



Cornell University
ILR School

Cornell University ILR School
DigitalCommons@ILR

Articles and Chapters

ILR Collection

2004


Methodological Challenges in Union Commitment Studies

Mahmut Bayazit
Cornell University

Tove Hammer
Cornell University, thh2@cornell.edu

David L. Wazeter
Pennsylvania State Education Association

Follow this and additional works at: <https://digitalcommons.ilr.cornell.edu/articles>

 Part of the [Applied Behavior Analysis Commons](#), [Industrial and Organizational Psychology Commons](#), [Organizational Behavior and Theory Commons](#), and the [Unions Commons](#)

Thank you for downloading an article from DigitalCommons@ILR.

Support this valuable resource today!

This Article is brought to you for free and open access by the ILR Collection at DigitalCommons@ILR. It has been accepted for inclusion in Articles and Chapters by an authorized administrator of DigitalCommons@ILR. For more information, please contact catherwood-dig@cornell.edu.

If you have a disability and are having trouble accessing information on this website or need materials in an alternate format, contact web-accessibility@cornell.edu for assistance.

Methodological Challenges in Union Commitment Studies

Abstract

Excerpt] Methodological problems in studies of union commitment were identified and illustrated with data from 4,641 members and 479 stewards in 297 local teachers' unions. Using a 20-item union commitment scale, results confirmed the existence of 3 substantive factors and 1 method factor at the individual level of analysis: loyalty to the union, responsibility to the union, willingness to work for the union, and a factor of negatively worded items. Tests of measurement invariance showed that the scale captured commitment for rank-and-file members but not for union stewards. The authors also found partial measurement invariance between long-time and newer members and full measurement invariance between men and women. Finally, the authors found that violation of the statistical assumption of independence reduced model fit when individual commitment scores were analyzed without attention to the hierarchical nature of the data.

Keywords

unions, membership, teachers, organizing, participation

Disciplines

Applied Behavior Analysis | Industrial and Organizational Psychology | Organizational Behavior and Theory | Unions

Comments

Suggested Citation

Bayazit, M., Hammer, T. H. & Wazeter, D. L. (2004). *Methodological challenges in union commitment studies* [Electronic version]. Retrieved [insert date] from Cornell University, ILR school site: <http://digitalcommons.ilr.cornell.edu/articles/287/>

Required Publisher Statement

Reprinted with permission of [American Psychological Association](#). Final version published as Bayazit, M., Hammer, T. H. & Wazeter, D. L. (2004). Methodological challenges in union commitment studies. *Journal of Applied Psychology*, 89(4), 738-747.

Methodological Challenges in Union Commitment Studies

Mahmut Bayazit, Tove Helland Hammer, David L. Wazeter

Methodological problems in studies of union commitment were identified and illustrated with data from 4,641 members and 479 stewards in 297 local teachers' unions. Using a 20-item union commitment scale, results confirmed the existence of 3 substantive factors and 1 method factor at the individual level of analysis: loyalty to the union, responsibility to the union, willingness to work for the union, and a factor of negatively worded items. Tests of measurement invariance showed that the scale captured commitment for rank-and-file members but not for union stewards. The authors also found partial measurement invariance between long-time and newer members and full measurement invariance between men and women. Finally, the authors found that violation of the statistical assumption of independence reduced model fit when individual commitment scores were analyzed without attention to the hierarchical nature of the data.

Union commitment, as originally defined by Gordon, Philpot, Burt, Thompson, and Spiller (1980), consists of members' loyalty to their union, the responsibility toward and willingness to work for the union, and a belief in unionism. After more than two decades of research, there are still a number of methodological problems related to measurement, samples, and use of analysis techniques that limit researchers' ability to test union commitment models. The purpose of this study was to identify these problems and show how the theoretical models can be tested with appropriate choice of methodology. We identify three areas of concern: the treatment of negatively worded items in union commitment scales, the populations sampled, and the use of individual-level analyses with multilevel data. We describe each problem briefly and demonstrate, with data from a study of members and stewards, in 297 locals belonging to a national union of public schoolteachers, how the fit of a union commitment model improves through a series of methodological steps.

Early research focused on replicating the four orthogonal factors of the Gordon et al. (1980) study, but results differed widely. After reanalyzing Gordon et al.'s data, Friedman and Harvey (1986) suggested a more parsimonious two oblique-factors solution, the first one resembling a combination of the loyalty and belief in unionism factors (i.e., attitudes and opinions) and the second one resembling a combination of responsibility and willingness to work for the union (i.e., behavioral intentions). Tetrick, Thacker, and Fields (1989), and Thacker, Fields, and Tetrick (1989), using confirmatory factor analysis (CFA), identified the original four factors but showed that the factors were highly intercorrelated. Fullagar (1986) showed that the original belief in unionism items (see Item 4 in the Appendix for an example) is not part of the union commitment construct but instead measures commitment to work (see also Klandermans, 1989; Morrow, 1983; and Sverke & Kuruvilla, 1995, for further discussions). In fact, Gordon et al. (1980) had also been concerned about the validity of this factor, which was composed entirely of negatively worded items. After discarding all of the belief in unionism

items, Kelloway, Catano, and Southwell (1992) argued for a three oblique- factors solution. Sverke and Kuruvilla (1995) looked for methodological and contextual correlates to explain the differences in factor solutions but found that there were no obvious effects on the factor structure of the number and type of items used to measure commitment, the research samples used (the type of union), the national context, or the response rates.

Despite disagreements about the number and nature of the dimensions of union commitment, there is no dearth of different models that describe its antecedents and consequences (e.g., Bamberger, Kluger, & Suchard, 1999; Barling, Fullagar, & Kelloway, 1992; Newton & Shore, 1992; Tetrick, 1995). A meta-analysis of model-testing studies showed that prounion attitudes (affect) and beliefs about union instrumentality predict union commitment but that there is ample room for improvement in model specification (Bamberger et al., 1999). Some of that improvement may have to come from conceptual analysis, but we argue that much can be gained from avoiding or solving methodological problems.

Negatively Worded Items as a Method Artifact

To avoid response styles such as acquiescence (e.g., Anastasi & Urbina, 1997), the 30-item scale developed by Gordon et al. (1980) contained 11 negatively worded items. The use of negatively worded items has been questioned because the potential, and sometimes negligible (e.g., Rorer, 1965), advantages they offer seem to be offset by problems with internal consistency (Schriesheim, Eisenbach, & Hill, 1991), factor structure (e.g., Marsh, 1996; Schriesheim & Eisenbach, 1995), and validity (e.g., Holden & Fekken, 1990) in the scales that contain them.¹ These problems may result more from respondents' lack of attention when reading the items, their cognitive capacity, the level of difficulty of the rating task, or the displacement of the neutral points of the items (Holden & Fekken, 1990; Marsh, 1996; Samelson & Yates, 1967) than from a response style.

