Aggregating Individual-Level Evaluations of the Organizational Social Climate: A Multilevel Investigation of the Work Environment at the Federal Bureau of Prisons

The opinions expressed in this paper represent the views of the authors and do not necessarily reflect the position of either the U.S. Department of Justice or the Federal Bureau of Prisons.

Appears in *Justice Quarterly* 14: 4: 739-762. Copyright owned by the Academy of Criminal Justice Sciences.

Prepared by
Scott D. Camp, Ph.D.
William G. Saylor, M.A.
and Miles D. Harer, Ph.D.
Federal Bureau of Prisons
320 First Street NW
Washington, D.C. 20534

Aggregating Individual-Level Evaluations of the Organizational Social Climate: A Multilevel Investigation of the Work Environment at the Federal Bureau of Prisons

Abstract

The social climate of the work site is typically regarded as highly influential for individual and organizational productivity. This research examines intrafirm variability in the social climate by studying how an organization, the Federal Bureau of Prisons, uses information it gathers about the social climate. Intrafirm variability in social climate is not usually examined, although it has clear implications for theory, methods, and policy. Multilevel models are used to investigate the empirical properties of three scales of the social climate: institutional operations, commitment to the organization, and satisfaction with supervision. The analysis points out the necessity of moving beyond traditional concerns about scale properties at the individual level when aggregating individual-level measures into organizational measures, especially when the theoretical and practical concerns pertain to intrafirm variability in the social climate.

Aggregating Individual-Level Evaluations of the Social Climate

Introduction

Organizational analysts are following the lead provided initially by human relations theorists more than 50 years ago and extended by theorists who stress currently relevant organizational theories such as total quality management (TQM). These analysts emphasize that social working conditions affect organizationally relevant individual outcomes such as absenteeism, turnover, work attitudes, and conflict, which ultimately affect individual and organizational productivity (Guillèn 1994). A typical observation might be that social climate affects productivity by influencing three types of behaviors: attachment, performance, and citizenship (Kopelman, Brief, and Guzzo 1990).

Many unresolved methodological and theoretical issues are inherent in evaluating whether the social climate has the hypothesized impact on individual and/or organizational productivity. The issues center largely around developing appropriate measures of outcomes and/or the social climate. Conceptualizing and measuring outcomes in a public-sector organization is especially problematic; it is also difficult to measure outcomes in private, for-profit organizations where profits are the ultimate measuring stick. We are not concerned with outcomes in this analysis, however. Instead we focus on measures of the social climate, itself a problematic area both theoretically and methodologically.

Theoretically, analysts typically point to international or (at the very least) interfirm differences as crucial for explaining differences in nations' or firms' productivity and survival (Guillèn 1994; Lincoln and Zeitz 1980; Raudenbush and Willrns 1991). Less often examined, though no less interesting, is intrafirm variability in creating and maintaining appropriate social

climates to support the attainment of organizational goals. An interesting question, for example, is whether the social climate observed at different sites of a single organization are predicated mainly on overall organizational practices and policies or on local implementations of those practices and policies. In other words, do local managers make a difference?

At a commonsense level, we usually think that local managers make a difference. Still, it is prudent to examine whether in fact we can measure the extent of local managers' control over the social climate at the locations for which they are responsible. This point is especially relevant in the public sector, given initiatives such as the National Performance Review (1993). Federal government agencies are being asked to measure results, including the effectiveness of staff and managers. Yet before constructing instruments measuring the social climate that can be used to compare managers, we must examine the theoretical and methodological issues inherent in the instruments. In this research we address these types of questions. Unless such research is conducted, executives and policy makers ultimately may hold local management responsible for processes over which they have little or no control.

Methodologically the most general issue raised in measuring the social climate revolves around potential problems with aggregate measures. Measures of the social climate are almost invariably aggregations of individual evaluations of that climate, or what Lazarsfeld and Menzel (1961) call "analytical" measures. Aggregate measures of organizational properties stand in contrast to "global" measures, such as the technology used by an organization, which cannot be disaggregated into individual evaluations or traits. The use of aggregate measures of organizational properties, such as social climate, has generated considerable methodological controversy. Some researchers claim that the individual is the only proper unit of analysis with

such measures; others claim that the organization is the only appropriate unit. The controversy continues even though Lincoln and Zeitz (1980) outlined rudimentary methods, 17 years ago, for investigating whether statistical relationships between variables arise from organization-level or individua-level processes.

Lincoln and Zeitz (1980) claim that not all aggregations of individual-level evaluations are theoretically valid measures of organizational properties. For example, they claim that job satisfaction is an individual "sentiment" for which there is no corresponding organizational trait (p.392). Whether or not Lincoln and Zeitz are correct about the proper use of job satisfaction measures, they rightly point out that we must give careful attention to the conceptual and methodological issues involved in constructing aggregate measures. Unfortunately their advice has often been ignored.

In this paper we examine the important theoretical issues surrounding intrafirm variation in social climate. Simultaneously we address the methodological issues of creating aggregate measures of the social climate by examining how one public-sector organization, the Federal Bureau of Prisons (BOP), tracks the climate at its different prisons. In addition to exploring how organizational analysts consider and measure the social climate, this paper has policy implications for the BOP and other agencies that have adopted similar tactics for evaluating this climate. In public-sector law enforcement agencies alone, several other state corrections departments, as well as the Correctional Services of Canada, have adopted in whole or in part the survey instrument that the BOP uses for this purpose. In addition, Immigration and Naturalization Services is preparing to deal with management issues created by the agency’s current and projected rapid growth.

Issues

Behavior occurs within, and is shaped by, social context (Hauser 1970; Lazarsfeld and Menzel 1961; Lincoln and Zeitz 1980). Expressed in these terms, the statement is practically a truism for social scientists. Even so, most empirical research conducted by social scientists focused until recently on behavior at either the individual or the organizational level, but rarely encompassed both levels at once. This was due largely to inadequate statistical methods for analyzing individual and organizational levels simultaneously. Recently, however, methods have been developed for incorporating both levels of analysis into one study. Much of the work in developing these new methods has been pioneered by education researchers, who are interested primarily in the effects of school differences—including the climates created by different mixes of teachers, students, and educational practices—on individ*ual* students' educational achievement (Bryk and Raudenbush 1992).

