about responses to informal control items is correct, and the processing of new information is paramount, a broader question concerns information flows: How do residents acquire the knowledge of the safety and behavior of their neighbors? Such information has a variety of sources, including direct observation (as we have found here), dissemination from other residents, and perhaps through reports and depictions from the mass media. Information about the neighborhood is likely to be structured in complex ways by the network connectivity of residents. For example, neighborhoods with a small number of extremely dense but unconnected local hubs may have heterogeneous responses compared to one with weak ties across hubs. The latter may function to increase consensus by introducing disparate forms of information across local networks. Future research is needed to examine the contextual sources of information leading to survey responses.

Over 45 years ago, Blalock (1968:21) called for the construction of "auxiliary measurement theories," which would identify the theoretical assumptions behind the process of measuring sociological concepts, link observable indicators to theoretical constructs, and allow researchers to distinguish between research findings that reject measurement theories versus reject conceptual theories. For example, auxiliary measurement theories would theorize not only a stimulus–response relationship within an interview setting, but also the intervening cognitive processes of information processing and retrieval. This would make theory construction and measurement construction part of the same process (Blalock 1982). We extend Blalock's call for auxiliary measurement theories to extend to multiple levels of nested data, such as neighborhoods, groups, cities, and nations. Such a call raises new substantive puzzles and methodological questions.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Supported by grants from the National Science Foundation (SES-0004324, SES-0966662). the National Consortium on Violence Research (SBR-9513040), and the Blumstein-Jordan Endowed Professorship in Sociology.

Notes

1. The models differ in where marginalization occurs: For item response theory (IRT) models, marginalization occurs conditional on subject abilities; for structural equation models (SEM), marginalization occurs on the continuous latent variable before dichotomizing or ordering the continuous variable (for details, see De Leeuw 1983; Takane and de Leeuw 1987). The models diverge when

estimating multiple-population models and testing for differences in both slopes and intercepts; in such cases, SEMs provide accurate estimates of slope differences but biased estimates of intercept differences (see Kankaras, Vermunt, and Moors 2011). This is not an issue for our models, which estimate single population models, and, accordingly, leave intercepts unconstrained. For applications of IRT to self-reported delinquency, see Osgood et al. (2002) and Piquero et al. (2000).

2. Much earlier, Maccoby et al. (1958:39) hypothesized that in disorganized neighborhoods, "individual adults will feel less responsible for guiding other people's children," and "will ignore deviant acts when they see them being committed, unless they themselves are directly involved." Maccoby et al. (1958:45) used a small-scale survey to operationalize informal control as "individual residents being unwilling to take action if they observed their neighbors children engaged in some sort of deviant behavior." Taylor (1996) altered this to refer to resident reports about neighbors intervening.
3. Sampson (2012) recently used two unobtrusive measures of altruism to validate collective efficacy, under the assumption that efficacious neighborhoods elicit altruistic behavior. One was the classic letter drop experiment and the other was the administration of cardiopulmonary resuscitation to heart attack victims. Both were correlated with informal social control.
4. Results for alternate measures of reliability appear in Online Appendix Table 1 (which can be found at http://smr.sagepub.com/supplemental/).
5. This provides estimates of reliability for each indicator because we are relaxing the assumption of τ-equivalence ($\lambda_i = \lambda_j$). By contrast, the usual α reliability must assume τ-equivalence and when it assumes parallel measures (equal measurement error variances ${\sigma_{\delta_i}}^2 = {\sigma_{\delta_j}}^2$) it can estimate reliability from the average intercorrelations among items in the scale.
6. Skrondal and Laake (2001:574) show that, in general, regressions among traditional factor scores yields inconsistent estimates of regression coefficients, whereas regressions among "covariance preserving factor scores" provide consistent estimates (see also Jöreskog 2000; Ten Berge et al. 1999). We compared regressions among latent scores using a two-step procedure with estimates based on estimating measurement models and regressions among latent variables simultaneously and obtained nearly identical results.
7. The latent variable scores allow us to incorporate information from our unconventional measurement model for the first two levels into our models at the final two levels. This two-step strategy has the advantage of minimizing potential cross-level bias, but the disadvantage of ignoring covariation of estimates across all levels, causing standard errors to be slightly underestimated. We estimated simple three-level models (assuming parallel measures) simultaneously and found this difference in standard errors to be negligible.