In union commitment research, Thacker et al. (1989) suggested that the negatively worded scale items contribute to lack of model fit. Kelloway et al. (1992) showed that the negatively worded, reverse-coded items created a method artifact, which led them to recommend eliminating the items altogether. They suggested a shorter, 13-item scale, which has been widely used (e.g., Fullagar, Gallagher, Gordon, & Clark, 1995; Kelloway, Catano, & Carroll, 2000; Tan & Aryee, 2002).

The evidence for a method factor does not preclude the meaningful existence of a scale that contains negatively worded items, however. In fact, the items may have both substantive variance and method variance, and eliminating them may jeopardize content validity of the scales. In the present study, we examined whether negatively worded items create a method artifact and undermine the variance explained by the substantive factors.

¹ Similar concerns have been raised for a number of other psychometric measures (e.g., Cordery & Sevastos, 1993; Harris, 1991; Harvey, Billings, & Nilan, 1985; Idaszak & Drasgow, 1987; Magazine, Williams, & Margaret, 1996; Marsh, 1996; McGee, Ferguson, & Seers, 1989).

Populations Sampled

In some studies, analyses have been based on samples that combine union members and stewards. Others compared the degree of union commitment between these two groups (e.g., Magenau, Martin, & Peterson, 1988), suggesting that stewards' level of commitment should be higher because they have already accepted an active union leadership role. Unfortunately, the basis for drawing such inferences is lacking without the establishment of evidence for metric and structural invariance between these two groups. A central question in measurement invariance is whether scale items measure an identical construct across groups of respondents with different characteristics (Reise, Widaman, & Pugh, 1993). We suspect that there would be variance in measurement parameters because stewards' responses to some of the items in the standard union commitment scales may reflect an assessment of their own performance instead of commitment, as the steward's role involves building union commitment among the members, for example through grievance handling (e.g., Barling et al., 1992).

There may also be variance in metric and structural parameters across groups within a single population (e.g., male vs. female union members or long-time vs. newer members). For example, studies have shown that men have higher mean scores on responsibility toward and willingness to work for the union (e.g., Gordon et al., 1980; Mellor, Mathieu, & Swim, 1994; Thacker, Fields, & Barclay, 1990), whereas women report higher mean scores on loyalty (e.g., Gordon et al., 1980; Iverson & Kuruvilla, 1995). Explanations of these differences have focused on local union structural characteristics such as formalization and centralization (Mellor et al., 1994), union socialization experiences (e.g., Fullagar, McCoy, & Shull, 1992), and differences in nonunion roles (Wetzel, Gallagher, & Soloshy, 1991). However, such differences between groups cannot be unambiguously interpreted without evidence of measurement invariance, which is scant and inconclusive (e.g., Gordon et al., 1980). Newer and older members may also differ in their metric and structural parameters because newer members may not be very knowledgeable about the union environment and their union roles (e.g., Fullagar et al., 1992; Fullagar et al., 1995; Gordon et al., 1980), which can bias their responses toward items that require such knowledge.

Homogeneity of Union Commitment Scores at the Local Union Level

When data come from different local unions within one national or international union (e.g., Gordon et al., 1980; Mellor, 1990) or from a number of unions that differ in type and size (e.g., Kelloway et al., 1992, Study 2; Tan & Aryee, 2002), it cannot be assumed that the commitment scores are independent of each other. Observations from the same organizational unit or cluster tend to be more alike than observations from different clusters, and the nonindependence in the data biases individual-level results if one does not control for contextual variance (Kenny & La Voie, 1985). Unfortunately, none of the previous studies predicting members' commitment to their union from individual- (e.g., Fullagar & Barling, 1989) or union-level variables (e.g., Deery, Iverson, & Erwin, 1994; Fullagar et al., 1995; Mellor et al., 1994) examined or controlled for the possible nonindependence in their data. The analysis of

multilevel data with conventional methods (e.g., ordinary least squares regression, analysis of variance) ignores nonindependence of observations and causes the underestimation of standard errors, increasing the Type I error rates (Zucker, 1990). The extent of this problem increases with the level of the intraclass correlation (ICC) between observations that are nested within clusters (e.g., local unions; Julian, 2001; B. O. Muthén, 1994). However, researchers who ignore the nesting, even with modest ICCs (e.g., ICC = .02), are likely to overestimate the existence of substantive relationships, especially when the number of observations within each group is high and independent variables are constrained to vary only at the group level (Barcikowski, 1981).

The multilevel data analysis takes the nested structure into consideration, improving the estimation of parameters within individual units and allowing the variance and covariance components to be partitioned among levels (e.g., decomposing the correlation matrix among a set of member-level variables into within and between-unions components; B. O. Muthén, 1994). In this study we examine the effect of between-unions variance on the fit of the measurement model using multilevel CFA.

Method

Participants

The data came from 4,641 teachers (union members) and 479 faculty representatives (union stewards) belonging to 297 local unions in a northeastern state of the United States, sampled out of a population of 605 local unions within that state in 1989.² The state association selected members randomly from each of the locals for inclusion in the study, with the number of subjects determined by the size of the local, which means that our data are unbalanced. The surveys were sent to the union members from the state association's headquarters. Code numbers were used to identify respondents by union and to match survey responses to demographic information from the state association's archives. Each survey was accompanied by two letters, one from the president of the state association endorsing the study and one from the researchers describing the purpose of the study and the procedures used to ensure confidentiality of responses and by a postage-paid envelope for returning completed surveys to the researchers at their university. No special incentives were provided to the members to complete the surveys.

Surveys were sent to 14,388 teachers, and 4,652 surveys were returned (32% response rate). The number of members from each local varied between 1 and 52, and the correlation between sample size from each group and actual union local size was .88, suggesting that each local was adequately represented in the final sample. A separate survey was sent to 1,357 stewards, and 492 were returned (36% response rate). No follow-up mailing or contacts were attempted. Eleven surveys from the members' sample and 13 surveys from the stewards' sample were discarded because of missing data.

² The sample included all union locals whose presidents were part of a union leadership study done in conjunction with the present research. Comparisons of school district characteristics (average salary, wealth, tax effort, and enrollment) and members' demographic characteristics (salary, education level, teaching experience, and gender) in the research sample with the other districts and members in the state showed that the research sample differed from the rest of the state only on gender ($M = 1.37$ vs. $M = 1.32$, respectively; $p < .05$).

Measures

Union commitment. Union commitment was measured with a modification of Friedman and Harvey's (1986) 20-item scale, part of the original 30-item scale (Gordon et al., 1980) that previous research had found to be a valid measure of loyalty, responsibility toward, and willingness to work for the union (Kelloway et al., 1992, Study 1). The scale, listed in the Appendix, also contains one of the original belief in unionism items. One of Friedman and Harvey's (1986) items (Item 46) was discarded because it did not apply to this union. It was replaced with another item representing the same dimension (Item 10 in the Appendix). Minor revisions were made on some of the other items, most of which involved replacing the word *union* with *local association*. Each item required a response on a 5-point scale (ranging from 1 = *strongly disagree* to 5 = *strongly agree*). Six of the items were negatively worded (see list in the Appendix).

