JAEPR, 1(*2*), 70-76

ISSN: 1930-9325

# METHODS FORUM

# Selecting Statistical Procedures for Multi-Level Data: Examining Relationships between Principal Authenticity and Teacher Trust and Engagement*

Chuang Wang and James J. Bird

*The University of North Carolina at Charlotte*

Using a case study of the relationship between principal authenticity and teacher trust and engagement, this article presents the theoretical rationale and statistical procedures to use hierarchical linear models (HLM) to analyze multi-level data. The dilemma of unit of analysis when working with multi-level data, the estimation method associated with HLM, and the conceptual fallacy during the interpretation of relationships at the individual level based upon aggregated data are also discussed.

*Keywords:* unit of analysis, conceptual fallacy, multi-level data analysis, hierarchical linear models.

## Problems Educational Researchers Encounter

It is a common practice in educational research to model treatment variables at the classroom level rather than at the student level. Students are naturally nested within classrooms in the educational system; therefore, randomly assigning students at the student level into treatment and control groups is often not practical. Data collected in educational settings are usually hierarchical in nature, for example, students nested within classrooms and classrooms nested within schools. Failure to consider the hierarchical structure of educational data could cause unreliable estimation of the effectiveness of school context or teacher quality on student learning outcomes and could misdirect educational policies and practices (Raudenbush & Bryk, 2002). Three major concerns should be taken into consideration when analyzing multi-level data: (a) unit of analysis, (b) statistical procedures, and (c) conceptual fallacy.

*Unit of Analysis.* One of the major assumptions of most statistical data analytic procedures (e.g., analysis of variance) is independent observation (i.e., students are independent from each other), but students are somewhat alike within a classroom and somewhat different across classrooms. Therefore, educational researchers face dilemma regarding the unit of analysis—whether to analyze the data at the classroom level or the student level. One way to

*This article was originally published in the *CEME Forum*, January 2011.

examine the relationships between a student-level variable and a classroom-level variable is to aggregate student-level variables within each classroom (using the classroom level as the unit of analysis) and then correlate these aggregated means with teacher-level variables. This method is limited in that it reduces the statistical power significantly due to a much smaller sample size, and it fails to consider within-classroom differences (Adams & Forsyth, 2006; Cheung, Keeves, Sellin, & Tsoi, 1990; Hand & Prather, 1986). Moreover, student-level variables aggregated to the classroom or school level are often highly correlated to each other and are likely to cause problems of multicollinearity in regression analysis (Keeves & Sellin, 1990). For example, a strong negative correlation between school average socioeconomic status and prior academic performance ($r$ = -.69) was noticed when the potential multicollinearlity concern was examined before investigating school contextual factors that could influence collective teacher efficacy (Adams & Forsyth, 2006).

Another way to examine these relationships is to conduct student-level analyses with all classroom- level variables assigned to individual students (using the student-level as the unit of analysis). This method is inappropriate for most commonly used inferential statistical tests because the residuals involving student-level variables cannot be assumed to be independently and randomly distributed. Students within a particular classroom are more like each other than they are like the students in any other classroom. In addition, the sample size for classroom-level variables is multiplied by the number of students in each classroom and therefore the estimation of the coefficients between classroom-level variables is likely to have a Type-I error. The predictive power of relationships between classroom-level and student-level variables, however, could be underestimated because they are evaluated in terms of the overall student-level variances (Keeves & Sellin, 1990).

*Statistical Procedures.* Statisticians often use variance estimates in data analysis, so it is important to consider the structure of the data (e.g., between-group and within-group variances) in statistical procedures. Many statistical models have been developed for multi-level data analyses so that relationships between both student-level and classroom-level variables can be analyzed simultaneously and at individual and group levels. For a more complete discussion of the range of statistical issues that justify the use of multi-level modeling, see Cheung et al. (1990), Hox (2002), O'Connell and McCoach (2008), and Raudenbush and Bryk (2002). Multi-level models specify effects among variables related to individuals at the individual-within-group level while group-related variables are assumed to affect parameters representing between-group differences. An application of linear regression using the multi-level approach is known as hierarchical linear modeling (HLM). A fundamental difference between HLM and simple regression lies in the estimation method. Three estimation methods are commonly used in regression analyses: (a) ordinary least squares (OLS); (b) generalized least squares (GLS); and (c) maximum-likelihood (ML).

