Group Dynamics: Theory, Research, and Practice
1998, Vol. 2, No. 4, 313–320

1089-2699/98/$3.00

Using Hierarchical Linear Modeling to Analyze Grouped Data

John B. Nezlek
College of William & Mary

Linda E. Zyzniewski
Virginia Commonwealth University

This article discusses how to use a random coefficient modeling technique known as hierarchical linear modeling to analyze data collected within groups. The article describes how to use this technique to examine group- and individual-level phenomena, including examination of how individual-level relationships vary as a function of group characteristics. A comparison of hierarchical linear modeling with more traditional, ordinary-least-squares techniques and a presentation of how to implement analyses to test specific hypotheses are included. Included are brief discussions of pertinent issues such as the impact of different centering options, the analysis of categorical variables, distinctions between random and fixed effects, and balanced and unbalanced designs.

A wide variety of psychologically meaningful phenomena occur within groups, and analyzing data describing group phenomena requires the use of techniques that take into account the multiple levels of analysis that characterize such data. For example, the simplest group design involves two levels of analysis: that of the group, and that of individuals within groups. This article describes how to use a random coefficient modeling technique known as *hierarchical linear modeling* to analyze data collected within groups. The article describes how to use this technique to examine group- and individual-level phenomena, including how to determine whether individual-level relationships vary as a function of group characteristics. In addition, Pollak (1998) describes the results of an analysis of grouped data using hierarchical linear modeling, and Moritz and Watson (1998) discuss general levels-of-analysis issues in groups research.

For present purposes, *groups* are defined as collections of individuals that occur either naturally, such as work groups in organizations (e.g., Podsakoff, Ahearne, & MacKenzie, 1997), or arbitrarily, such as groups created in experiments (e.g., Stasson & Bradshaw, 1995). For present purposes, the basis used to create groups is not of primary importance. Although the techniques described in this article provide numerous advantages over more traditional techniques, they do not provide dramatically different information about the internal, external, or ecological validity of a study. Such attributes are more a function of how studies are designed than of how data are analyzed.

For present purposes, *group-level phenomena* are defined as outcomes that exist only at a group or aggregate level. For example, researchers have examined differences among groups in the completion of collective tasks (Leary & Forsyth, 1987). In contrast, individual-level phenomena are examined using data describing individuals. For example, researchers have examined the relationships between how well individuals within a group solve problems and individual differences such as attributional biases (Leary & Forsyth, 1987). Finally, researchers may also be interested in how individual-level phenomena vary as a function of group characteristics. For example, does the relationship between attributional biases and problem-solving ability vary across different types of groups?

Traditionally, data collected within groups have been analyzed using different types of ordinary-least-squares (OLS) techniques. Within such a framework, relationships concerning group means can be examined using means aggregated for each group, in essence, treating

John B. Nezlek, Department of Psychology, College of William & Mary; Linda E. Zyzniewski, Department of Psychology, Virginia Commonwealth University.

The logic of the analyses presented in this article follows closely that presented by Bryk and Raudenbush (1992), and every effort has been made to acknowledge this. We are grateful to Mathilda duToit of Scientific Software International for her comments on an earlier version of this article.

Correspondence concerning this article should be sent to John B. Nezlek, Department of Psychology, College of William & Mary, P.O. Box 8795, Williamsburg, Virginia 23187-8795. Electronic mail may be sent to john.nezlek@wm.edu.

the group as the unit of observation. Differences among groups can be compared using analyses of variance (ANOVAs) in which groups are the units of observation, and relationships between group means can be examined using correlational techniques in which groups are also the unit of observation.

Such analyses are straightforward and have been very popular. Nonetheless, OLS analyses using aggregated group means typically ignore at least two important differences that may exist among groups: group size and the consistency of responses of members within groups (hereafter referred to as *reliability*). That is, smaller and larger groups and more and less reliable groups contribute equally to group-level analyses. Within the OLS framework, such differences can be incorporated using what is called *weighted-least-squares* (WLS) analyses in which observations (e.g., group means) are weighted by whatever adjustment is deemed appropriate: numbers of observations, reliability, and so on.