Demographic variables. The state association provided information on members' union tenure and gender. Sixty-seven percent of the member sample were women ($n = 3,081$), and 33% were men ($n = 1,501$). The mean for union tenure was 15.4 years ($SD = 8.21$).

Data Analysis Procedure

To examine whether (a) negatively worded items contribute meaningful variance to their latent factors, (b) measurement parameters differ across populations, and (c) there is contextual variance in the union commitment scores, we present a number of different sets of models.³

In the first set we established the factor structure of the union commitment measure at the individual level with four models identified as alternative theoretical representations of union commitment (see above for a description of these theoretical models). We compared a model with three oblique factors (Model 1a), with three models hypothesizing (a) three orthogonal factors (Model 1b), (b) two oblique factors (Model 1c), and (c) three oblique factors and an orthogonal method factor (Model 1d).⁴ Finally, we examined the modification indices for possible misspecifications and then present a final model (Model 1e), which we used in the next two sets of models. We also examined the same set of five models for the steward sample.

In the second set of models we examined the measurement and structural invariance of the union commitment scales between union members and union stewards, between long-time members and newer members, and between women and men, using multiple-group CFA. For each group comparison we present (a) a configural invariance (baseline) model in which all factor loadings are freely estimated, (b) a metric invariance model in which all factor loadings

³ For both member and steward samples, Mardia's normalized estimates for multivariate kurtosis (EQS, 5.7b; Bentler, 1995) were high (160 for members and 40 for faculty representatives), indicating multivariate non-normal distributions. We were therefore not able to use maximum likelihood, which assumes multivariate normality, to estimate the parameters in our models. Instead we used Satorra and Bentler's (1994) rescaled test statistics, which correct normal theory maximum-likelihood standard error estimates and test statistics for high multivariate kurtosis and have been shown to work well under a variety of distributions (Curran, West, & Finch, 1996; Hu, Bentler, & Kano, 1992). We report Satorra-Bentler rescaled statistics for all CFA throughout the article unless specified otherwise. We have analyzed our data using the maximum-likelihood mean-adjusted estimation method in Mplus software unless specified otherwise. All chi-square difference tests reported were done using the calculations for the maximum-likelihood mean-adjusted chi-square difference testing described on page 360 in the *Mplus User's Guide* (L. K. Muthén & Muthén 1998).

⁴ The second-order factor model as an alternative model is equivalent to the three oblique-factors model (with the same degrees of freedom). Therefore, these two models are indistinguishable from each other in terms of goodness of fit to data (MacCallum, Wegener, Uchino, & Fabrigar, 1993). A recent meta-analysis on antecedents and consequences of union commitment suggested that multidimensional measures are better representations of union commitment than are unidimensional measures (Bamberger et al., 1999). Hence we report CFA results for only the three oblique-factors model.

are hypothesized to be equal, and if warranted, (c) a partial-metric invariance model (Byrne, Shavelson, & Muthén, 1989). To set the scale for each factor in our configural invariance model, the first factor loading for each factor was constrained to equality across groups, and the variance of the latent variable was fixed at unity in one sample and estimated in the other (Reise et al., 1993). The constraint on the first factor loading was assumed to lead to comparability of factor loadings across groups if the item whose factor loading was constrained to invariance was shown to provide equivalent measurement across groups.⁵

In the third set we present models using the multilevel CFA procedure, recommended by B. O. Muthén (1994), to understand the level of nonindependence at the individual level of analysis and model fit when this nonindependence is controlled. We used three steps in this set. Step 1 was the estimation of the best fitting model at the individual level (Model 1e). We used this model in Step 2 to analyze the sample pooled-within covariance matrix (SPW), which is created by group-mean centering the commitment data of each member. Group-mean centering removes the between groups variation from each respondent's data. Analysis of the SPW matrix is preferred when examining the individual-level factor structure because this matrix is not ill-defined by the between-groups covariation (B. O. Muthén, 1994). Step 3 was the estimation of between-groups variations in the union commitment items. In this step, we first examined the item ICCs, which present the amount of variance in the individual scores that is due to group membership. In addition, we present two independence models. In Independence Model 1, all between-levels variances and covariances were fixed at zero, but means were freely estimated, and all within-level parameters were estimated as in Model 1e. In Independence Model 2, all between-levels variances in addition to means were freely estimated, whereas covariances were fixed at zero, and all within-level parameters were estimated as in Model 1e. Because the two models differ only in the estimation of between-levels variances, comparing them is a more stringent test of whether it is necessary to consider the between-levels variance. If Independence Model 2 fits the data better, we would conclude that between-levels variance is worth further consideration.

Results

Descriptive Statistics

Item descriptive statistics and interitem correlations revealed that two of the items were problematic in both samples. The item "I plan to remain a member. . ." (Item 12 in the Appendix) had the highest skewness and kurtosis indices in both samples. This is not surprising given strong norms encouraging union membership for teachers in the state. Planning to stay a union member most likely does not indicate union loyalty in this context. Therefore, we decided to exclude the item from further analysis of sample data.

The second item, "I talk up the local. . ." (Item 11 in the Appendix) was originally an indicator of union loyalty. However, it can also measure the member's behavior with regard to praising the local to others. In fact, it correlated highly with all of the items from the willingness to work for the union scale ($r = .36-.48$; $p < .01$) almost as highly as with the other items in the

⁵ To test this assumption we ran another baseline model in which the factor variances were set to unity and were hypothesized to be equal across groups to define the scale for each factor. The results from this analysis (not reported here) supported the assumption that the factor loadings for the items that were constrained to be equal in the baseline model were, in fact, invariant.

loyalty scale ($r = .38-.56$; $p < .01$). We therefore discarded this item from further analysis. In previous studies it loaded on both the loyalty and willingness to work dimensions (Gordon et al., 1980; Ladd, Gordon, Beauvais, & Morgan, 1982).

Individual-Level CFA

With the remaining 18 items we ran a single-level CFA, hypothesizing three oblique factors (Model 1a in Table 1). All parameter estimates were significant. Squared multiple correlations (R^2) for the items ranged from .34 to .67 for loyalty to the union, from .30 to .56 for responsibility to the union, and from .44 to .57 for willingness to work for the union factor. The belief in unionism item (Item 4 in the Appendix) had the lowest R^2 and standardized loading on the loyalty factor. The model fit indices showed a lack of fit of the model to the data (Table 1), being lower than their recommended cutoff points (Hu & Bentler, 1999), which are .95 for the Tucker–Lewis Index and comparative fit index, .06 for root-mean-square error of approximation, and .08 for standardized root-mean-square residual.