Managers at the BOP are interested in tracking the social climate of the working environment at the BOP. Each year since 1988, the Office of Research and Evaluation at the BOP has administered the Prison Social Climate Survey (PSCS) to a stratified, proportional probability sample of employees. In 1994, the survey year analyzed here, the survey was administered to 9,228 staff members working in 74 different federal prisons. Usable surveys were collected from 8,115 respondents for a response rate of 87.9 percent. Results from the PSCS, along with a broad array of other operational data collected by the BOP, are disseminated to BOP managers via an executive information system known as KI/SSS, the Key Indicators/Strategic Support System (Gilman 1991; Saylor 1988). One component of KI/SSS is ESMI, the Executive Staff Management Indicators module (Muth 1995); this is a collection of measures regarded by

executive staff members at the BOP as the most important for tracking the performance of BOP institutions.

The organizational elements tracked by ESMI include three measures of the organizational social climate. These institutional measures are aggregated from individual-level scales that measure evaluations of institutional operations, commitment to the BOP, and satisfaction with supervision. Research by Office of Research and Evaluation staff has demonstrated the validity of these scales as individual-level measures (Saylor, Gilman, and Camp 1996) and their usefulness in predicting objective outcomes at the individual level (Camp 1994). The individual-scale scores are aggregated in ESMI by reporting the percentage of staff members at each institution who report on average that they agree with the items used to construct the scales. Agreement with the items signifies high levels of the respective latent properties.

Developers of the executive information system warn BOP managers that the institution scores for the three social climate scales are summary scores of individual responses. Despite this warning, however, the information system has features that strongly encourage viewing the institutional results as if they represent organizational properties. First, summarization itself shifts the unit of analysis to the organization. Second and more important, the institution scores are used in the information system to create low-to-high rankings of the prisons. This almost certainly creates an impetus to view the institution scores as representing organizational properties rather than simply summaries of individual scores.

In light of the preceding discussion, it is essential to investigate the organizational properties of the summary measures of the social climate used by the BOP because these measures are used in that manner in any case. This examination presents the opportunity to

theoretically and methodologically examine the organizational properties of aggregate scales and their ability to reveal the extent of local managers' control over local social climate. The intrafirm aspect of the social climate typically is neglected in organizational studies. Methodologically this research allows us to examine whether accurate organizational measures can be derived simply by aggregating accurate individual-level measures. As stated above, there are theoretical reasons to believe that this is not necessarily the case; Griffith (1995), in his investigation of school social climate in a large metropolitan school district, found that the variation was described more appropriately as related to the characteristics of individuals than of aggregates.

The social climate is one of those concepts that we all recognize intuitively but find difficult to define precisely. Generally it is thought to be the sum of the working conditions associated with a company, location, or work site that are created by specific organizational policies, practices, and behaviors.

The BOP tracks three scales that are believed to be related to social climate. Two of the scales, commitment to the BOP and satisfaction with supervision, can be regarded as individual viewpoints that grow out of experiences with the social climate. The third scale, evaluation of institutional operations, is a more direct assessment of organizational flexibility; this is thought to be a subdimension of the organizational practices that make up the social climate.

The items that constitute the institutional operations scale cover subjects ranging from the adequacy of formal communications channels to the formal authority structure. It is fairly obvious that this scale attempts to measure individual assessments of an organizational feature, namely the adequacy of the institution's formal organizational structure. (A complete list of the items making up the respective scales is presented in the appendix.)

The scale for commitment to the BOP measures respondents' identification with and intent to remain a part of the BOP. One could argue that commitment is an individual trait. This is not to suggest that the BOP commitment scale is not important, but only that it is not so dear that this scale captures an organizational as well as an individual phenomenon.[1]

The scale measuring satisfaction with supervision is designed to measure how well the respondents feel their immediate supervisors perform. Respondents are asked to rate how well their supervisors provide feedback about the respondents' work performance and solicit respondents' input about work-related matters. Thus it appears that a reasonable aggregate measure of the general satisfaction with supervision could be obtained simply by aggregating the individual scale scores for each institution.

The goals of the present analysis are to use the multilevel analytical techniques (Mason, Wong, and Entwisle 1983) developed since the seminal piece by Lincoln and Zeitz (1980), especially as codified by Bryk and Raudenbush (1992), (1) to examine the measurement properties of the social climate scales used by the BOP in the ESMI module, (2) to examine the scales for individual-level and organization-level sources of variation, and (3) to compare the rankings of the BOP facilities on the aggregate social climate scales with the rankings currently used by BOP managers, which are based on unadjusted aggregate scores.

Following the terminology of Bryk and Raudenbush (1992), we refer to multilevel models as hierarchical linear models (HLM). As Rowan, Raudenbush, and Kang (1991:208) observe,

---

[1]At the individual level, for example, Lincoln and Kalleberg (1990) argue that commitment to the organization is more important in getting maximum effort from workers than measures such as job satisfaction. Camp (1994), analyzing data from the Federal Bureau of Prisons, found that organizational commitment is indeed an important predictor of voluntary turnover among individuals.

"HLM procedures allow for explicit modeling at two levels of analysis so that the estimated effects of independent variables at one level of analysis are adjusted simultaneously for effects at the other level of analysis." Such an examination of the social climate scales more fully informs BOP management, and the management of other organizations, about how aggregate scales of the social climate can be used. In particular, the examination allows us to determine more precisely the extent of variation in social climates in the correctional facilities operated by the BOP and the reasons for any variation discovered, at least with respect to these three social climate measures. In addition, our examination creates the potential for further analyses that appropriately relate the individual and aggregate scores of social climate to theoretically interesting outcome variables.

Methodological Approach

In the statistical analysis presented here, we use simple two- level hierarchical models to examine the measurement properties of the social climate scales, working similarly to Rowan et al. (1991) in their examination of school climate. As in a traditional analysis of variance, the respondents' values are expressed as the mean value for the institutions in which they work. The within-correctional- institution model is

$$Y_{ij} = \beta_{0j} + r_{ij}$$

where Y_{ij} is the response for worker i at correctional institution j, β_{0j} is the mean for correctional institution j, and r_{ij} is the random error term.