Using OLS techniques to analyze individual-level phenomena such as the relationship between two variables is a bit more complicated. Some researchers have ignored the fact that their data are organized within groups and have analyzed their data only at the individual level. Although appealing in their simplicity, such analyses are fundamentally incorrect because they confound individual- and group-level effects and provide potentially inaccurate estimates of individual-level relationships. For example, Cashin, Presley, and Meilman (1998) used individual-level analyses to compare the alcohol-related beliefs and behaviors of United States collegians who were members of fraternities and sororities to the beliefs and behaviors of students who were not. The individual-level differences they reported between members and nonmembers may have varied widely across the schools and Greek organizations in their sample. For example, there may have been schools or organizations for which such differences were smaller, larger, or even reversed compared to the differences they reported.

Within an OLS framework, such individual-level relationships can be examined by analyzing summary statistics (such as correlations or subgroup mean differences) describing the individual-level relationships within each group. Individual-level relationships can then be examined controlling for group-level differences in means.

OLS-based multilevel analyses have been the mainstay of social science researchers, and research using such analyses has been very productive.[1] Nonetheless, advances in statistical theory and computational algorithms combined with the availability of inexpensive high-speed data processing have made a set of techniques generally referred to as *random coefficient modeling* readily available. Moreover, these techniques provide numerous advantages over multilevel OLS techniques, particularly for the estimation of parameters.

For example, although WLS analyses of group- and individual-level relationships can incorporate different weights such as group size and reliability, such analyses are essentially OLS procedures. Moreover, although WLS may present some advantages over standard OLS, maximum-likelihood procedures such as those discussed here provide more accurate parameter estimates than OLS-based procedures; these issues are discussed in Kenny, Kashy, and Bolger (1998).

Hierarchical Linear Modeling

Hierarchical linear modeling (HLM) is a random coefficient modeling technique that can be used to analyze data collected within groups. The term *hierarchical* refers to the fact that sets of observations are treated as hierarchically nested within other sets. For example, data describing individuals are analyzed as nested within the groups to which the individuals belong. Although there are different implementations of hierarchical linear modeling, in this article we discuss the specific implementation known as HLM (Bryk & Raudenbush, 1992). HLM was chosen because the algorithms are well understood and documented and because the software (HLM4; Bryk, Raudenbush, & Congdon, 1998) is readily available and easily accessible. Regardless, preliminary comparisons suggest that differences in the algorithms used by different hierarchical linear modeling programs do not produce meaningfully different results (Bryk & Raudenbush, 1992, p. 233). For a discussion of the importance of distinguishing

[1] In this article, OLS analyses are assumed to be multilevel analyses of some kind.

HLM, the model and program, from hierarchical linear modeling, the technique, see de Leeuw and Kreft (1995).

Some of the important differences between HLM and OLS are that HLM uses a combination of maximum-likelihood and Bayesian procedures to estimate parameters, and in HLM, parameter estimates from different levels of analysis are not independent. For example, the reliability of responses within groups (an individual-level datum) is used to estimate the variance of group-level parameters. The more reliable the responses are in a group, the greater the weight assigned to the group mean in estimating variances. The less reliable the responses are in a group, the smaller the weight assigned to the population mean in estimating variances. This convention, known as *precision weighting,* is part of the procedure that HLM uses to produce "empirical Bayes estimates" (EBEs) of parameters. These EBEs allow HLM to separate fixed (or true) and random (or error) parameter variance, whereas these two sources of variance are combined in OLS analyses. These procedures are described in detail in Bryk and Raudenbush (1992, pp. 32–59).

Using HLM, data are analyzed with a series of regression-like hierarchically nested models in which parameters from one level of analysis are analyzed at the next level of analysis. Theoretically, hierarchical models can have any number of levels; however, for simplicity's sake, this discussion focuses on two-level models. In these models, data describing individuals are Level 1 observations (or units of analysis), and these data are treated as nested within groups, which are Level 2 units of analysis. Application of the technique is described using the nomenclature developed by Bryk and Raudenbush (1992).