Compared with Model 1a, the fit of both Model 1b, with three orthogonal factors, $\Delta\chi^2(3, N = 4,641) = 2,685.1, p < .01$, and Model 1c, with two oblique factors, $\Delta\chi^2(2, N = 4,641) = 2,308.6, p < .01$, were significantly worse. Model 1d, in which we hypothesized an orthogonal factor for the method artifact created by the reverse-coded negatively worded items (Items 4, 6, 7, 8, 9, and 14 in the Appendix) in addition to three oblique factors, showed a better fit to the data compared with Model 1a, $\Delta\chi^2(6, N = 4,641) = 582.9, p < .01$. All six negatively worded items, except for the belief in unionism item, had significant loadings on the method factor.

Next we examined the modification indices (modification index estimated using the maximum-likelihood estimation method) for possible misspecifications in our final model (Model 1d) because the model fit was less than optimal. The errors of the two loyalty items (Items 3 and 5 in the Appendix) were highly correlated (modification index = 924; estimate for parameter change = 0.17). Because both items concern an appreciation of the benefits of joining the union, we decided to free this correlation (Model 1e). This modification led to a significant improvement in model fit, $\Delta\chi^2(1, N = 4,641) = 553.1, p < .01$. In the final model, all factors correlated significantly with one another, with loyalty willingness (r_{LW}) = .55; loyalty responsibility (r_{LR}) = .36; and responsibility willingness (r_{RW}) = .53.

We estimated the same set of models using the union steward sample data. Model 1e was also the best fit to these data, as reported in the second portion of Table 1. It is interesting that the belief in unionism item was also the weakest link in the loyalty and method factors. All item R^2 s were lower for the steward sample than they were for the member sample. Factor correlations were as follows: loyalty willingness (r_{LW}) = .44, loyalty responsibility (r_{LR}) = .28, and responsibility willingness (r_{RW}) = .55.

Next we examined the modification indices (modification index estimated using the maximum-likelihood estimation method) for possible misspecifications in our final model (Model 1d) because the model fit was less than optimal. The errors of the two loyalty items (Items 3 and 5 in the Appendix) were highly correlated (modification index = 924; estimate for parameter change = 0.17). Because both items concern an appreciation of the benefits of joining the union, we decided to free this correlation (Model 1e). This modification led to a significant improvement in model fit, $\Delta\chi^2(1, N = 4,641) = 553.1, p < .01$. In the final model, all

Table 1

Goodness-of-Fit Indices for Six Confirmatory Factor Analysis Models Using Union-Member Data

Statistic	Null model (Model 0)	Three oblique factors (Model 1a)	Three orthogonal factors (Model 1b)	Two oblique factors (Model 1c)	Method factor (Model 1d)	Correlated error variance (Model 1e)
Member sample ($n = 4,641$)						
χ^2	3,3079.72	3,486.23	5,223.72	5,024.67	2,826.10	2,272.99
df	153	132	135	134	126	125
SCF	1.25	1.22	1.21	1.22	1.21	1.21
$\Delta\chi^2$			2,685.1 ($df = 3$)**	2,308.6 ($df = 2$)**	582.9 ($df = 6$)**	553.1 ($df = 1$)**
CFI	0.00	0.90	0.85	0.85	0.92	0.94
TLI	0.00	0.88	0.83	0.83	0.90	0.92
RMSEA	0.21	0.07	0.09	0.09	0.07	0.06
SRMR	0.34	0.06	0.18	0.07	0.06	0.06
Faculty representative sample ($n = 479$)						
χ^2	2,269.11	349.41	481.86	490.77	294.45	279.67
df	153	132	135	134	126	125
SCF	1.24	1.21	1.21	1.21	1.19	1.19
$\Delta\chi^2$			164.2 ($df = 3$)**	173.1 ($df = 2$)**	20.9 ($df = 6$)*	14.8 ($df = 1$)**
CFI	0.00	0.90	0.84	0.83	0.92	0.93
TLI	0.00	0.88	0.81	0.81	0.90	0.91
RMSEA	0.17	0.06	0.07	0.08	0.05	0.05
SRMR	0.25	0.06	0.13	0.08	0.06	0.06

Note. Chi-square values are Satorra–Bentler scaled chi-square; SCF = scaling correction factor; CFI = comparative fit index; TLI = Tucker–Lewis fit index; RMSEA = root-mean-square error of approximation; SRMR = standardized root-mean-square residual.

* $p < .01$. ** $p < .001$.

factors correlated significantly with one another, with loyalty willingness (r_{LW}) = .55; loyalty responsibility (r_{LR}) = .36; and responsibility willingness (r_{RW}) = .53.

We estimated the same set of models using the union steward sample data. Model 1e was also the best fit to these data, as reported in the second portion of Table 1. It is interesting that the belief in unionism item was also the weakest link in the loyalty and method factors. All item R^2 s were lower for the steward sample than they were for the member sample. Factor correlations were as follows: loyalty willingness (r_{LW}) = .44, loyalty responsibility (r_{LR}) = .28, and responsibility willingness (r_{RW}) = .55.

Measurement Invariance

We first examined measurement invariance between rank-and-file members and the stewards. As shown in Table 2 (Column 2), constraining the factor loadings to invariance across these two groups led to a statistically significant decrease in model fit, $\Delta\chi^2(20) = 135.24$; $p < .01$. Because the subjective model fit indices are affected by the covariance structure of the larger sample, these indices do not change across models, because the better fitting sample data come from the larger members' sample. There were considerable differences in a number of factor loadings between the two groups (Column 1). Seven items measuring loyalty (Items 4 through 10), one item measuring willingness to work (Item 2), and one item measuring responsibility to the union (Item 3) had smaller factor loadings in the steward sample than in the member sample. In contrast, three of the factor loadings for the method factor (Items 2, 5, and 6; see Table 2) were lower for the member sample, suggesting that the groups also differed in their reactions to the negatively worded items. These results led us to reject the full invariance model. We did not find partial invariance, so there is no need to examine structural invariance (Byrne et al., 1989).

Our second test of measurement invariance was between longtime union members (union tenure greater than 3 years) and newer union members (union tenure less than or equal to 3 years). As shown in Table 2 (Column 4), constraining the factor loadings to invariance across the two groups led to a statistically significant decrease in model fit, $\Delta\chi^2(20) = 70.94$, $p < .01$. Consistent with this finding, some of the practical indices of model fit indicated that the baseline model was a better fit to the data. The two groups had different factor loadings (Column 3 in Table 2) in four items from the loyalty factor (Items 2, 4, 8, and 9), three items from the responsibility factor (Items 2, 3, and 4), and four items from the method factor (Items 2, 3, 4, and 6). The willingness to work factor loadings did not differ between the two groups, indicating metric invariance for that subscale.