The modeling of between-institution effects, in addition to the modeling of within-institution effects, makes this model hierarchical. The institution means are modeled as deviations from a grand mean. The between-institution model is expressed as:

$$\beta_{0j} = \gamma_0 + u_{0j}$$

where β_{0j} is the mean from the within institution model, γ_0 is the overall grand mean of all institutions, and u_{0j} is the random error term.

The simple within-institution and between-institution models provide the basis for all of the analysis reported here. In the random intercepts models considered in this analysis, the within-institution model is expanded to incorporate variables describing the individual characteristics of the respondents. The within-institution equations in the random intercepts models take the form:

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_1 + \beta_{2j}X_2 + \beta_{3j}X_3 + \ldots + \beta_{Pj}X_{Pij} + r_{ij}$$

where Y_{ij} is still the expected value of worker *i* at institution *j*. Y_{ij} is now conditioned not only on the intercept for the institution, β_{0j}, and the error, r_{ij}; Y_{ij} is also a linear function of the *P* independent variables describing the respondents' individual characteristics.

In the random-intercepts models, the between-institution models remain the same. This fact gives the models their name: Only the intercepts, the β_{0j}, at level 2 are allowed to vary. In hierarchical models, any of the β coefficients can be allowed to vary randomly. In this analysis, however, we are concerned primarily with differences in expected values between the different institutions after controlling for relevant variables; in other words, we are interested in the variation of the β_{0j}. In addition, we have no reason to suspect that the effects of the individual-level variables are not fixed across institutions. Therefore the effects of the other β coefficients are fixed in all analyses reported here.

After considering the random-intercepts models, we examine models that attempt to represent the between-institution differences. Thus the within-institution models remain the same as in the random-intercepts models, but now we further develop the level-2 models:

$$\beta_{0j} = \gamma_0 + \gamma_1 W_{1j} + \gamma_2 W_{2j} + \ldots + \gamma_Q W_Q + u_{0j}$$

where β_{0j} is the grand mean, the intercept, of the aggregate variables considered. The coefficients γ_1 through γ_Q are the fixed effects of the institution-level characteristics, the W_Q, on the intercepts. Stated differently, γ_1 through γ_Q represent the main effects of the group-level variables in raising or lowering the intercepts of the individual-level equations.

Data and Variables

The univariate statistics for the variables used in this analysis are presented in Table 1. The level-1, individual-level data come from the 1994 Prison Social Climate Survey (PSCS). The PSCS is described more completely in a conceptual paper by Saylor (1984); there Saylor explains how the PSCS data are collected specifically to allow for the type of multilevel analysis performed here.[2]

The dependent variables of interest are the three social climate scales measuring institutional operations, commitment to the HOP, and satisfaction with supervision.[3] Instead of using the percentage of favorable responses as an aggregate measure of the social climate scales, as in ESMI, we analyze the aggregate mean scores of the social climate scales. The scales range in value from 0 (most unfavorable) to 6 (most favorable), reflecting the seven choices presented to respondents for the Likert items constituting the scales. By using the mean scores we can capture the full range of information provided by respondents. The scales for commitment to BOP and

[2]When Saylor developed the PSCS, the HLM procedures used here had not been developed. Saylor, however, designed the data collection to facilitate analysis with ANCOVA models, which also separate total aggregated variable relationships into individual- and organizational-level components.

[3]The scales were first confirmed with factor analysis by Saylor; research results using the social climate scales were first presented by Saylor and Wright (1992; Wright and Saylor 1991, 1992).

satisfaction with supervision are transformed further by squaring the scale values to help normalize the measures.

Preliminary univariate analysis of the scales and analysis of the residuals from ordinary least squares models (not reported here) demonstrate that the transformations are desirable. We considered other transformations, but squaring the terms produces the most stable statistical results. The substantive conclusions, however, remain the same with and without transformation of the measures of BOP commitment and satisfaction with supervision.

We use several variables in this analysis to control for the respondents' potentially different socialization experiences: ethnicity, gender, age, job tenure, education, transfer status, and race. Most of these variables affect outcomes in social science research. Ethnicity takes a value of 1 for Hispanic respondents.[4] Gender takes a value of 1 for females. Age and BOP job tenure are modified by logarithmic transformations to manage skewed distributions. Education is measured with a dummy variable: A score of 1 indicated that the respondent has at least a four-year college degree. Transfer status is coded 1 for individuals who have transferred from one BOP facility to another: Much of the folklore among workers at the BOP concerns the differences between "homesteaders," who make their careers at one institution, and career employees, who are willing to move for promotions. Two dummy variables for race compare blacks with whites, and others (everyone who is neither black nor white) with whites.

The rest of the individual-level variables we use are entered to control for the respondents' different locations in the structures of their respective correctional facilities. We enter dummy

[4]The question about ethnic status, Hispanic or not, is separate from the item about race. Therefore ethnicity is not confounded with race in this analysis.

variables to control for supervisory status, work site, and department. Respondents who identify themselves as supervisors receive a score of 1 on the supervisor dummy variable. The variable measuring location of work site, SATELLITE, takes a value of 1 for respondents reporting that they spent most of the last six months assigned to sites other than the main facility. Many BOP institutions with inmates at a higher security level have satellite facilities such as minimum-security camps and intensive confinement centers (ICCs), commonly known as boot camps. It is not possible, however, to separate respondents working at ICCs from respondents working at camps.

In the Prison Social Climate Survey, respondents choose one of 10 broad departmental designations as representing their work assignment. We constructed nine dummy variables to compare workers in nine general department designations with workers in correctional services, the excluded category. Especially interesting here are the comparisons of workers in unit management and programming with those in correctional services. Although all workers in federal correctional facilities treat custody as a major priority, workers in correctional services are primarily responsible for that function. As a result they sometimes have different priorities than workers in unit management and programming, who also work closely with inmates.

