

Using Inmate Survey Data in Assessing Prison Performance:  
A Case Study Comparing Private and Public Prisons

Scott D. Camp, Ph.D.  
Gerald G. Gaes, Ph.D.  
Jody Klein-Saffran, Ph.D.  
Dawn M. Daggett, M.S.  
William G. Saylor, M.A.

April 17, 2002

Office Research and Evaluation  
Federal Bureau of Prisons  
320 First Street NW  
400 Building, Room 3025  
Washington, D.C. 20534

contact: [scamp@bop.gov](mailto:scamp@bop.gov)  
202-616-7248 (Camp)

The views expressed in the paper are those of the authors and do not necessarily represent the views of either the Federal Bureau of Prisons or the US Department of Justice. We would like to acknowledge and thank the editor, Michael Vaughn, and the anonymous reviewers at *CJR* for their extremely helpful comments. The discussion is much more focused as a result of their input.

This paper appeared in *Criminal Justice Review* 27: 1: 26-51 (2002).

Using Inmate Survey Data in Assessing Prison Performance:  
A Case Study Comparing Private and Public Prisons

## Using Inmate Survey Data in Assessing Prison Performance: A Case Study Comparing Private and Public Prisons

### **Abstract**

The public sector needs to monitor the performance of the private prisons, and it is necessary to conduct the monitoring as objectively as possible. This paper demonstrates that an often overlooked source of data, surveys of inmates, can be used to differentiate prisons on areas such as gang activity, safety and security, sanitation, and food service delivery. Hierarchical linear models were used to generate the prison performance measures. Second, we show that inmates and staff largely agreed in their assessments of conditions at the prison. Finally, we demonstrate that while there is considerable consistency for different measures within the topical areas examined, there is no necessary correspondence in performance across the different topical areas of gang management, safety and security, sanitation, and food service delivery. While surveys will and should never replace operational reviews and audits, we demonstrate that they can be effectively used to obtain information about operational differences between prisons.

## **Using Inmate Survey Data in Assessing Prison Performance: A Case Study Comparing Private and Public Prisons**

The emergence of private adult prisons in the United States in the 1980s generated an industry of experts to debate the relative merits of private and public prisons. Participants in the debate included academics, policy-makers, elected officials, stockholders, and others. Those involved in the discussions about private prisons tended to take firm stands either for or against private prisons, and often there was more smoke than fire in the positions advanced.<sup>1</sup> Even when attention turned to empirical matters, the results were less than conclusive (see the reviews by Austin & Coventry, 2001; Gaes, Camp, & Saylor, 1998; General Accounting Office, 1996).<sup>2</sup> Rather than review the often contentious writings on prison privatization, we focus upon one of the beneficial results that evolved out of the debates about prison privatization, the increased attention on measuring prison performance. We frame the discussion of prison performance within the context of a case study of prison privatization.

---

<sup>1</sup>Proponents of prison privatization typically focus on costs, claiming that cost savings are practically guaranteed by free market competition (Moore, 1998; Segal & Moore, 2002; Thomas, 1997). Proponents also claim that prison quality is maintained or more typically improved in comparison to public prisons. A more sophisticated argument is that too much emphasis has been placed on costs. Instead, attention should be focused on how competition introduced by privatization improves overall system performance in the public sector (Harding, 2001). Critics of privatization deny these claims on ethical, legal, and empirical grounds (Shichor, 1995). Empirically, they point that competition for prison beds occurs in a constrained market with parallels to contracting problems in the defense industry. Likewise, they fear that the focus upon profits threatens the level of care as spending may be cut to insure profits. Other analysts point to the inability of the private sector to recruit and maintain an experienced workforce (Camp & Gaes, 2002).

<sup>2</sup>Many of the cost studies done to date have lacked scientific rigor and generalizability (General Accounting Office, 1996; Nelson, 1998). For example, a recent policy piece reviewed cost studies comparing public and private prisons (Segal & Moore, 2002), but it ignored a major study by Nelson (1999) that found evidence counter to their claim that private prisons save money. Likewise, the Segal and Moore study ignored the work done by the lead author and his colleagues on prison performance measurement as related to comparing public and private prisons (Camp, 1999; Camp & Gaes, 2000a, 2000b, 2002; Camp, Gaes, & Saylor, 2002; Camp, Saylor, & Wright, 1999). As McDonald and his colleagues noted (1998), one of the more noticeable aspects of research on privatization is that few studies have been conducted to date.

The present analysis addresses two points. First, the present study validates a source of data typically not used in evaluating prison performance in the United States, inmate survey data. In prior analyses, Camp and colleagues (Camp et al., 2002; Camp, Saylor, & Harer, 1997) used hierarchical linear models (Bryk & Raudenbush, 1992) to analyze staff survey data to compare prisons. The same techniques were used in the present analysis to see if these data were amenable to the same approach. Where most of their work focused on survey data collected from staff, we focused upon data collected from inmates. Where possible, the present analysis went beyond the internal validity of the inmate survey data to compare and triangulate the results with performance measures generated from staff survey data. We spent most of our efforts on the issue of data reliability and validity since a demonstration of using data collected from inmates does not exist currently in the literature with the sole exception of Camp (1999). The second point addressed by the paper is much more interesting substantively. We use the method demonstrated here to generate comparisons of one private prison to three public prisons. Data were also collected from other comparable institutions in the prison system to provide context in evaluating the institutions with which we were most concerned. The additional prisons were necessary to fulfill the methodological requirements of our statistical techniques. From a theoretical and practical perspective, it is much more informative to evaluate the targeted prisons, public or private, in relation to the performance of the entire system of prisons.