The first step in any HLM analysis is the specification of the Level 1 model, in this instance, a model describing individuals' responses. The simplest Level 1 model is

$$y_{ij} = \beta_{0j} + r_{ij}.$$

In such a model, β_{0j} is a random coefficient representing the mean of y for group j (across the i individuals appearing in each group), r_{ij} represents the error associated with each observation, and the variance of r_{ij} constitutes the Level 1 residual (or error) variance.

An important difference between this model and the model underlying a traditional OLS analysis relying on aggregated means is the r_{ij} term. In HLM, the r_{ij} term is part of the precision-weighting procedure used to estimate variance components. In contrast, in OLS, variance estimates from different levels of analysis are independent, such as the independence of the sums of squares for between- and within-subjects factors in ANOVA.

The next step is the specification of a Level 2 model, in this instance, a model describing group-level effects. The simplest Level 2 model is

$$\beta_{0j} = \gamma_{00} + u_{0j}.$$

In this model, γ_{00} represents the mean of the group means (β_{0j}) from the Level 1 model; u_0^j represents the error of β_{0j}, the deviation from the grand mean; and the variance of u_0^j constitutes the Level 2 residual variance. Bryk and Raudenbush (1992) referred to a set of two such models as "totally unconditional" because individual responses (ys) and group means (β_{0j}s) are not being modeled as a function (conditionally) of other individual- or group-level variables.

Although such a totally unconditional analysis does not test any hypotheses about the relationships between variables at either level, the results it provides are very important. First, the total variance of a measure is divided into individual- and group-level components. This suggests where "the action is" in the data and may indicate the level of analysis at which introducing additional terms is likely to account for variance. Directly related to this conceptually, the analysis also provides an estimate of the consistency of responses within groups (the average reliability). Second, the variance estimates provided by such an analysis serve as the basis for estimating the strength of individual- and group-level effects.

Before discussing other details of HLM analyses of grouped data, it is important to note that in HLM, tests of the statistical significance of effects (are effects nonzero?) are based on fixed parameters, whereas estimates of the strength of effects (variance accounted for) are based on random parameters. In contrast, in OLS analyses, significance tests and estimates of effect strength are based on the same parameter estimates, that is, the sums of squares for effects.

Group-level effects are examined at Level 2 by modeling of the variability of β_{0j}, the coefficient from the Level 1 model representing the group mean. Such analyses are referred to as "means as outcomes" analyses because the outcome of the Level 1 model is a mean. For example, to test the hypothesis that the mean scores of male and female groups differed, the following Level 2 model would be examined:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(\text{SEX}) + u_{0j}.$$

In such a model, SEX could be a dummy (0, 1) or contrast (1, −1) coded variable. If γ_{01} is significantly different from 0, the means of female and male groups are significantly different. For example, assume that an analysis using a contrast-coded variable (female groups assigned a value of 1) produced an intercept (γ_{00}) of 4 and a γ_{01} of 2. The mean score for female groups would be 6, and for male groups it would be 2. If γ_{01} is significantly different from 0, the difference between 2 and 6 is statistically significant.

In HLM, the significance of the SEX effect is tested in the fixed effect portion of the model. To evaluate the strength of the SEX effect, the residual Level 2 variance from the unconditional model needs to be compared to the residual variance of the model including SEX at Level 2. The percentage by which the residual variance from the first model is reduced by the second is a measure of the percentage of variance in group means accounted for by group sexual composition. This procedure is discussed in Bryk and Raudenbush (1992, p. 65).

Individual-level relationships or effects are examined similarly, albeit using various Level 1 models. Assume a study in which the prime dependent measure is the effort individuals contribute to a group task, and individual differences in sociability are measured as a possible predictor or covariate. To determine if effort and sociability are related, the following Level 1 model could be used:

$$y_{ij} = \beta_{0j} + \beta_{1j}(\text{SOC}) + r_{ij}.$$

In such a model, β_{1j} is a random coefficient representing the relationship between effort (y) and sociability for group j (across the i individuals appearing in each group), r_{ij} represents the error associated with each observation, and the variance of r_{ij} constitutes the Level 1 residual variance for this model.