8. The empirical Bayes's residuals, U^*_{0k}, are equal to the least squares residuals, U_{0k}, from equation (16) shrunk toward zero by an amount equal to the reliability, λ_k (intersubjective correspondence among respondents across neighborhoods) from equation (14): $U^*_{0k} = \lambda_k U_{0k}$. It is well known that shrinkage estimators minimize mean squared error (see Raudenbush and Bryk 2002:47).
9. The American Association for Public Opinion Research, Council of American Survey Research Organizations-4 response rate was 51.3 percent with a cooperation rate of 97 percent.
10. Results based on the full sample were nearly identical. An additional sample of 355 cases was drawn from an enumeration of households not listed in the updated white pages, and surveyed with questionnaire surveys. This sample along with a group of 545 respondents drawn in the original three samples but who opted to complete the survey by mail rather than phone were employed in separate analyses as a check on our main findings.
11. Missing values dropped the sample by less than 13 percent. Our analyses suggest that these missing values do not affect our substantive results in any appreciable way.
12. Maximum likelihood estimation with Satorra-Bentler scaled statistics sometimes work better when sample sizes are not large (e.g., Lei and Wu (2012). We estimated our models using this procedure and found similar results.
13. Because latent variables are normalized by fixing a metric slope to unity for a reference indicator, other slopes are identified only relative to the reference slope. Therefore, estimated slopes that depart from unity indicate correlation between error variance and true score, which substantively implies different uses of the scale relative to the reference indicator (e.g., Bielby 1986).
14. Coefficients for the neighborhood-level portion of the model presented in Table 1 appear in in column 3 of Table 5. The individual-level coefficients for the models presented in columns 1 and 2 of Table 5 are virtually identical, and therefore, are not presented.

Supplementary Material

The online appendix is available at http://smr.sagepub.com/supplemental/.

References

Anderson, J. R. 1983. *The Architecture of Cognition*. Cambridge, MA: Harvard University Press.

Anderson, T. W. and Herman Rubin. 1949. "Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations." *Annals of Mathematical Statistics* 20:46-63.

Bandura, Albert. 1986. *Social Foundations of Thought & Action: A Social Cognitive Theory*. Englewood Cliffs, NJ: Prentice-Hall.

Bandura, Albert. 1997. *Self Efficacy: The Exercise of Control*. New York: W. H. Freeman.

Bentler, Peter M. 2009. "Alpha, Dimension-free, and Model-based Internal Consistency Reliability." *Psychometrika* 74:137-43.

Bielby, William T. 1986. "Arbitrary Metrics in Multiple Indicator Models of Latent Variables." *Sociological Methods and Research* 15:3-23.

Blalock, Hubert M. 1968. "The Measurement Problem: A Gap between the Languages of Theory and Research." Pp. 5-27 in *Methodology in Social Research*, edited by H. M. Blalock and A. B. Blalock. New York: McGraw-Hill.

Blalock, Hubert M. 1982. *Conceptualization and Measurement in the Social Sciences*. Beverly Hills, CA: Sage.

Bollen, Kenneth A. 1989. *Structural Equations with Latent Variables*. New York: John Wiley.

Broege, Nora, Ann Owens, Anthony P. Graesch, Jeanne E. Arnold, and Barbara Schneider. 2007. "Calibrating Measures of Family Activities between Large- and Small-scale Data Sets." Pp. 119-50 in *Sociological Methodology*, Vol. 37, edited by Yu Xie. Boston, MA: Blackwell.