Item 2 in the loyalty subscale ("The record of this local is a good example of what dedicated people can get done") may not be a good indicator of loyalty for newer members because it assumes the respondent has knowledge about the history of the local union. Item 4 in the loyalty subscale ("As long as I am doing the kind of work I enjoy, it does not matter if I belong to a teacher's association"), originally an indicator of the belief in unionism factor, is also problematic. This finding is consistent with Gordon et al.'s (1980), which suggested that the belief in unionism factor might only exist for long-time members and not for newer members.

Items loading on the method factor also behaved differently across older and newer members. Item 2 in the method factor (which is Item 4 in the loyalty subscale) had a significant negative loading for the newer members' sample, indicating that newer members may consider this a positive item rather than a negative one. To explore partial invariance, we tested a third

model in which we freely estimated the factor loadings of Items 2 and 4 in the loyalty subscale and all method factor loadings for both groups and kept all of the other items invariant across groups. This partial invariance model fit just as well as did our baseline model, $\Delta\chi^2(13) = 5.7$, $p > .05$, suggesting that all three subscales (minus two items in the loyalty subscale) are valid measures for both groups.

Structural invariance between long-time and newer member subsamples on the four latent variables (see Model 1e in Table 1) can be evaluated by contrasting means and variance estimates across groups relative to their standard errors (Reise et al., 1993). In both the partial and the full invariance models, the factor means for responsibility and willingness factors were significantly lower for newer members. The variances of the loyalty and willingness factors were also smaller for newer members, indicating possible range restriction.

Our third analysis examined measurement invariance between male and female member subsamples. As shown in Table 2 (Column 6), constraining the factor loadings to invariance (forcing them to be equal) across groups led to a statistically nonsignificant decrease in model fit, $\Delta\chi^2(20) = 4.13$, $p > .05$. In addition, factor loadings (Column 5 in Table 2) were quite similar between men and women, and some fit indices indicated that the full invariance model was a better fit. Given these results, we fail to reject the full invariance model.

Results for structural invariance showed differences in factor means and variances across the two groups. The means on the loyalty factor were significantly lower for men than for women. In contrast, means on willingness to work for and responsibility to the union factors were significantly higher for men than for women. Factor variances for the loyalty and willingness factors were higher for men than for women.

Multilevel CFA

We analyzed the factor structure at the individual level, taking into account the hierarchical nature of the data in our CFA. First, we repeated our analysis of an individual-level CFA (Model 1e) using the (S_{PW}) matrix (see Column 1 in Table 3), which was created by group-mean centering the commitment data to remove the between-group variation from each individual's scores. Second, we estimated Model 1e using the S_T matrix (raw total covariance) with the maximum-likelihood estimation method, $\chi^2(125, N=4,345) = 2,750.1$, $p < .01$; comparative-fit index = .94; Tucker–Lewis Index = .92; root-mean-square error of approximation = 0.07; and standardized root-mean-square residual = 0.06. To examine whether the between-levels covariation distorted the data, we compared the model fit using S_{PW} with the model fit using the S_T matrix. Comparing these two models, we can state that analysis with S_{PW} , which removes the between-levels variance, provides a better fit, $\Delta\chi^2(0, N = 4,345) = 530.1$, $p < .01$.

Third, we estimated the level of between-groups homogeneity using the ICCs, or the proportion of variance explained by group membership. ICCs for the 18 items ranged from .01 to .15. The ICCs for all responsibility to the union and willingness to work items were lower than .05. Finally, to determine whether the between-levels variance biased the interpretation of the individual level results, we compared two models, hypothesizing independence in the between level. Independence Model 2, in which the between-levels variances were freely estimated, provided a better fit to the data, $\Delta\chi^2(18) = 364.1$, $p < .01$, indicating that there was a significant amount of variance in the between levels that reduced the fit of the measurement model at the individual level.

Table 2

Tests of Measurement Invariance Comparing Faculty Representatives and Members, Long-Time Members and Newer Members, and Men and Women

Model Group	Union status				Union tenure				Gender			
	Baseline model		Full invariance		Baseline model		Full invariance		Baseline model		Full invariance	
	Faculty reps. (n = 479)	Members (n = 4,536)	Faculty reps. (n = 479)	Members (n = 4,536)	Long-time members (n = 4,047)	Newer members (n = 233)	Long-time members (n = 4,047)	Newer members (n = 233)	Women (n = 3,081)	Men (n = 1,501)	Women (n = 3,081)	Men (n = 1,501)
χ^2	2,634.69		2,755.44		2,310.87		2,381.81		2,481.32		2,485.43	
df	264		284		264		284		264		284	
SCF	1.19		1.21		1.15		1.15		1.20		1.20	
$\Delta\chi^2(df = 20)$			135.24**				70.94**		4.13			
CFI	0.93		0.93		0.94		0.93		0.93		0.93	
TLI	0.92		0.92		0.93		0.93		0.92		0.93	
RMSEA	0.06		0.06		0.06		0.06		0.06		0.06	
SRMR	0.06		0.06		0.06		0.06		0.06		0.06	
$\lambda_{1,1}$	0.76		0.59		0.73		0.73		0.70		0.71	
$\lambda_{2,1}$	0.77	0.69	0.54		0.68	0.44	0.67		0.64	0.65	0.64	
$\lambda_{3,1}$	0.74	0.80	0.62		0.77	0.70	0.77		0.75	0.72	0.74	
$\lambda_{4,1}$	0.26	0.57	0.43		0.53	0.75	0.53		0.53	0.49	0.52	
$\lambda_{5,1}$	0.34	0.73	0.55		0.70	0.68	0.70		0.68	0.67	0.68	
$\lambda_{6,1}$	0.48	0.61	0.47		0.60	0.56	0.60		0.58	0.55	0.57	
$\lambda_{7,1}$	0.32	0.59	0.44		0.57	0.56	0.57		0.55	0.52	0.54	
$\lambda_{8,1}$	0.50	0.73	0.55		0.69	0.78	0.69		0.66	0.67	0.66	
$\lambda_{9,1}$	0.43	0.74	0.56		0.69	0.82	0.70		0.67	0.67	0.68	
$\lambda_{10,1}$	0.62	0.82	0.63		0.80	0.76	0.80		0.76	0.75	0.76	
$\phi_{1,1}$	1.00 ^a	0.91	1.00 ^a	1.55	1.00 ^a	0.61	1.00 ^a	0.58	1.00 ^a	1.20	1.00 ^a	1.17
$\kappa_{1,1}$	0.00	-0.53	0.00	-0.45	0.00	-0.08	0.00	-0.08	0.00	-0.15	0.00	-0.14
$\lambda_{1,2}$	0.59		0.55		0.70		0.70		0.67		0.67	
$\lambda_{2,2}$	0.39	0.63	0.57		0.74	0.73	0.74		0.71	0.73	0.72	
$\lambda_{3,2}$	0.63	0.60	0.56		0.72	0.67	0.72		0.71	0.67	0.70	
$\lambda_{4,2}$	0.72	0.61	0.57		0.74	0.73	0.74		0.72	0.71	0.72	
$\phi_{2,2}$	1.00 ^a	1.41	1.00 ^a	1.65	1.00 ^a	0.91	1.00 ^a	0.87	1.00 ^a	1.16	1.00 ^a	1.14
$\kappa_{2,2}$	0.00	-1.20	0.00	-1.16	0.00	-0.15	0.00	-0.17	0.00	0.16	0.00	0.16
$\lambda_{1,3}$	0.58		0.53		0.55		0.55		0.54		0.54	
$\lambda_{2,3}$	0.47	0.45	0.42		0.43	0.36	0.43		0.42	0.40	0.42	
$\lambda_{3,3}$	0.55	0.66	0.60		0.62	0.53	0.62		0.61	0.58	0.61	
$\lambda_{4,3}$	0.43	0.47	0.43		0.45	0.35	0.45		0.45	0.41	0.45	
$\phi_{3,3}$	1.00 ^a	0.90	1.00 ^a	1.06	1.00 ^a	1.24	1.00 ^a	0.95	1.00 ^a	1.10	1.00 ^a	1.01
$\kappa_{3,3}$	0.00	-0.39	0.00	-0.35	0.00	-0.59	0.00	-0.50	0.00	0.30	0.00	0.29
$\lambda_{1,4}$	0.29		0.33		0.35		0.36		0.34		0.35	
$\lambda_{2,4}$	0.13	-0.01	0.01		0.02	-0.53	0.01		0.01	0.00	0.01	
$\lambda_{3,4}$	0.36	0.33	0.39		0.42	0.12	0.42		0.38	0.37	0.38	
$\lambda_{4,4}$	0.18	0.18	0.21		0.24	-0.08	0.23		0.24	0.21	0.24	
$\lambda_{5,4}$	0.31	0.14	0.18		0.20	0.25	0.20		0.21	0.15	0.19	
$\lambda_{6,4}$	0.28	0.04	0.07		0.08	-0.16	0.07		0.06	0.08	0.07	
$\phi_{4,4}$	1.00 ^a	1.52	1.00 ^a	1.16	1.00 ^a	0.43	1.00 ^a	0.51	1.00 ^a	1.28	1.00 ^a	1.09
$\kappa_{4,4}$	0.00	0.47	0.00	0.21	0.00	0.40	0.00	0.16	0.00	-0.31	0.00	-0.30