To examine whether this relationship is significantly different from 0 across the groups in the study, the following Level 2 model is examined:

$$\beta_{0j} = \gamma_{00} + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + u_{1j}.$$

Such models are described as unconditional at Level 2 because Level 1 parameters are not modeled as a function of other variables. In such models, the significance of γ_{10} indicates whether, on average, the relationship between sociability and effort is 0. The strength of the relationship between effort and sociability is determined by comparing the residual Level 1 variance from an unconditional model with the residual Level 1 variance of the model with sociability at Level 1. The percentage by which the residual variance from the first model is reduced by the second is a measure of the percentage of variance in effort accounted for by sociability. This procedure is discussed in Bryk and Raudenbush (1992, p. 70). Note that this model also provides a reliability estimate of β_{1j}, the coefficient representing the relationship between sociability and effort.

To examine whether individual-level relationships vary as a function of group-level characteristics, the variability of the β_{1j} coefficients is analyzed. For example, is the strength of the effort–sociability relationship the same for female and male groups? Assuming that SEX was coded as in the previous example, such a hypothesis would be tested by the significance of the γ_{11} coefficient in the following model:

$$\beta_{1j} = \gamma_{10} + \gamma_{11}(\text{SEX}) + u_{1j}$$

It is important to note that there may be predictable variability in a β_{1j} coefficient even if the average β_{1j} coefficient (γ_{10} from the unconditional Level 2 model, above) is not significantly different from 0. For example, the mean coefficient for female groups may be positive, whereas the mean coefficient for male groups may be negative and different from the female coefficient, and such a difference could exist when the mean coefficient across all groups is not significantly different from 0.

Analyses in which the covariation between variables is modeled at level 1 and then analyzed at level 2 are referred to as "slopes as outcomes" analyses because the outcome of the Level 1 model is a slope (or regression coefficient). Although HLM provides important advantages over OLS for analyzing means, the advantages it provides for analyzing slopes are more pronounced. The accuracy of parameters describing slopes depends on the joint reliability of outcomes and predictors, whereas the accuracy of parameters describing means depends on the reliability of the outcome only. HLM incorporates reliabilities into parameter estimates, and the advantage it provides over OLS is a direct function of the number of potential sources of unreliability in an analysis.

The previous description has assumed a two-level design in which individuals were nested within groups, which is a fairly typical design. There are times, however, when multiple observations or ratings are made of group members (by each other or outside observers), and hypotheses of interest concern these ratings. Such data can be analyzed in HLM through the use of three-level models in which ratings are treated as nested within individuals and individuals are nested within groups. The basic, unconditional model for the analysis of response y for which there are i observations about j persons in k groups is as follows:

$$y_{ijk} = \pi_{0jk} + e_{ijk}$$

$$\pi_{0jk} = \beta_{00k} + r_{0jk}$$

$$\beta_{00k} = \gamma_{00_0} + u_{00k}.$$

If one understands that π_{0jk} refers to the average rating for the jth person in the kth group and that the variance of e_{ijk} represents the random variance of ratings, the logic of such an analysis is identical to the logic of the two-level model, and the same types of parameters are estimated using the same procedures. A three-level model such as the preceding one decomposes the total variance of y into observer, individual, and group components, produces a reliability coefficient for ratings, and provides the basis for evaluating the effect sizes found in other analyses. Predictors can be included at any level, for example, a zero-mean-centered, contrast-coded variable representing observer versus member could be included at Level 1 to adjust mean ratings of group members for differences in member and observer ratings.