Browne, Michael W. 1984. "Asymptotic Distribution Free Methods in Analysis of Covariance Structures." *British Journal of Mathematical and Statistical Psychology* 37:62-83.

Cameron, Colin A. and Pravin K. Trivedi. 1998. *Regression Analysis of Count Data*. New York: Cambridge University Press.

Cronbach, L. J. 1951. "Coefficient Alpha and the Internal Structure of Tests." *Psychometrika* 16:297-334.

Coleman, James C. 1990. *Foundations of Social Theory*. Cambridge, MA: Harvard University Press.

Coppedge, Michael, John Gerring, David Altman, Michael Bernhard, Steven Fish, Allen Hicken, Matthew Kroenig, Staffan I. Lindberg, Kelly McMann, Pamela Paxton, Holli A. Semetko, Svend-Erik Skaaning, Jeffrey Staton, and Jan Teorell. 2011. "Conceptualizing and Measuring Democracy: A New Approach." *Perspectives on Politics* 9:247-67.

De Leeuw, Jan. 1983. "Models and Methods for the Analysis of Correlation Coefficients." *Journal of Econometrics* 22:113-37.

Gibson, C. L., J. Zhao, N. P. Lovrich, and M. J. Gafffney. 2002. "Social Integration, Individual Perceptions of Collective Efficacy, and Fear of Crime in Three Cities." *Justice Quarterly* 19:537-54.

Green, Samuel B. and Yanyun Yang. 2009. "Reliability of Summed Item Scores Using Sructural Equation Modeling: An Alternative to Coefficient Alpha." *Psychometrika* 74:155-67.

Hasher, L. and R. Zacks. 1984. "Automatic Processing of Fundamental Information: The Case of Frequency of Occurrence." *American Psychologist* 39:1372-88.

Hindelang, Michael J., Travis Hirschi, and Joseph G. Weis. 1981. *Measuring Delinquency*. Beverly Hills, CA: Sage.

Howell, W. C. 1973. "Representation of Frequency in Memory." *Psychological Bulletin* 80:44-53.

Jobe, J. B. and D. Hermann. 1996. "Implications of Models of Survey Cognition for Memory Theory." Pp. 193-205 in *Basic and Applied Memory Research: Volume 2: Practical Applications*, edited by D. Herrmann, M. Johnson, C. McEvoy, C. Herzog, and P. Hertel. Hillsdale, NJ: Erlbaum.

Jöreskog, Karl G. 1971. "Statistical Analysis of Congeneric Tests." *Psychometrika* 36:109-33.

Jöreskog, Karl G. 2000. *Latent Variable Scores and Their Uses*. Chicago, IL: Scientific Software International. Retrieved October 30, 2009 (http://www.ssicentral.com/lisrel/techdocs/lvscores.pdf).

Jöreskog, Karl G. and Dag Sörbom. 2001. *LISREL 8: User's Reference Guide*. Lincolnwood, IL: Scientific Software.

Jöreskog, Karl G. and Dag Sörbom. 2002. *PRELIS 2: User's Reference Guide*. Lincolnwood, IL: Scientific Software.

Kan, Man Yee and Stephen Pudney. 2008. "Measurement Error in Stylized and Diary Data on Time Use." Pp. 101-32 in *Sociological Methodology*, Vol. 38, edited by Yu Xie. Boston, MA: Blackwell.

Kankaras, Milos, Jeroen K. Vermunt, and Guy Moors. 2011. "Measurement Equivalence of Ordinal Items: A Comparison of Factor Analytic, Item Response Theory, and Latent Class Approaches." *Sociological Methods and Research* 40:279-310.

Land, Kenneth, Patricia McCall, and Lawrence Cohen. 1990. "Structural Covariates of Homicide Rates: Are There Any Invariances across Time and Space?" *American Journal of Sociology* 95:922-63.