Unit management was first adopted by the BOP in 1973. Under this system, a correctional facility is divided into a number of small, self-contained housing units. Unit managers, whose offices are located in housing units, are responsible for maintaining and reviewing inmates' records (Ingram 1987:205-207). Workers in programming departments, education, recreation, religion and psychology services provide programs to enhance the lives of inmates both within the institution and upon release. Many writers on corrections point to the tensions between custody

and programming functions (Cullen, Lutze, and Link 1989; Jacobs 1977; McShane and Williams 1989).

The level-2, organizational-level data reported in Table 1 are taken from operational databases used to generate statistical information reported in the BOP's Key Indicators"Strategic Support System and the Executive Staff Management Indicator module. Two of the institutional variables considered here represent organizational properties often used in organizational analyses. Size of the workforce controls for the amount of bureaucratization in daily operations. Although all institutions are part of a common bureaucratic entity, the BOP, they differ in size; the numbers of workers at institutions vary from 78 to 723. Age of the facility is considered as a rough control for the quality of working conditions.[5]

Controls for security level are considered for two reasons. First, the security level designation of an institution depends in part on the security procedures, or technology, employed. Obviously, workers at maximum-security institutions do some things differently than workers at minimum-security camps, though there are probably more similarities than differences. In any case, security level tends to be a control for technology. Second and more important, security level is a control for the dangerousness of inmates. Just as Rowan, et al. (1991:217) find that teachers' ratings of students' abilities influence evaluations of a school's social climate, it is reasonable to expect that the perceived dangerousness of inmates affects the social climate of a prison.

Region of the BOP is included as a general control for geographic location in the United States. We include the rest of the organization-level variables to test for any composition effects

[5]We thank one of the reviewers for this insight into the use of the facility'' age as a control variable.

of selected variables entered at the individual level. These variables include measures of the proportion of the workforce who are black, who are female, who have a college degree, and who belong to the 'other' racial category. The composition effects of median age and median BOP job tenure are also investigated. Composition effects measure the effects above and beyond individual-level effects. For example, does an institution's racial composition affect the social climate scales above and beyond the effects expressed by individuals of different racial backgrounds?

RESULTS

Scale Measurements

The results presented in Table 2 demonstrate some important features of the social climate scales. First, researchers traditionally have reported Cronbach's alpha, which measures the internal consistency, or reliability, of the items constituting the scales for individual respondents. The internal consistency, as reported by Cronbach's alpha in Table 2, is at least .865 for all of the social climate scales. Many analysts, with this information in hand, would simply proceed to aggregate the scales, which have high reliability at the individual level, into aggregate measures. They would assume that mean differences in the aggregate measures are reliable indicators of the variability in attributes of the organizations under investigation. This procedure, however, is not adequate as Saylor (1984) observes in his outline of the use of scale measures derived from PSCS data. The development of new technology, particularly the hierarchical linear model (HLM) procedures used here, allows for more complete investigation of the measurement properties of the scales.

Table 2, following the lead presented by Rowan et al. (1991) in their analysis of the working climate in the U.S. public high schools, presents additional information about the measurement properties of the social climate measures. The information, generated by HLM methods described in Bryk and Raudenbush (1992) and implemented in the HLM/2L software distributed by Scientific Software, Inc. (Byrk, Raudenbush,and Congdon 1994), allows us to assess the reliability of the scales when used as aggregated organizational measures.

The first piece of additional information about the usefulness of the scales, aggregated to the organizational level, is provided by the intraclass correlation. One way of thinking about this correlation is to view it as measuring how consistently the respondents within each organizational unit (in our case, correctional facilities) rate the organizational measures of interest: Institutional operations, commitment to BOP, and satisfaction with supervision. As shown in Table 2, the intraclass correlations are quite low, ranging from .017 for satisfaction with supervision to .058 for institutional operations. These values indicate that there is very little consistency in the ratings of the organizational properties provided by respondents working in the same institutions.

Another way of considering the intraclass correlation is to view it as a measure of the relative variation explained between and within organizational units. By definition, the intraclass correlation is computed as the ratio of between-organization variance to total variance. Thus, for institutional operations, the intraclass correlation coefficient of .058 means that 5.8 percent of the total variance in institutional operations is explained by variation between correctional facilities. Conversely, 94.2 percent of the variation in the evaluations of institutional operations occurs within correctional facilities, or at the individual level. Only 4.6 percent of the variance in BOP commitment occurs between correctional facilities, and only 1.7 percent of the variance in

satisfaction with supervision is explained in this way. Although these amounts are only small percentages of the total variance in our measures of interest, they are statistically significant; also, we examine the between-organization variance in the next step of the analysis.

The second additional piece of information provided in Table 2, concerning the usefulness of the scales aggregated to the organizational level, is the reliability measure. In Table 2 this refers to the reliability of the correctional facility means, or aggregate measures. For each institution, the reliability of the institutional mean is computed as

$$\lambda_j = J / (J + F^2/n_j)$$

where λ_j is analogous to a ratio of true score variance to total score variance (Byrk et al. 1994). The component for between-institution variance is J, F^2 is the within-institution variance, and n_j is the number of respondents at correctional institution *j*. As shown in Table 2, the adjusted averages of the reliabilities of the means are less than the internal consistencies for each measure. In the case of satisfaction with supervision, the reliability of the aggregate measure, .554, is low enough to cause concern about the suitability of using an aggregate measure for this variable. The lower reliabilities of the aggregate measures are to be expected, however, because the reliabilities are affected by intraclass correlation.[6] As noted above, the agreement among individuals at the different correctional facilities (the intraclass correlation) is low for all of the measures.

[6]Other factors that affect the reliabilities are the number of items in the scale, the internal consistency of items in the scale, and the number of respondents sampled at each correctional facility (Rowan, Raudenbush, and Kang 1991:213). None of these factors unduly influence the reliabilities of the scales examined here.

Random Intercepts Models

The second type of model considered here is the random-intercepts model, into which all of the individual-level variables are entered. No group-level variables are entered; the intercepts for the different institutions are allowed to vary randomly. The random-intercepts models allow us to determine whether significant between-group variation remains after controlling for the individual-level effects.