OLS estimation gives unbiased estimates of regression coefficients associated with fixed effects (slopes/relationships between independent and dependent variables) by minimizing the sum of squared residuals and assumes independent observation and homoscedasticity. GLS estimation is appropriate in multi-level analyses because it allows the residuals to be autocorrelated and heteroscedastic and uses the estimated error covariance from these residuals to estimate both the fixed and random effects through an iterative procedure (Singer & Willett, 2003). ML estimation method, maximizing the joint probability of observing a particular sample of data by estimating both fixed effects and the random effects (variance components) of

population parameters, is the default estimation method in the HLM software. The use of a multi-level approach allows the researcher to examine relationships among variables within and between classrooms (Goddard, Tschannen-Moran, & Hoy, 2001; Stapleton, 2006).

*Conceptual Fallacy*. A conceptual fallacy occurs when researchers make interpretations of relationships at the individual level based upon aggregated data (Hox, 2002). For example, Hand and Prather (1986) tested the impact of aggregation on the relationship between grade point average (GPA) and high school average scores of 45,000 students within 30 colleges and noticed that the regression coefficients having opposite signs at the individual level (e.g., .22) and at the college level (e.g., -.02). The notion of aggregation bias is also known as "ecological fallacy" when researchers draw conclusions at one level while analyzing the data at another level (Robinson, 1950).

Using an example of the application of both simple regression and HLM methods to examine the relationships between a principal's authenticity and a teacher's trust and engagement, this paper calls for careful selection of statistical methods and best practices for research.

## CONTEXT

As the leader of a school, a principal's behavior and management skills have an impact on teachers' job satisfaction, efficacy, trust in the organization, and engagement levels (Bird, Wang, Watson, & Murray, 2009; Bogler, 2001; Dipaola & Tschannen-Moran, 2001; Hipp & Bredeson, 1995; Koh, Steers, & Terborg, 1995). Hence, there is a need to examine relationships between a principal's authenticity and a teacher's trust and engagement.

In the context of education, "authentic leader" refers to one who self-regulates his/her behaviors through self-reflection and fosters balanced processing of information and relational transparency with teachers (Walumbwa, Avolio, Gardner, Wernsing, & Peterson, 2008). Trust is defined as "the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other party will perform a particular action important to the trustor, irrespective of the ability to monitor or control that party" (Mayer, Davis, & Schoorman, 1995, p. 712). The concept of engagement is defined as employees' commitment and the positive emotions they experience (Ostrem & Wheeler, 2006). Extensive research has been conducted to examine each of these constructs separately in both business and educational settings (e.g., Dipaola & Tschennan-Moran, 2001; George, Sims, McLean, & Mayer, 2007; Hoy, Hoy, & Kurz, 2008; Ostren & Wheeler, 2006; Tschannen-Moran & Hoy, 2000). Very few studies, however, have considered employees' trust and engagement and the employer's leadership style simultaneously (Yammarino, Dionne, Schriesheim, & Dansereau, 2008). While studying teacher ratings of the principal authenticity, it is difficult to assume that teachers within a particular school are independent from each other because their ratings might be somewhat similar while evaluating the performance of the same person (the principal). As a result, a series of research projects were conducted to investigate the relationships between principal authenticity and teacher trust and engagement using both single-level and multi-level approaches at various stages of the project.

# SYNOPSIS OF AUTHENTIC LEADERSHIP RESEARCH*

## Method

The authors conducted two studies to investigate the relationships between principal authenticity and teacher trust and engagement. Study I consisted of 156 teachers and 22 principals. A second study was conducted because the first used a small sample and the between-school variances had an average intra-class correlation (ICC) coefficient across all dependent variables of only 4% (Bird et al., 2009). Study II consisted of 917 teachers and 60 principals. The demographic information of the participants in Study I and Study II was similar. The collection of additional data and an increase in between-school variances by adding more school districts resulted in a re-examination of the potential to employ HLM by calculating the percentage of variance of each construct at the teacher-level and the school-level.

The instrument selected to measure authenticity was a 16-item questionnaire that consisted of four subscales: (a) self-awareness, (b) relational transparency, (c) internalized moral reasoning, and (d) balanced processing. The questionnaire was developed and validated by Walumbwa et al. (2008). Trust was measured by a 32-item survey that was composed of three subscales: (a) trust the principals, (b) trust co-workers, and (c) trust the school. The survey was developed and validated by Ferris and Travaglione (2003). Finally, engagement was measured by the 12 items developed by Buckingham and Coffman (1999) related to employee engagement as a result of thousands of focus groups and interviews. The survey was validated by Harter, Schmidt, and Keyes (2002). Although all selected scales were validated in previous studies, no study had examined the multi-level structural validity (Wang & Bird, 2011).