Lei, Pui-Wa and Qiong Wu. 2012. "Estimation in Structural Equation Modeling." Pp. 164-80 in *Handbook of Structural Equation Modeling*, edited by R. H. Hoyle. New York: Guilford Press.

Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables*. Thousand Oaks, CA: Sage.

Maccoby, Eleanor, J. Johnson, and R. Church. 1958. "Community Integration and the Social Control of Juvenile Delinquency." *Journal of Social Issues* 14:38-51.

Manski, Charles N. 2004. "Measuring Expectations." *Econometrica* 72:1329-76.

Manzoni, Anna, Jeroen K. Vermunt, Ruud Luijkx, and Ruud Muffels. 2010. "Memory Bias in Retrospectively Collected Employment Careers: A Model-based Approach to Correct for Measurement Error." Pp. 39-74 in *Sociological Methodology*, Vol. 40, edited by Tim F. Liao. Washington, DC: American Sociological Association.

Matsueda, Ross L. 2010. *Seattle Neighborhoods and Crime Survey, 2002-2003*. ICPSR28701-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2010-12-10. doi:10.3886/ICPSR28701.v1.

Matsueda, Ross L. and William T. Bielby. 1986. "Statistical Power in Covariance Structure Models." Pp. 120-58 in *Sociological Methodology 1986*, edited by Nancy B. Tuma. Washington, DC: American Sociological Association.

Messick, Samuel. 1995. "Validation of Psychological Assessment: Validation of Inferences from Persons' Responses and Performances as Scientific Inquiry into Score Meaning." *American Psychologist* 50:741-49.

Miethe, Terrance D. 1991. *Testing Theories of Criminality and Victimization in Seattle, 1960-1990* [Computer File]. 2nd ICPSR version. Blacksburg: Virginia Polytechnic Institute [producer], 1991. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 1998.

Nunally, Jum C. 1967. *Psychometric Theory*. New York, NY: McGraw-Hill.

Olsson, Ulf. 1979. "Maximum Likelihood Estimation of the Polychoric Correlation Coefficient." *Psychometrika* 44:443-60.

Osgood, D. Wayne, Barbara J. McMorris, and Maria T. Potenza. 2002. "Analyzing Multiple-item Measures of Crime and Deviance I: Item Response Theory Scaling." *Journal of Quantitative Criminology* 18:267-96.

Paxton, Pamela. 2002. "Social Capital and Democracy: An Interdependent Relationship." *American Sociological Review* 67:254-77.

Piquero, Alex R., Randall MacIntosh, and Matthew Hickman. 2000. "Does Self-control Affect Survey Response? Applying Exploratory, Confirmatory, and Item Response Theory Analysis to Grasmick et al.'s Self Control Scale." *Criminology* 38:897-930.

Piquero, Alex R., Randall MacIntosh, and Matthew Hickman. 2002. "The Validity of a Self-reported Delinquency Scale: Comparison across Gender, Age, Race, and Place of Residence." *Sociological Methods and Research* 30: 492-529.

Poon, Wai-Yin and Kik-Yum Lee. 1987. "Maximum Likelihood Estimation of Multivariate Polyserial and Polychoric Correlation Coefficients." *Psychometrika* 52: 409-30.

Raudenbush, Stephen W. and Anthony S. Bryk. 2002. *Hierarchical Linear Models: Applications and Data Analysis Methods*. 2nd ed. Thousand Oaks, CA: Sage.

Raudenbush, Stephen W. and Robert Sampson. 1999. "Ecometrics: Toward a Science of Assessing Ecological Settings, with Application to the Systematic Social Observation of Neighborhoods." *Sociological Methodology* 29:1-41.

Raykov, Tenko. 2001. "Bias of Coefficient Alpha for Fixed Congeneric Measures with Correlated Errors." *Applied Psychological Measurement* 25:69-76.