The results of the random-intercepts models are presented in Table 3. As shown in the random-effects portions of the results presented for the three measures, the between-institution portions of the variances (τ_0^2) remain statistically significant even after con- trolling for individual-level sources of variation.[7]

Table 3 reveals some interesting effects for the individual-level variables. Hispanics and non-Hispanics rate only one social climate scale differently. Hispanics report more commitment to the BOP than do non-Hispanics. Males and females do not rate any of the scales differently. Supervisors, on the other hand, consistently rate the scales higher than nonsupervisors. Race is statistically significant at the .05 level in two of the models. In the model of institutional operations, blacks rate institutional operations more favorably than whites; in the model of satisfaction with supervision, members of the "other" racial category report less satisfaction than whites. Tenure has a consistent, negative effect on evaluations of social climate. Respondents with at least a bachelor's degree report lower BOP commitment and less satisfaction with supervision. Respondents who report that they spend most of their time working at a satellite of the main

[7]The proportion of the variance (not shown here) explained by the between- institution portion is still small for all of the measures; it does not change much from the values presented for the measurement results reported in Table 2.

institution, either a minimum-security camp or an intensive confinement center (ICC), make the same evaluations of the social climate scale as respondents working in the main facility. Older workers evaluate the institutional operations and BOP commitment scales more favorably. Finally, workers in different departments within the institution often evaluate the work environment differently. Particularly interesting are the comparisons of workers in correctional services with those in unit management and programming: Correctional service workers evaluate all three social climate scales lower than workers in either of the other areas.

Table 4 displays a comparison of the institutions on the unadjusted and adjusted scores for the social climate scales. The unadjusted scores are the mean values for the respondents at each of the 74 institutions. The adjusted scores are the intercepts, the $\$_{0j}$ for each of the 74 institutions as produced by the random-intercepts models. Because of the coding of the independent variables, the intercepts in these model. are the expected scores of the "typical" BOP workers at each institution.[8] The adjusted and unadjusted scores are ranked and compared with Spearman's rank correlation coefficients. (Spearman's coefficients are nonparametric comparisons of the rank orders, not the actual scores, as in Pearson product-moment correlations.)

As demonstrated in Table 4, the unadjusted and the adjusted scores produce comparable but not identical rankings. The Spearman's rank correlation coefficients are near unity for each of the measures. Even so, this table presents some interesting substantive differences. Wardens at BOP facilities are particularly eager to keep their institutions from falling into the bottom decile.

[8]Typical is used here in a nonevaluative sense and refers only to the most common sociodemographic characterization of a worker at the BOP. That is, a typical worker is white, non-Hispanic, and male, does not have a college degree, has never transferred within the BOP, works at the main facility, is a nonsupervisor, has average BOP amounts of tenure and age, and works in correctional services.

Table 4 contains a list of the institutions in the bottom decile for each of the social climate scales on the basis of the adjusted scores. For institutional operations, Institution A has the lowest aggregate, adjusted score; Institution G has the seventh lowest score on institutional operations. Institution A also has the lowest score, as indicated by the ran]dng in parentheses, for institutional operations on the original values, the unadjusted means. Institution G has the fifth lowest score for institutional operations on the basis of the original, unadjusted scores. Again, this information is presented in parentheses.

Given the high values for the Spearman's rank correlation coefficients, we are not surprised to see so much overlap between institutions in the bottom decile, whether the rankings are based on adjusted or unadjusted scores. For BOP commitment, the same institutions fall into the bottom decile whether adjusted or unadjusted scores are used in constructing the rankings, though the order is not the same in each case. For the other two scales, however, different institutions occupy the bottom decile when adjusted scores are used. Institutions E and F are not in the bottom decile for institutional operations when unadjusted scores are used, but are so when ranking is based on adjusted scores. Similarly, Institutions K, L, and M escape the bottom decile of the rankings of satisfaction with supervision unless adjusted scores are used to construct the rankings. In fact, for the unadjusted scores of satisfaction with supervision, Institution M ranks twentieth among the 74 BOP institutions on satisfaction with supervision, well above the lowest decile cutoff of 7. Thus, in light of the results presented above, especially the significance of J in all of the models, it is necessary to model the between-institution variance.

Models of Organization-Level Variation in the Social Climate Scales

Modeling the between-institution variance fails completely when the organizational characteristics considered are included in the hierarchical linear models. The organizational-level variance is modeled in three ways; all produce equally disappointing results. First, all of the organizational variables listed in Table 1 are entered into level 2 as predictors of the β_0. Almost none of these variables are statistically significant.[9] Consequently we reestimate the models of between-institution variance with subsets of the entire list of organizational variables. One set of models includes only the security-level indicators; another set includes only the variables measuring the compositional effects of race, gender, age, BOP job tenure, and college education. None of the variables are statistically significant in any of the models examined.[10]

We are surprised by the lack of finding for any of the organizational variables considered here; obviously it calls for explanation. We expected in particular to find effects for security level. The lack of finding for the organizational variables, however, is consistent with the influence of local management over local social climates. Although it is less than desirable to eliminate other factors to conclude that a factor not directly tested is important, this is our only alternative in this analysis. We discuss these points in greater detail below. For now, however, we mention that the small number of institutions, 74, and the small amount of variation to be explained also may contribute to the nonfindings.

---

[9]The only minor exception to this general pattern is the model of BOP commitment. In the model with all of the organizational variables, two of the indicators for region assume statistical significance. Yet because of the nonsignificance of the other variables In this model and in the models of institutional operations and satisfaction with supervision, we attach little importance to these results.

[10]To save space, we do not report the results of these models; they are available on request.

Discussion

The most surprising result of our analysis is the inability to model any of the between-institution variance. Part of this failure, as we mentioned above, may be due to the small amount of variance to be explained at the organizational level. Even for the institutional operations scale—the aggregate social climate measure that most clearly measures an organizational property—the portion of variance due to organizational-level processes is only about 5.8 percent. For the other scales, the portion of the variance due to organizational-level processes is even smaller. Most of the variance exists among individual raters.