Considerations When Using HLM

There are numerous subtleties and nuances that need to be considered when analyzing data using HLM, and space does not permit consideration of many of them. Nonetheless, brief consideration of some issues relevant to the analysis of grouped data is possible. More detailed discussions of these and other relevant issues can be found in Bryk and Raudenbush (1992) and Kreft and de Leeuw (1998). In addition, analysts interested in interactive hypotheses should be certain they are familiar with different ways of testing and interpreting interactions within a regression framework (Aiken & West, 1991).

When one is examining individual-level relationships, decisions need to be made about how individual-level predictors (or covariates) are centered. *Centering* refers to the basis used to estimate slopes and intercepts. When using continuous variables as individual-level predictors, many researchers will probably want to "group mean center" predictors. This means that the slope and intercept for each group are based on deviations from each group's mean on the predictor variable, and the intercept for each group represents the predicted score on the dependent measure for a person at the group mean on the predictor. This eliminates the influence on slope and intercept estimates of group differences in means of predictors. Other options include zero-mean centering, in which case slopes and intercepts are based on deviations from 0 (something to consider when using contrast-coded variables as predictors), and grand mean centering, in which case slopes and intercepts are based on deviations from the grand mean. Regardless, analysts need to be careful when selecting centering options because different options can produce dramatically different slopes and intercepts.

Similarly, it is important for analysts to keep in mind how different ways of coding categorical variables influence the interpretation of intercepts at any level in a model. Although significance tests of slopes *at a particular level* *of analysis* do not vary across different types of coding (e.g., dummy vs. contrast), the meaning of intercepts and *the results of tests of these parameters at higher levels* will vary. For example, assume a Level 1 model in which individual sex is dummy-coded 0 for women and 1 for men. If this predictor is zero-mean centered, the intercept will represent the mean for women, and the slope will represent the sex difference. In contrast, if a contrast-coded variable (-1 for women, 1 for men) is zero-mean centered, the intercept will represent the group mean adjusted for sex differences. The intercepts produced by these two models are different, and the results of Level 2 (group-level) analyses of these intercepts will, accordingly, not be the same.

The previous discussion has tacitly assumed that the dependent (y) variables of interest have been continuous measures, such as scores on scales. Nonetheless, group researchers also study dichotomous outcomes such as verdicts, decisions to adopt Plan X or Y, and so on. The analysis of dichotomous measures is complicated by the fact that distributions of dichotomous measures vary as a function of the mean proportion of a measure, and this covariation between mean and standard deviation violates the assumptions of many analytic techniques. Often, this problem is circumvented by transforming group proportions to eliminate such covariation.

In HLM, dichotomous outcomes can be properly analyzed using special Level 1 models. Assume a study in which individuals within groups must choose one of two options (X or Y) and the dependent measure is operationalized as whether or not an individual chose Option X. Individual-level effects could be examined with the following Level 1 Bernoulli model with $n = 1$:

$$\text{Prob}\,(y = 1 | \beta_0) = \phi.$$

The coefficients from the individual-level analyses, in this case the log-odds of adopting X, could be analyzed at the group level as a function of group-level variables. A description of the use of HLM to analyze categorical outcomes (binary and otherwise) can be found in Bryk, Raudenbush, and Congdon (1996, pp. 117–159).[2]

When planning and conducting HLM analyses, it is important to keep in mind that HLM and other similar programs were designed to meet the needs of researchers primarily in the fields of education and sociology, and the needs of such researchers shaped choices about how the program estimates parameters. For example, HLM analyzes covariance, not correlation, matrices and produces estimates of unstandardized, not standardized, coefficients.[3] Such a convention is critical for researchers for whom it is important to estimate coefficients in original metrics, for example, increases in reading scores associated with increased days in school, and such researchers tend to be interested more in unstandardized than in standardized coefficients. For other researchers, including many psychologists, original metrics are not particularly important. Differences in the number of points on attitude scales may reflect differences in researchers' preferences more than differences in the constructs being measured. Consistent with these preferences, such researchers tend to be interested more in standardized than in unstandardized coefficients.