Note. The first subscript for factor loadings (λ s) denotes the item numbers (see the Appendix for the order), and the second subscript denotes factors (1 = loyalty to the union, 2 = willingness to the union, 3 = responsibility to the union, and 4 = method factor). The subscripts for factor means (κ) and variances (ϕ) identify factors. Chi-square values are Satorra-Bentler scaled chi-square; SCF = scaling correction factor; CFI = comparative fit index; TLI = Tucker-Lewis fit index; RMSEA = root-mean-square error of approximation; SRMR = standardized root-mean-square residual. Reps. = representatives.

^a Parameters fixed at unity to identify each model.

** $p < .001$.

Table 3

Results of Two-Level Confirmatory Factor Analysis (CFA) on Union Commitment Using Members' Data

Model	Individual-level CFA	Independence Model 1	Independence Model 2
χ^2	2,220.91 ^a	3,428.20	3,074.10
<i>df</i>	125	296	278
SCF	—	1.03	1.04
<i>N</i> (within)	4,345	4,345	4,345
<i>N</i> (between)		296	296
$\Delta\chi^2$			364.1 (<i>df</i> = 18)**
CFI	0.94	0.92	0.93
TLI	0.93	0.92	0.92
RMSEA	0.06	0.05	0.05
SRMR _w	0.06	0.06	0.06
SRMR _b	—	0.19	0.20

Note. Chi-square values are Satorra–Bentler scaled chi-square; SCF = scaling correction factor; CFI = Comparative fit index; TLI = Tucker–Lewis fit index; RMSEA = root-mean-square error of approximation; SRMR_w = standardized root-mean-square residual within value; SRMR_b = standardized root-mean-square residual between value. Individual level CFA = analysis using pooled-within matrix in which all individual-level data are group-mean centered; there is no between-level in the model, and the model is same as Model 1e in Table 1. Independence Model 1 = All variances and covariances in between level are fixed at zero, and means are free. Independence Model 2 = All covariances are fixed at zero and all variances and means are free in between level. For all two-level models, within-level portion of the model is the same as with Model 1e in Table 1.

^a A pooled-within covariance matrix is used in this analysis; therefore, the estimation method available for such summary data is normal theory maximum likelihood.

** $p < .001$.

Teachers may differ from blue-collar union members in that their commitment could be influenced by professional as well as bread-and-butter issues. To test the generalizability of some of our findings we analyzed union commitment data (on 14 of the items representing all three factors) from 1,055 members of the National Association of Letter Carriers collected by Kuruvilla and Sverke (1993).⁶ Because of data limitations we were able to cross-validate our findings only with regard to factor structure and measurement invariance between men and women and new and long-time members. Initially we established that the three oblique-factors model was also the best fitting model for the mail carriers sample. Male and female samples once again did not differ in their measurement parameters. Two items that required knowledge of the union's history (Item 2 in the Appendix) and knowledge of union members' responsibilities to other members (Item 19 in Appendix) created a bias in the measurement of union commitment for new members. These results increase our confidence in the generalizability of some of our findings to members of other unions.

Discussion

In this study we examined negatively worded items, measurement invariance, and homogeneity of variance at the union level as methodological challenges in union commitment studies. Our first set of results show that three oblique factors and an orthogonal method factor for the reverse-coded negatively worded items were the best fit to the data in both the union member and union steward samples. The existence of a method factor indicates nonrandom measurement error shared by the negatively worded items. Controlling for such shared variance in the items did not decrease their loadings on the three substantive factors, suggesting that they are an important part of the loyalty subscale.

The original belief in unionism item loaded on the loyalty factor, but the factor explained very little variance for this item. Also, the item did not load significantly on the method factor, suggesting that it was not perceived as negatively worded. It also loaded highly on the loyalty factor for newer members but not for longtime members, which Gordon et al. (1980) suggested would happen.

Our second set of results shows that we cannot assume measurement invariance between rank-and-file union members and union stewards for the commitment scales in general and the loyalty subscale specifically. If these scales are used to measure the commitment of union stewards, the estimates would likely be biased and potentially misleading indicators of their union commitment. In fact, we suggest that stewards' responses to these items may be assessments of their own job performance more than of their commitment. Our results apply to those who were stewards at the time of the data collection. It is not clear how much members who previously occupied that position or other important positions differ from other members.