The failure to model between-institution variance also may be due in part to the relatively small number of institutions, which reduces the statistical power of the tests at the organizational level. In addition, however, the failure is probably due to the lack of theoretical understanding about the choice of variables to include at the organizational level. This is true especially in an analysis such as this, in which only one bureaucratic entity is under examination. The BOP is a highly centralized agency that continually reviews institutions to ensure policy compliance. By way of comparison, Rowan et al. (1991) found that 7 to 24 percent of the variance in their measures of the working environment of schools resides between schools. Admittedly they use different measures than we do, but they also analyze a random sample of *all* schools in the United States. Different school districts are organized, funded, and managed differently; this probably has much to do with the between-school variance they find. Similarly, to examine variance between correctional institutions, we need a comparative analysis of correctional institutions under different forms of bureaucratic control, such as the BOP in the United States and the Correctional

Services of Canada (CSC). A comparison with CSC is feasible because CSC already has collected similar data. In addition, states that collect comparable data could serve as comparison points.

As noted above, the inability to explain the organizational sources of variance in the aggregate measures, especially institutional operations, is consistent with the notion that local management has some influence on the local social climate. Although we cannot directly test this idea in our models, we think it likely that at least part of the organizational variance is caused by differences in implementation of policy by different managers at the different BOP prisons. Historically, wardens have exercised much control over populations at their facilities.

Conclusions

The use of the three social climate measures examined here as aggregate evaluations of organizational properties is not without ambiguity. Conceptually the institutional operations scale most clearly evaluates an organizational property. It also appears to be a reasonable candidate empirically. Although the amount of between-institution variance is not particularly large, it is statistically significant. The measurement properties of the scale, at both the individual and the organizational level, are sound. Probably it is not surprising that this appears to be the best measure organizationally. This measure has items that ask most directly about organizational properties, particularly the flexibility and effectiveness of the institution's formal structure in allowing respondents to perform their jobs.

Because of the relatively small amount of organizational-level variance in the aggregate measure of institutional operations, we still urge some restraint in the use of institutional operations to uncover intrafirm organizational differences. If an aggregate measure of institutional operations is to be used, we recommend constructing an adjusted score to partial out the

explained sources of individual-level variability. In fact, these organizational measures are probably useful. Given the results presented here, we cannot rule out the possibility (and we think it is highly likely) that the aggregate value for institutional operations reflect aspects of the effectiveness of institutional management.

We think that more caution is warranted in using the other two scales as aggregate measures. The measurement properties of the commitment scale are adequate; that is not the problem. Rather, the problem with commitment lies in how to interpret the aggregate measure. Institutions do not have commitment; commitment is a property of individuals. Institutions, however, have practices and structures that encourage organizational commitment. Yet the commitment scale does not directly measure these features, which would provide more appropriate organizational measures. We encourage researchers to more directly measure organizational elements that increase individuals' commitment to their organization rather than using aggregations of organizational commitment as a proxy for those elements.

The scale for satisfaction with supervision presents clear measurement problems. The intraclass correlation for this scale is much lower than for the other two scales; the reliability of the scale measure, as measured by

$\rho_j = J / (J + F^2)$ and reported in Table 2, is so low that we cannot recommend its use as an organizational measure.[11]

[11]Our investigation into the organizational properties of the aggregate social climate scales does not necessarily negate the use of the measures as aggregate summary scores. As long as one understands that a summary score is a summarization of individual properties, there is no problem in using any of the aggregate social climate measures. We believe, however, that when aggregate information is presented in a manner such as that used by the BOP, it encourages use of the measures as performance indicators of the effectiveness of institution management. We cannot state too strongly that the analysis presented here does not provide justification for using any but the aggregate institutional operations scale in this manner. institutional operations scale to rank BOP facilities. We also recommend the continued use of the BOP commitment scale at an aggregate level as long as

We believe that this analysis provides a strong methodological caution against the unthinking use of aggregate measures. Whether or not aggregate measures are explicitly intended to represent organizational properties, we emphasize that computing such measures encourages those who use them to view them as if they represent organizational properties. Therefore it is necessary to move beyond examining the properties of the measures at the individual level, and to examine those properties at the organizational level.

Theoretically we see that there is some hope of using aggregate measures of the social climate to capture intrafirm differences created by differences in local managers implementations of policy and practice. Yet this statement is tempered by the relatively small amount of variance that is associated with organizational-level processes, at least for the measures considered here. Because of the focus on only one organization, the BOP, one cannot distinguish whether the small amount of variance at the organizational level is due to limitations in our measures or to limitations in the variability across institutions. We must assess the measures' ability to capture variability across different organizations before we can begin to sort out these issues. Such an analysis will allow us to understand more clearly how strongly the social climate is influenced by local management practices, as opposed to overall corporate policies, practices, and the like. Although we lack the comparative framework in this case study of BOP practices, we think we have shown a way to address this important theoretical and policy topic.

---

those using the scale understand it is a quasi-organizational measure at best. And, finally, there are too many methodological and measurement problems associated with the satisfaction with supervision scale to recommend its use as an aggregate measure, although it may be useful as a summarization of the individual attitudes of respondents towards their immediate supervisors.

References

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical Linear Models: Applications and Data Analysis Methods*. Newbury Park, CA: Sage.

Bryk, A. S., Raudenbush, S. W., & Congdon, R. T. (1994). *Hierachical Linear Modeling with the HLM/2L and HLM/3L Programs*. Chicago, IL: Scientific Software International.

Camp, S. D. (1994). Assessing the Effects of Organizational Commitment and Job Satisfaction on Turnover: An Event History Approach. *The Prison Journal, 74*(3), 279-305.

Cullen, F.T., F.E. Lutze, and B.G. Link. 1989. "The Correctional Orientation of Prison Guards: Do Officers Support Rehabilitation?" *Federal Probation* 53: 33-42.

Gilman, E. (1991). Implementing Key Indicators. *Federal Prisons Journal*, *2*(3), 48-56.

Griffith, J. 1995. "Does the School Make a Difference? Statistics for Assessing Group-Level Effects." Paper presented at the meeting of the Society oo the Study of Social Problems, Washington, D.C.

Guillén, M. (1994). *Models of Management: Work, Authority, and Organization in a Comparative Perspective*. Chicago: University of Chicago Press.