## Results

In Study I, independent samples *t*-tests failed to identify any statistically significant differences between elementary, middle, and high schools. As a result, school-level was not used in further analyses. MANOVA also failed to note any statistically significant differences due to gender, ethnicity, and highest education level achieved. Therefore, the data were considered as from a single group in Pearson correlation. Teacher ratings of principal authenticity was found to be statistically and significantly related to teachers trust.

Based upon the results from Study I, group differences related to demographic information were not considered in Study II. Similar results were found with regard to the relationships among authenticity, trust, and engagement. Comparisons of the mean scores and standard deviations of the first-order and second-order constructs of authenticity between principal self-report and teacher ratings also revealed that principals rated themselves significantly higher than teachers rated them. Follow-up multiple comparisons of multivariate analysis of variance with Tukey's Honestly Significant Difference (HSD) method showed that principals rated themselves significantly higher than teachers rated them on all first-order and second-order constructs of authenticity.

Confirmatory factor analysis was employed to examine the measurement validity of the instruments. The measurement models for teacher trust and engagement fit the data very well. The authenticity measurement model was confirmed at the teacher level (teachers' ratings of their principal's authenticity) but not at the principal level (principal self-report of authenticity).

ICC showed that 14% of the variance in teacher ratings of their principal's authentic leadership, 15% of the variance in teacher ratings of trust in their principals, and 6% of the variance in teacher ratings of engagement existed between schools. As a result, the multi-level structural equation modeling approach was adopted to examine the relationships between these constructs.

HLM analyses in Study II again suggest that school building principal leadership authenticity is positively and statistically related to their teachers' levels of trust and engagement. This is consistent with Study I results; however, the aggregated teacher ratings of principal authenticity had a significant impact on the relationships between teacher trust and principal authenticity and between teacher engagement and principal authenticity, respectively. This suggested a potential aggregated bias in Study I. The correlation coefficients for the relationships were underestimated although the trend of associations stayed the same. At the teacher level, teachers were more than likely to have lower levels of trust and lower levels of engagement when they rated their principal's authenticity low. Moreover, teacher level of engagement was no longer statistically related to their ratings of principal authenticity when their trust in their principal was controlled. That is, for teachers who had the same level of trust in their principal, their levels of engagement were not related to their ratings of the principal's authenticity anymore.

The larger sample size also allowed the researchers to partition the participants into three groups along a continuum of principal-teacher agreement concerning the principal authenticity. The three groups were composed of schools whose principals over-estimated their own authenticity vis-à-vis how their teachers rated their authenticity, schools whose principals and teachers agreed with each other concerning principal authenticity, and schools whose principals under-estimated their own authenticity relative to how their teachers rated their authenticity. The over-estimated schools had significantly less teacher trust and engagement levels than the under-estimated schools.

## IMPLICATIONS FOR RESEARCH

The larger sample size and the multi-level structure of the data allowed for the examination of structural validity of the surveys to measure authenticity, trust, and engagement by considering between-school and within-school variances. The larger sample size also provided opportunities to measure differences in teacher trust and engagement along a continuum of the differences between teacher ratings and principal self-ratings of authenticity. Although statistically significant relationships were found in both studies between authenticity, trust, and engagement, the ML estimation method in HLM revealed additional information: the average teacher ratings of principal authenticity had a statistically significant impact on the relationships between individual teacher trust and their ratings of principal authenticity and between individual teacher engagement and their ratings of principal authenticity. While the multi-level SEM confirmed the relationships between authenticity, trust, and engagement between and within schools, HLM analyses suggested that most of the variance was between authenticity and trust. When teachers had the same level of trust in the principal, their levels of engagement were no longer related to their ratings of principal authenticity. This result highlights the importance of principal authenticity. Teachers trust their principals more when the principals are self-aware, relationally transparent, use balanced processing, and act with integrity.

The use of multi-level analyses also suggested the limitations of single-level analyses. For example, the between-school variances of authenticity, trust, and engagement in Study I was ignored with the Pearson correlation method. The statistically significant impact of average ratings at the school level on the relationships at individual teacher levels also suggested a potential aggregation bias in the single-level analyses. Researchers interested in examining relationships between variables that are measured in multi-levels are encouraged to select a multi-level approach whenever possible (Cheung et al., 1990).