⁶ The two samples were different in union tenure ($M = 12.2$ years vs. $M = 15.4$ years for letter carriers and teachers, respectively; $p < .01$) and gender distribution (there was a majority of men in letter carriers sample). Furthermore, the sample of letter carriers came from a number of different states, and respondents were asked about their commitment to the national union. We also examined measurement invariance between teachers and mail carriers; however, because of the sampling differences mentioned above, the results of this comparison are not conclusive. Results of the analyses on the letter carrier sample are available from Mahmut Bayazit.

Furthermore, tests of measurement invariance suggest that one of the items on the loyalty subscale (“The record of the local is a good example of what dedicated people can get done”) is not suitable for newer members. The use of this item or similar items that require knowledge of union history or practices to measure the commitment of new members, or perhaps of all members in newly formed unions, is problematic.

Our results also suggest that it is safe to assume measurement invariance between male and female members of the union, but nonstructural invariance. This is consistent with previous results showing similar factor solutions but different mean scores between these groups (Gordon et al., 1980). We also found that women had smaller factor variances than men on two of the factors, which suggests that there are variables constraining women’s commitment.

Finally, the between-levels variance in union commitment scores was low, but it still distorted the fit of the individual-level commitment model in this study. The fit of the measurement model improved significantly when we controlled for the nonindependence at the individual level of analysis, indicating the presence of contextual effects on union commitment.

The primary sample used in this study consisted of teachers from 297 local unions in one state, which may have constrained the between-groups variance in our multilevel CFA analysis. It is possible that there is important between-groups variance at the state level as opposed to at the local union level. If that is the case, our results are conservative and should be considered a lower limit for union level effects.

In light of the findings presented above, we make the following recommendations for future researchers using the 20-item union commitment scale or its variants:

1. Do not include items from the original belief in unionism scale in union commitment measures.
2. Replace negatively worded items with positively worded items or control for the method variance. Control of response styles may be better achieved through other means (see, e.g., Barnette, 2000; Marsh, 1986).
3. Researchers should collect data on union status and on level of involvement within the union. A different set of items generated using a different conceptualization of union commitment may be necessary to measure stewards’ to their union or their commitment to their union role.
4. Researchers should collect data on gender and tenure of union members to control for their effects in analysis of substantive relationships.
5. Researchers should be aware of naturally occurring hierarchies in the data when testing hypotheses at the individual level.
6. Before testing individual-level or cross-hypotheses, researchers should examine the level of nonindependence by calculating the variance in union commitment scores attributed to union membership.
7. If the between-unit variance is significant, even if it is small ($ICC < .05$), researchers should account for it in the measurement model (e.g., multilevel CFA) as well as in the estimation of theoretical models (e.g., random coefficient modeling).

References:

- Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice Hall.
- Bamberger, P. A., Kluger, A. N., & Suchard, R. (1999). The antecedents and consequences of union commitment: A meta-analysis. *Academy of Management Journal*, *42*, 304–318.
- Barcikowski, R. S. (1981). Statistical power with group mean as the unit of analysis. *Journal of Educational Statistics*, *6*, 267–285.
- Barling, J., Fullagar, C., & Kelloway, K. (1992). *The union and its members: A psychological approach*. New York: Oxford University Press.
- Barnette, J. J. (2000). Effects of stem and Likert response option reversals on survey internal consistency: If you feel the need, there is a better alternative to using those negatively worded stems. *Educational and Psychological Measurement*, *60*, 361–370.
- Bentler, P. M. (1995). *EQS structural equations program manual*. Encino, CA: Multivariate Software.
- Byrne, M. B., Shavelson, R. J., & Muthén, B. O. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin*, *105*, 456–466.
- Cordery, J. L., & Sevastos, P. P. (1993). Responses to the original and revised Job Diagnostic Survey: Is education a factor in responses to negatively worded items? *Journal of Applied Psychology*, *78*, 141–143.
- Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, *1*, 16–29.
- Deery, S. J., Iverson, R. D., & Erwin, P. J. (1994). Predicting organizational and union commitment: The effect of industrial relations climate. *British Journal of Industrial Relations*, *32*, 581–597.
- Friedman, L., & Harvey, R. J. (1986). Factors of union commitment: The case for a lower dimensionality. *Journal of Applied Psychology*, *71*, 371–376.
- Fullagar, C. J. A. (1986). A factor analytic study on the validity of a union commitment scale. *Journal of Applied Psychology*, *71*, 129–136.
- Fullagar, C. J. A., & Barling, J. (1989). A longitudinal test of a model of the antecedents and consequences of union loyalty. *Journal of Applied Psychology*, *74*, 213–227.

- Fullagar, C. J. A., Gallagher, D. G., Gordon, M. E., & Clark, P. F. (1995). Impact of early socialization on union commitment and participation: A longitudinal study. *Journal of Applied Psychology, 80*, 147–157.
- Fullagar, C. J. A., McCoy, D., & Shull, C. (1992). The socialization of union loyalty. *Journal of Organizational Behavior, 13*, 13–26.
- Gordon, M. E., Philpot, J. W., Burt, R. E., Thompson, C. A., & Spiller, W. E. (1980). Commitment to the union: Development of a measure and an examination of its correlates [Monograph]. *Journal of Applied Psychology, 65*, 479–499.
- Harris, M. (1991). Role conflict and role ambiguity as substance versus artifact: A confirmatory factor analysis of House, Schuler, and Levanoni's (1983) scales. *Journal of Applied Psychology, 76*, 122–126.
- Harvey, R., Billings, R., & Nilan, K. (1985). Confirmatory factor analysis of the Job Diagnostic Survey: Good news and bad news. *Journal of Applied Psychology, 70*, 461–468.
- Holden, R. R., & Fekken, G. C. (1990). Structured psychopathological test item characteristics and validity. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 2*, 35–40.
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55.
- Hu, L. T., Bentler, P. M., & Kano, Y. (1992). Can test statistics in covariance structure analysis be trusted? *Psychological Bulletin, 112*, 351–362.
- Idaszak, J. R., & Drasgow, F. (1987). A revision of the Job Diagnostic Survey: Elimination of measurement artifact. *Journal of Applied Psychology, 72*, 69–74.
- Iverson, R., & Kuruvilla, S. (1995). Antecedents of union loyalty: The influence of individual dispositions and organizational context. *Journal of Organizational Behavior, 16*, 557–582.
- Julian, M. W. (2001). The consequences of ignoring multilevel data structures in nonhierarchical covariance modeling. *Structural Equation Modeling, 8*, 325–352.
- Kelloway, E. K., Catano, V. M., & Carroll, A. E. (2000). Psychological involvement in the union. *Canadian Journal of Behavioural Science, 32*, 163–167.
- Kelloway, E. K., Catano, V. M., & Southwell, R. R. (1992). The construct validity of union commitment: Development and dimensionality of a shorter scale. *Journal of Occupational and Organizational Psychology, 65*, 197–211.