Hauser, R. M. 1970. "Context and Consex: A Cautionary Tale." *American Journal of Sociology* 74: 587-611.

Ingram, G. (1987). The American Federal System: Control and Classification. In A. E. Bottoms & R. Light (Ed.), *Problems of Long-Term Imprisonment*(pp. 200-213). Brookfield: Gower.

Jacobs, J.B. 1977. *Stateville: The Penitentiary in Mass Society*. Chicago: University of Chicago Press.

Kopelman, R. E., Brief, A. P., & Guzzo, R. A. (1990). The Role of Climate and Culture in Productivity. In B. Schneider (Ed.), *Organizational Climate and Culture*(pp. 282-318). San Francisco: Jossey-Bass.

Lazarsfeld, P. F., & Menzel, H. (1961). On the Relation Between Individual and Collective Properties. In A. Etzioni (Ed.), *Complex Organizations: A Sociological Reader*. Glencoe, IL: Free Press.

Lincoln, J. R., & Kalleberg, A. L. (1990). *Culture, Control, and Commitment: A Study of Work Organization and Work Attitudes in the United States and Japan*. New York: Cambridge University Press.

Lincoln, J. R., & Zeitz, G. (1980). Organizational Properties from Aggregate Data: Separating Individual and Structural Effects. *American Sociological Review, 45*, 391-408.

Mason, W. M., Wong, G. Y., & Entwisle, B. (1983). Contextual Analysis through the Multilevel Linear Model. In S. Leinhardt (Ed.), *Sociological Methodology*. San Francisco: Jossey Bass.

McShane, M.D. and F.P. Williams, III. 1989. "Running on Empty: Creativity and the Correctional Agenda." *Crime and Delinquency* 35: 562-75.

Muth, W. R. (1995). *Implementing a Management Information Tool for Executives in the Federal Prison System* . Federal Bureau of Prisons.

National Performance Review. 1993. *Creating Government that Works Better and Costs Less*. Washington, D.C.: U.S. Government Printing Office.

Raudenbush, S.W. and J.D. Willms, eds. 1991. *Schools, Classrooms, and Pupils: International Studies of Schooling from a Multilevel Perspective*. New York: Academic Press.

Rowan, B., Raudenbush, S. W., & Kang, S. J. (1991). School Climate in Secondary Schools. In S. W. Raudenbush & J. D. Willms (Ed.), *Schools, Classrooms, and Pupils: International Studies of Schooling from a Multilevel Perspective*. San Diego, CA: Academic Press.

Saylor, W. G. (1984). *Surveying Prison Environments* . Federal Bureau of Prisons.

Saylor, W. G. (1988). *Developing a Strategic Support System: Putting Social Science Research into Practice to Improve Prison Management* . Federal Bureau of Prisons.

Saylor, W. G., & Wright, K. N. (1992). Status, Longevity, and Perceptions of the Work Environment among Federal Prison Employees. *Journal of Offender Rehabilitation, 17*(3/4), 133-160.

Wright, K. N., & Saylor, W. G. (1991). Male and Female Employees' Perceptions of Prison Work: Is There a Difference. *Justice Quarterly, 8*(4), 505-524.

Wright, K. N., & Saylor, W. G. (1992). A Comparison of Perceptions of the Work Environment Between Minority and Nonminority Employees of the Federal Prison System. *Journal of Criminal Justice, 20*, 63-71.

Table 1
Univariate Statistics

Level-1 Descriptive Statistics

Var. Name	Description	N	Mean	S.D.	Min	Max
INSTOP SCALE	Institutional operations scale*	8,086	3.34	1.18	0.00	6.00
SUPERV SCALE	Squared satisfaction with supervision*	8,078	14.57	9.65	0.00	36.00
BOPCOM SCALE	Squared BOP commitment*	8,043	18.42	9.38	0.00	36.00
ETHNIC	Ethnicity (Hispanic coded 1)	8,019	0.11	0.31	0.00	1.00
GENDER	Gender (female coded 1)	8,007	0.25	0.43	0.00	1.00
SUPERVISOR	Supervisor (supervisor coded 1)	8,041	0.18	0.38	0.00	1.00
OTHER RACE	Race (other coded 1)	8,035	0.11	0.31	0.00	1.00
BLACK	Race (black coded 1)	8,035	0.16	0.36	0.00	1.00
BOP TENURE	Natural logarithm of BOP tenure	7,740	1.82	0.69	0.00	3.71
COLL_ED	College education (> B.S. coded 1)	7,819	0.35	0.48	0.00	1.00
SATELLITE	Work location (satellite facility coded 1)	8,115	0.10	0.30	0.00	1.00
AGE	Natural logarithm of age	8,020	3.56	0.19	3.00	4.23
UNIT_MAN	Unit management coded 1	8,022	0.10	0.30	0.00	1.00
MECH_SVC	Mechanical services, safety coded 1	8,022	0.08	0.28	0.00	1.00
UNICOR	Prison industries coded 1	8,022	0.05	0.21	0.00	1.00
PUB_HLTH	Public Health Services coded 1	8,022	0.03	0.17	0.00	1.00
HLTH_SVC	Health services coded 1	8,022	0.08	0.26	0.00	1.00
FOOD_SVC	Food services coded 1	8,022	0.05	0.21	0.00	1.00
PROGRAMS	Programming† coded 1	8,022	0.08	0.26	0.00	1.00
FIN_MAN	Financial management, etc.‡ coded 1	8,022	0.18	0.38	0.00	1.00
OTHER	Other department coded 1	8,022	0.02	0.14	0.00	1.00

* Measured on a 7 point Likert scale from unfavorable evaluation (0) to favorable evaulation (6).

† Education, Recreation, Religion, and Psychology Services.