R Development Core Team. 2010. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Sampson, Robert J. 1997. "Collective Regulation of Adolescent Misbehavior: Validation Results from Eighty Chicago Neighborhoods." *Journal of Adolescent Research* 12:227-44.

Sampson, Robert J. 2012. *Great American City: Chicago and the Enduring Neighborhood Effect*. Chicago, IL: University of Chicago Press.

Sampson, Robert J., Jeffrey D. Morenoff, and Felton Earls. 1999. "Beyond Social Capital: Spatial Dynamics of Collective Efficacy for Children." *American Sociological Review* 64:633-60.

Sampson, Robert J., Stephen W. Raudenbush, and Felton Earls. 1997. "Neighborhoods and Violent Crime: A Multilevel Study of Collective Efficacy." *Science* 277:918-24.

Shaw, Clifford R. and Henry D. McKay. [1942] 1969. *Juvenile Delinquency in Urban Areas*. Chicago: University of Chicago Press.

Skrondal, Anders and Petter Laake. 2001. "Regression Among Factor Scores." *Psychometrika* 66:563-576.

St. Jean, Peter K. B. 2007. *Pockets of Crime: Broken Windows, Collective Efficacy, and the Criminal Point of View*. Chicago: University of Chicago Press.

Takane, Yoshio and Jan de Leeuw. 1987. "On the Relationships between Item Response Theory and Factor Analysis of Discretized Variables." *Psychometrika* 52:393-408.

Taylor, Ralph B. 1996. "Neighborhood Responses to Disorder and Local Attachments: The Systemic Model of Attachment, Social Disorganization, and Neighborhood Use Value." *Sociological Forum* 11:41-74.

Taylor, Ralph B. 2002. "Fear of Crime, Social Ties, and Collective Efficacy: Maybe Masquerading Measurement, Maybe déj à vu All Over Again." *Justice Quarterly* 19:773-92.

Ten Berge, Joas M. F., Wim P. Krijnen, Tom Wansbeck, and Alexander Shapiro. 1999. "Some New Results on Correlation-preserving Factor Scores Prediction Methods." *Linear Algebra and its Applications* 289:311-18.

Tourangeau, Roger, Lance J. Rips, and Kenneth Rasinski. 2000. *The Psychology of Survey Response*. Cambridge, UK: Cambridge University Press.

Tversky, Amos and Daniel Kahneman. 1973. "Availability: A Heuristic for Judging Frequency and Probability." *Cognitive Psychology* 5:207-32.

Tversky, Amos and Daniel Kahneman. 1974. "Judgment Under Uncertainty: Heuristics and Biases." *Science* 185:1124-31.

Author Biographies

Ross L. Matsueda is the Blumstein-Jordan Endowed professor of Sociology at the University of Washington. His current research projects include the Seattle Neighborhoods and Crime Project, modeling life course transitions, rational choice, and desistance, and a vignette study of political protest. In the area of methods, he is working with Elena A. Erosheva and Donatello Telesca on a unimodal curve registration (UCR) approach to modeling behavioral trajectories. Their recent publications on the topic include "Modeling Criminal Careers as Departures from a Unimodal Population Age-Crime Curve: The Case of Marijuana Use," *Journal of the American Statistical Association*, 2012; and "Breaking Bad: Two Decades of Life Course Data Analysis in Criminology, Developmental Psychology, and Beyond," *Annual Review of Statistics and Its Applications*, 2014.

Kevin M. Drakulich is an assistant professor in the School of Criminology and Criminal Justice at Northeastern University. His research focuses on questions related to race, racism, and perceptions of inequality, crime, and justice. He is a 2014 recipient of the National Institute of Justice's W. E. B. Du Bois Fellowship as well as the 2014 New Scholar Award from the American Society of Criminology's Division of People of Color and Crime. His recent work appears in *Social Problems*, *Social Science Research*, and the *Journal of Interpersonal Violence*.