Researchers who are primarily interested in, and familiar with analyses that rely on, standardized coefficients need to be careful when conducting and evaluating HLM analyses. For example, in HLM the equality of coefficients or sets of coefficients can be tested, and, unlike tests of standardized regression coefficients, differences in the variances of predictor variables contribute to the results of tests of unstandardized coefficients. Nevertheless, such differences can be eliminated by standardization of measures.

As noted previously, HLM provides separate estimates of fixed and random variances. The program uses sample (group, in terms of our ongoing example) sizes, means, and reliabilities to do this, and aspects of these procedures have implications for the similarity of the results of HLM and OLS analyses for balanced and unbalanced designs. For balanced designs, HLM analyses and OLS analyses produce identical tests of the statistical significance of group-level (Level 2) effects such as the sex effect in the previous examples. Recall that in HLM such tests are based on estimates of fixed parameters.

---

[2] More comprehensive procedures to analyze categorical data will be available in HLM Version 5, to be released in the near future.

[3] There are also other, more purely statistical reasons for analyzing covariance rather than correlation matrices.

HLM and OLS analyses tend to differ, however, in their estimates of the strength of these effects, expressed as the variance in group coefficients (means or slopes) accounted for by a group-level variable. These differences are due in part to the fact that HLM incorporates Level 1 reliability estimates in the estimation of random variance, whereas OLS does not. If all groups were equally sized and equally reliable, the results of OLS and HLM analyses would be identical.

The separation of fixed and random variance also means that HLM can treat effects as fixed or random. The distinction between fixed and random effects may be puzzling for many analysts who are familiar only with techniques that test fixed effects, and the present discussion considers this important issue only superficially. Nonetheless, one of the important distinctions between random and fixed effects concerns the "inference space" of an effect. Tests of fixed effects make inferences about differences involving the *specific* effects being studied, whereas tests of random effects assume that the specific effects being studied were sampled from a larger population of effects. For a brief discussion of this issue, see Littell, Milliken, Stroup, and Wolfinger (1996, pp. 230–234).

Decisions to specify effects as fixed or random can be made on three bases: theoretical, empirical, or practical. Theoretically, an effect can be defined as fixed if an analyst is interested only in a specific effect, such as individual-level differences in decisions of Black and White jurors. If, however, that same analyst is interested in making inferences about individual-level racial differences, the Level 1 effect representing Black–White should be random because Black–White differences are only one type of racial difference. Empirically, effects can be fixed if an analysis determines that the random variability of an effect is not significantly different from 0, and HLM provides such tests by default. Also, effects may need to be fixed if a model estimating many parameters has difficulty converging. Finally, it is important to note that group-level differences in individual-level effects can be examined even when individual-level effects are fixed. Such analyses concern what is called *nonrandom variability in slopes.*

Regardless of the basis used to fix an effect, analysts need to keep in mind that fixing an effect changes the results of significance tests of the effect because of a change in the breadth of inference. More subtlely, analysts need to recognize that because HLM fits covariance matrices, fixing an effect changes some of these covariance matrices, and estimates of parameters of other effects in a model may also change. Although general guidelines do not, by definition, apply equally well to all situations, most analysts should probably think first of their effects as random and fix them only if there are compelling reasons to do so.

When evaluating the relative advantages of HLM and OLS analyses, it is important to keep in mind that, as a general rule, differences in the results of HLM and OLS analyses vary directly as a function of differences among the numbers of observations in groups (sample sizes) and the group-level reliabilities of these observations. If a study has many Level 2 units (groups) and many Level 1 observations (individuals) in each group, the results will be more similar than if either of these is smaller. As sample sizes approach infinity, multilevel OLS analyses and the maximum-likelihood Bayesian analyses that constitute HLM provide increasingly similar results. Regardless, it would seem that given the realities of social science research on groups, HLM provides numerous advantages over comparable multilevel OLS analyses, and researchers would be well advised to use such techniques.

References

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions.* Newbury Park, CA: Sage.

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models.* Newbury Park, CA: Sage.