‡ Financial Management, Human Resources, Inmate Systems, Research, Institution Executive Staff, Facilities, Information Systems, and Legal

Table 1 - Continued

Level-2 Descriptive Statistics

Var. Name	Description	N	Mean	S.D.	Min	Max
SIZE	Size of work force	74	305.91	132.53	78.00	723.00
MEDIAN AGE	Median age of work force	74	35.36	1.83	31.00	39.80
PCT_BLACK	% black in work force	74	0.18	0.15	0.00	0.61
PCT_FEMALE	% female in work force	74	0.25	0.09	0.06	0.54
MEDIAN_TEN	Median tenure of work force	74	4.95	1.63	1.00	8.90
PCT_COLL	% of work force with college degree	74	0.30	0.06	0.16	0.44
PCT_OTHER	% other race in work force	74	0.14	0.16	0.02	1.00
AGE_FAC	Natural logarithm of age of facility	74	2.60	1.22	0.69	4.52
RG_MXR	Indicator for Mid-Atlantic region	74	0.00	0.57	-1.00	1.00
RG_NCR	Indicator for North Central region	74	0.00	0.57	-1.00	1.00
RG_NER	Indicator for North East region	74	0.04	0.61	-1.00	1.00
RG_SCR	Indicator for South Central region	74	0.00	0.57	-1.00	1.00
RG_SER	Indicator for South East region	74	-0.01	0.56	-1.00	1.00
SL_HI	Indicator for high security	74	-0.07	0.53	-1.00	1.00
SL_LO	Indicator for low security	74	0.07	0.65	-1.00	1.00
SL_MD	Indicator for medium security	74	0.11	0.67	-1.00	1.00
SL_MN	Indicator for minimum security	74	0.01	0.61	-1.00	1.00

Table 2
Measurement Properties of the Aggregate Social Climate Measures

	Institutional Operations	BOP Commitment	Satisfaction with Supervision
Within-Institution Variance	1.267	81.616	89.350
Between-Institution Variance	0.078	3.951	1.583
Intraclass correlation	.058	.046	.017
Reliability	.807	.768	.554
Cronbach's "	.865	.897	.922
Number of Items in Scale	10	5	10

Table 3
Random Intercepts Models for the Social Climate Measures

	Institutional Operations	BOP Commitment†	Satisfaction with Supervision†
Fixed Effects			
INTERCEPT	3.085**	16.934**	11.443**
ETHNIC	0.100	1.391**	0.831
GENDER	-0.059	-0.068	0.325
SUPERVISOR	0.505**	3.613**	3.298**
OTHER RACE	-0.111*	-0.855*	-1.063**
BLACK	0.165**	0.072	0.342
BOP TENURE (LOG)	-0.223**	-1.388**	-0.599**
COLL_ED	-0.005	-0.671**	-0.745**
SATELLITE	-0.017	-0.205	0.157
AGE (LOG)	0.467**	2.406**	0.895
TRANSFER	0.054	1.028**	-0.347
UNIT_MAN	0.298**	1.285**	5.542**
MECH_SVC	0.000	0.039	3.555**
UNICOR	0.039	0.807	2.525**
PUB_HLTH	0.012	-1.211	6.219**
HLTH_SVC	0.063	-0.271	4.059**
FOOD_SVC	0.156**	0.417	2.938**
PROGRAMS	0.316**	1.750**	6.010**
FIN_MAN	0.318**	1.883**	4.116**
OTHER	0.555**	2.616**	6.745**
Random Effects (Variance Estimates)			
F^2	1.258	80.539	85.299
J	0.067**	3.198**	1.677**

† BOP Commitment and Satisfaction with Supervision are transformed by squaring the terms.
* p # .10
** p # .05

Table 4
Correspondence Between Original (Unadjusted) Rankings
of Institutions and Adjusted Rankings

	Institutional Operations	BOP Commitment†	Satisfaction with Supervision†
Lowest Decile, Rankings on Adjusted Scores, Low to High (Original Rank):	A* (1)	A (2)	D (1)
	B (2)	H (1)	C (2)
	C (4)	I (6)	K (10)
	D (3)	C (7)	L (9)
	E (8)	B (3)	B (4)
	F (15)	J (5)	M (20)
	G (5)	G (4)	N (5)
Spearman's Rank Correlation Coefficient	.95	.96	.91

* To protect the confidentiality of information about the individual BOP facilities, we have replaced the institution names with letter designations. The letter designations are consistent across the scales, so that institution A represents the same institution in the rankings of the institutional operations scale and the BOP commitment scale.

Appendix 1
Questionnaire Items in the PSCS Used to Create Social Climate Scales

The Commitment to the Bureau of Prisons scale is composed of the following 5 items from the work environment section of the Prison Social Climate Survey (PSCS).

1. I have a good opinion of the BOP most of the time.
2. Most of the time the BOP is run very well.
3. I am usually satisfied with the BOP.
4. The BOP is better than any of the other correctional agencies (e.g., State).
5. If I remain in corrections, I would prefer to remain the BOP.

The Institutional Operations scale is composed of the following 10 items from the work environment section of the PSCS.

1. The information I get through formal communications channels helps me to perform my job effectively.
2. In the BOP, it is often unclear who has the formal authority to make a decision.
3. It's really not possible to change things in the institution.
4. I am told promptly when there is a change in policy, rules, or regulations that affects me.
5. I have the authority I need to accomplish my work objectives.
6. Employees do not have much opportunity to influence what goes on in the BOP.
7. Under the present system, promotions are seldom related to employee performance.
8. Management at this institution is flexible enough to make changes when necessary.
9. In the BOP, authority is clearly delegated.
10. In general, this institution is run very well.

Note: The negatively oriented items (#2, 3, 6, and 7) are reversed in order to combine them with the positively oriented items.

The Satisfaction with Supervision scale is composed of the following 10 individual items in the work environment section of the PSCS.

1. My supervisor engages me in the planning process, such as developing work methods and procedures for my job.
2. My supervisor gives me adequate information on how well I am performing.
3. My supervisor asks my opinion when a work-related problem arises.
4. I have a great deal of say over what has to be done on my job.
5. On my job, I know what my supervisor expects of me.
6. The standards used to evaluate my performance have been fair and objective.
7. The information I receive about my performance usually comes too late for it to be of any use to me.
8. My last annual performance rating presented a fair and accurate picture of my actual job performance.
9. My own hard work will lead to recognition as a good performer.
10. I often receive feedback from my supervisor for good performance.

Note: The negatively oriented item #7 is reversed in order to combine it with the positively oriented items.