## REFERENCES


Adams, C. M., & Forsyth, P. B. (2006). Proximate sources of collective teacher efficacy. *Journal of Educational Administration, 44,* 625-642.

Bird, J. J., Wang, C., Watson, J., & Murray, L. (2009). Relationships among principal authentic leadership and teacher trust and engagement levels. *Journal of School Leadership, 19,* 153-171.

Bogler, R. (2001). The influence of leadership style on teacher job satisfaction. *Educational Administration Quarterly, 37,* 662-683.

Cheung, K. C., Keeves, J. P., Sellin, N., & Tsoi, S. C. (Eds.) (1990). The analysis of multilevel data in educational research: Studies of problems and their solutions. *International Journal of Educational Research, 14*(3).

Dipaola, M., & Tschannen-Moran, M. (2001). Organizational citizenship behavior in schools and its relationship to school climate. *Journal of School Leadership, 11*, 424-447.

George, B., Sims, P., McLean, A., & Mayer, D. (2007). Discovering your authentic leadership. *Harvard Business Review, 85*, 129-138.

Goddard, R. D., Tschannen-Moran, M., & Hoy, W. K. (2001). A multilevel examination of the distribution and effects of teacher trust in students and parents in urban elementary schools. *The Elementary School Journal, 102*, 3-17.

Ferres, N., & Travaglione, T. (2003, December). The development and validation of the workplace trust survey: Combining qualitative and quantitative methodologies. Paper presented at the 9th Asia Pacific Researchers in Organization Studies Conference, Oaxaca, Mexico.

Hand, C. A., & Prather, J. E. (1986, August). *The impact of aggregation bias upon the interpretation of test scores across schools.* Paper presented at the Annual Meeting of the American Statistical Association, Chicago, IL.

Harter, J. K., Schmidt, F. L., & Keyes, C. L. (2002). Well-being in the workplace and its relationship to business outcomes: A review of the Gallup studies. In C. L. Keyes & J. Haidt (Eds.), Flourishing: The positive person and the good life (pp. 205-224). Washington D. C.: American psychological Association.

Hipp, K. A., & Bredeson, P. V. (1995). Exploring connections between teacher efficacy and principal's behavior. *Journal of School Leadership, 5*, 136-150.

Hox, J. J. (2002). *Multilevel analysis: Techniques and applications*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Hoy, A. W., Hoy, W. K., & Kurz, N. M. (2008). Teachers' academic optimism: The development and test of a new construct. *Teaching and Teacher Education, 24*, 821-835.

Keeves, J. P., & Sellin, N. (1990). Some problems of analysis. *International Journal of Educational Research, 14,* 219-224.

Koh, W. L., Steers, R. M., & Terbog, J. R. (1995). The effects of transformational leadership on teacher attitudes and student performance in Singapore. *Journal of Organizational Behavior, 16*, 319-333.

Mayer, R., Davis, J., & Schoorman, F. (1995). An integrative model of organizational trust. *Academy of Management Review, 20*, 709-734.

O'Connell, A. A., & McCoach, D. B. (2008). *Multilevel modeling of educational data.* Charlotte, NC: Information Age Publishing, Inc.

Ostrem, L., & Wheeler, D. (2006, October) *Engagement and trust as outcomes of servant leadership.* Paper presented at the Gallup Leadership Summit, Washington, DC.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods.* Thousand Oaks, CA: Sage.

Robinson, W. S. (1950). Ecological correlations and the behavior of individuals. *American Sociological Review, 15,* 351-357.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence.* New York, NY: Oxford University Press.

Stapleton, L. M. (2006). Using multilevel structural equation modeling techniques with complex sample data (pp. 345-383). In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course.* Charlotte, NC: Information Age Publishing.

Tschannen-Moran, M., & Hoy, W. K. (2000). A multidisciplinary analysis of the nature, meaning, and measurement of trust. *Review of Educational Research, 70,* 547-593.

Walumbwa, F. O., Avolio, B. J., Gardner, W. L., Wernsing, T. S., & Peterson, S. J. (2008). Authentic leadership: Development and validation of a theory-based measure. *Journal of Management, 34,* 89-126.

Wang, C., & Bird, J. J. (2011). Multi-level modeling of principal authenticity and teachers' trust and engagement. *Academy of Educational Leadership Journal, 15,* 125-147.

Yammarino, F. J., Dionne, S. D., Schriesheim, C. A., & Dansereau, F. (2008). Authentic leadership and positive organizational behavior: A meso, multi-level perspective. *The Leadership Quarterly, 19,* 696-707.