Bryk, A. S., Raudenbush, S. W., & Congdon, R. (1996). *HLM: Hierarchical linear modeling with the HLM/2L and HLM/3L programs.* Chicago: Scientific Software.

Bryk, A. S., Raudenbush, S. W., & Congdon, R. (1998). *HLM: Version 4.03a.* Chicago: Scientific Software.

Cashin, J. R., Presley, C. A., & Meilman, P. W. (1998). Alcohol use in the Greek system: Follow the leader? *Journal of Studies on Alcohol, 59,* 63–70.

de Leeuw, J., & Kreft, I. G. G. (1995). Questioning multilevel models. *Journal of Educational and Behavioral Statistics, 20,* 171–189.

Kenny, D. A., Kashy, D. A., & Bolger, N. (1998). Data analysis in social psychology. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 1, pp. 233–265). Boston: McGraw-Hill.

Kreft, I. G. G., & de Leeuw, J. (1998). *Introducing multilevel modeling*. Newbury Park, CA: Sage.

Leary, M. R., & Forsyth, D. R. (1987). Attributions of responsibility for collective endeavors. In C. Hendrick (Ed.), *Review of personality and social psychology: Group processes* (Vol. 8, pp. 167–188). Newbury Park, CA: Sage.

Littell, R. C., Milliken, G. A., Stroup, W. W., & Wolfinger, R. D. (1996). *SAS system for mixed models*. Cary, NC: SAS Institute.

Moritz, S. E., & Watson, C. B. (1998). Levels of analysis issues in group psychology: Using efficacy as an example of a multilevel model. *Group Dynamics, 2,* 285–298.

Podsakoff, P. M., Ahearne, M., & MacKenzie, S. B. (1997). Organizational citizenship behavior and the quantity and quality of work group performance. *Journal of Applied Psychology, 82,* 262–270.

Pollak, B. N. (1998). The benefits of HLM in modeling relationships between groups and individuals: A cross-level demonstration. *Group Dynamics, 2,* 299–311.

Stasson, M. F., & Bradshaw, S. D. (1995). Explanations of individual-group performance differences: What sort of "bonus" can be gained through group interaction? *Small Group Research, 26,* 296–308.

Acknowledgment of Ad Hoc Reviewers: 1997–1998

Group Dynamics: Theory, Research, and Practice extends its thanks to the following individuals who contributed their time and expertise to the journal by serving as reviewers.

Lynley Hicks Anderman
James Balkwell
Sally Barlow
Wendi Barnder
Paul Bliese
Charles F. Bond, Jr.
Nyla R. Branscombe
Jim Bright
Pauline Rose Clance
Russell D. Clark, III
Terry Connolly
Beth Dietz-Uhler
Julie Duck
Mark F. Ettin
Deborah Feltz
Gary Alan Fine
Julie Fitness
Lawrence Frey
Carolyn Green
Bernard Guerin
Alex Halsam
Judy Harackiewicz
A. Paul Hare
David Harrison
Helen C. Harton
Richard Heaps
Terry Hight
Andrea Beth Hollingshead
William T. Hoyt
Susan E. Jackson
Karen Jehn
Norb Kerr
John Kircher
Dennis Kivlighan
Robert Klein
Robin Kowalski
Mike Lambert
Annette Lee-Chai
Frederick Leong
James Liu
David F. Lopez
Robin Martin
Mark McGavin
Lou Medvene
Greg Myers
Steve Nida
Penelope Oakes
Lisa Pelled
Kate Reynolds
Charles E. Rice
Richard J. Riding
Steven G. Rogelberg
Meg Rohand
Gad Saad
Larry Sanna
Janet Schofield
Donald A. Sharpe
William S. Silver
Henry Sims
Kristin Sommer
John Sosik
Russell Spears
Gerry Stone
Jill Sweeney
Shiraz Tata
Patrick Vargas
Bill von Hippel
Janine Wanlass
Jackie Wellen
Ladd Wheeler
Karen Williams
Gwen Wittenbaum
Stephen Worchel
Angelita Yu