- Kenny, D. A., & La Voie, L. (1985). Separating individual and group effects. *Journal of Personality and Social Psychology*, *48*, 339–348.
- Klandermans, B. (1989). Union commitment: Replications and tests in the Dutch context. *Journal of Applied Psychology*, *74*, 869–875.
- Kuruvilla, S., & Sverke, M. (1993). Two dimensions of union commitment based on the theory of reasoned action: Cross cultural comparisons. *Research & Practice in Human Resource Management*, *1*, 1–16.
- Ladd, R. T., Gordon, M. E., Beauvais, L. L., & Morgan, R. L. (1982). Union commitment: Replication and extension. *Journal of Applied Psychology*, *67*, 640–644.
- MacCallum, R. C., Wegener, D. T., Uchino, B. N., & Fabrigar, L. R. (1993). The problem of equivalent models in applications of covariance structure analysis. *Psychological Bulletin*, *114*, 185–199.
- Magazine, S. L., Williams, L. J., & Margaret, L. (1996). A confirmatory factor analysis examination of reverse-coding effects in Meyer and Allen's affective and continuance commitment scales. *Educational and Psychological Measurement*, *56*, 241–245.
- Magenau, J. M., Martin, J. E., & Peterson, M. M. (1988). Dual and unilateral commitment among stewards and rank and file union members. *Academy of Management Journal*, *31*, 359–376.
- Marsh, H. W. (1986). The bias of negatively worded items in rating scales of young children: A cognitive developmental phenomenon. *Developmental Psychology*, *22*, 37–49.
- Marsh, H. W. (1996). Positive and negative global self-esteem: A substantively meaningful distinction or artifacts? *Journal of Personality and Social Psychology*, *70*, 810–819.
- McGee, G. W., Ferguson, C., & Seers, A. (1989). Role conflict and role ambiguity: Do the scales measure these two constructs? *Journal of Applied Psychology*, *74*, 815–818.
- Mellor, S. (1990). The relationship between membership decline and union commitment: A field study of local unions in crisis. *Journal of Applied Psychology*, *75*, 258–267.
- Mellor, S., Mathieu, J. E., & Swim, J. K. (1994). Cross-level analysis of the influence of local union structure on women's and men's union commitment. *Journal of Applied Psychology*, *79*, 203–210.
- Morrow, P. C. (1983). Concept redundancy in organizational research: The case of work commitment. *Academy of Management Review*, *8*, 486–500.

- Muthén, B. O. (1994). Multilevel covariance structure analysis. *Sociological Methods & Research*, 22, 376–398.
- Muthén, L. K., & Muthén, B. O. (1998). *Mplus user's guide*. Los Angeles, CA: Muthén & Muthén.
- Newton, L., & Shore, M. L. (1992). A model of union membership: Instrumentality, commitment, and opposition. *Academy of Management Review*, 17, 157–164.
- Reise, S. P., Widaman, K. F., & Pugh, R. H. (1993). Confirmatory factor analysis and item response theory: Two approaches for exploring measurement invariance. *Psychological Bulletin*, 114, 552–566.
- Rorer, L. G. (1965). The great response style myth. *Psychological Bulletin*, 63, 129–156.
- Samelson, F., & Yates, J. F. (1967). Acquiescence and the F scale. *Psychological Bulletin*, 68, 91–103.
- Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structural analysis. In A. von Eye & C. C. Clogg (Eds.), *Latent variable analysis: Applications to development research* (pp. 399–419). Newbury Park, CA: Sage.
- Schriesheim, C. A., & Eisenbach, R. J. (1995). An exploratory and confirmatory factor-analytic investigation of item wording effects on the obtained factor structures of survey questionnaire measures. *Journal of Management*, 21, 1177–1193.
- Schriesheim, C. A., Eisenbach, R. J., & Hill, K. D. (1991). The effect of negation and polar opposite item reversals on questionnaire reliability and validity: An experimental investigation. *Educational and Psychological Measurement*, 51, 67–78.
- Sverke, M., & Kuruvilla, S. (1995). A new conceptualization of union commitment: Development and test of an integrated theory. *Journal of Organizational Behavior*, 16, 505–532.
- Tan, H. H., & Aryee, S. (2002). Antecedents and outcomes of union loyalty: A constructive replication and an extension. *Journal of Applied Psychology*, 87, 715–722.
- Tetrick, L. E. (1995). Developing and maintaining union commitment: A theoretical framework. *Journal of Organizational Behavior*, 16, 583–595.
- Tetrick, L. E., Thacker, J. W., & Fields, M. W. (1989). Evidence for the stability of the four dimensions of the commitment to the union scale. *Journal of Applied Psychology*, 74, 819–823.
- Thacker, J. W., Fields, M. W., & Barclay, L. A. (1990). Union commitment: An examination of

antecedents and outcome factors. *Journal of Occupational Psychology*, 63, 33–48.

Thacker, J. W., Fields, M. W., & Tetrick, L. E. (1989). The factor structure of union commitment: An application of confirmatory factor analysis. *Journal of Applied Psychology*, 74, 228–232.

Wetzel, K., Gallagher, D. G., & Soloshy, D. E. (1991). Union commitment: Is there a gender gap? *Relations Industrielles*, 46, 564–583.

Zucker, D. M. (1990). An analysis of variance pitfall: The fixed effects analysis in a nested design. *Educational and Psychological Measurement*, 50, 731–768.

Research Reports

Appendix

Union Commitment Items Used in the Study

Loyalty

1. I feel a sense of pride being part of the local association.
2. The record of this local is a good example of what dedicated people can get done.
3. There is a lot to be gained by joining the local association.
4. As long as I am doing the kind of work I enjoy, it does not matter if I belong to a teacher's association. (R)A1
5. Deciding to join the local association was a smart move on my part.
6. I have little confidence and trust in most members of my local association. (R)
7. Very little that the membership wants has any real importance to the local association. (R)
8. My values and local associations' are not very similar. (R)
9. I feel little loyalty to the local association. (R)
10. The local association adequately represents the interests of all members.
11. I talk up the local association to my friends as a great organization to be a member of.
12. Based on what I know now and what I believe I can expect in the future, I plan to be a member of the association for the rest of the time I work in the district.

Willingness to Work for the Union

13. If asked I would serve on a committee for the local association.
14. I doubt that I would do any special work to help the local association. (R)
15. I am willing to put in a great deal of effort beyond that normally expected of a member to make the local association successful.

16. If asked I would run for an elected office in the local association.

Responsibility to the Union

17. Every association member must be prepared to take the risk of filing a grievance.

18. It is the duty of every member "to keep his/her ears open" for information that might be helpful to the local association.

19. It is every member's duty to support or help another member use the grievance procedure.

20. It is every member's responsibility to see to it that the school administration "lives up to" all the terms of the contract.

Note. (R) denotes reverse-coded negatively worded items. A1Originally a Belief in Unionism scale item (Gordon et al., 1980).