The inmate data reported in the current analysis were collected at the Federal Bureau of Prisons (BOP) and at a private prison operated by Wackenhut Corrections Corporation (WCC) in Taft, California for the BOP. The BOP is one of the larger consumers of private-prison beds in the world, but at the time these data were collected, the Taft prison was the only private prison contracted by the BOP that held general population inmates. Most private prisons under contract

with the BOP at the time, and currently, held illegal aliens who were subject to deportation at the expiration of their sentences.

Details about the operation of the Taft facility are provided for those unfamiliar with the issues surrounding the operation of this prison. Following this brief discussion, general methods for comparing prisons are quickly reviewed. The paper then moves to the methods used in this analysis, the results of the analysis, and the conclusions.

### **Taft Correctional Institution**

After modifying the intent of an executive order by then-President Clinton to privatize all new low- and most minimum-security prisons under the jurisdiction of the BOP, the U.S. Congress directed the BOP in 1997 to contract with a private vendor for the operation of a prison owned by the Federal government in Taft, California (see Conference Report to Accompany H.R. 3610, Making Omnibus Consolidated Appropriations for Fiscal Year 1997, Public Law 104-208). The Taft contract broke new ground for the BOP as the prison was built to house general population, low-security inmates. Prior to that time, the BOP had only contracted with the private sector for special populations of inmates, such as illegal aliens and pre-release prisoners. Since the Taft contract, the U.S. Congress has mandated that the BOP place a portion of minimum- and low-security inmates convicted of crimes in the District of Columbia in private prisons. As of mid-year 2001, the BOP had placed slightly over 14,000 inmates in secure private prisons. Another 6,000 inmates were in half-way (pre-release) houses operated by the private sector. The Taft facility alone held over 1,800 low-security inmates in a secure facility and another 500 in an attached prison camp.

At the same time that the Taft prison was built, the BOP constructed three other low-security prisons that used the same architectural design as the Taft facility. These prisons were

located in Elkton, Ohio, Forrest City, Arkansas, and Yazoo City, Mississippi. The construction and activation of these four prisons approximated a natural experiment, and these facilities were examined in this analysis along with several other low-security prisons operated by the BOP. Inmates are assigned to the four respective facilities in the same manner that they are assigned other BOP prisons. As such, the inmate populations were not randomly assigned, and it was necessary for us to control for the individual-level characteristics of the inmates in our statistical models as described below. Nonetheless, it was fortuitous to have three comparison prisons to compare to Taft that were built upon the same architectural design, that were activated at the same, and that held the same general types of inmates, and that generally incarcerated the same number of inmates.

In 1999, the National Institute of Justice (NIJ) contracted with Abt Associates to conduct a formal evaluation of the Taft prison. Douglas McDonald, Ph.D., a well-respected researcher in the area of private prisons, is the lead researcher on that project. An interim cost and quality assessment from the Abt team is due at any time. In addition to the research contracted for by NIJ, the BOP contracted with Julie Nelson, Ph.D., an independent contractor with expertise in economic and cost matters, to conduct cost analyses comparing Taft to BOP operations. To date, the BOP has performed well with respect to cost. Nelson (1999), using an avoidable cost methodology patterned largely on the Office of Management Budget A-76 circular (OMB, 1983),<sup>3</sup> found that the BOP could have saved \$1.9 million by operating the Taft prison in its first year of operations, fiscal year 1998. In unpublished tables provided to the BOP, Nelson (2001) reported that the BOP could have saved an additional \$2.2 million in the following fiscal year if the BOP had operated the prison. In the latest fiscal year for which complete data were available

---

<sup>3</sup>The A-76 methodology is the official mechanism used by the U.S. government in making purchase or manufacture decisions.

(fiscal year 2001), Nelson found that costs at the Taft facility and the three BOP prisons were converging. These findings by Nelson are important for the present analysis because it has been argued that there are economic incentives that push down both costs and quality at private prisons. In an interesting theoretical work, Hart, Shleifer, and Vishny (1997) argued that as a general rule differences in incentives in the private and public spheres push private prisons to operate more cheaply, but the same incentives, in the absence of very strong and precise contract language, generate lower levels of quality in operations. Such pressures should not have biased the quality comparisons examined here because at the time of this study the Taft prison spent as much or greater amounts per inmate day as the comparable BOP prisons at Elkton, Forrest City, and Yazoo City.<sup>4</sup>

### **Methods for Evaluating Prison Performance**

Proponents of private prisons claim that private prisons operate more efficiently with higher quality services (see especially Logan, 1990; Moore, 1998), and opponents argue that quality is threatened by private operations (Shichor, 1995) or that the role of government oversight is altered (Stolz, 1997). As noted above, Hart, Shleifer, and Vishny (1997) have developed an economic argument as to why there are pressures on private-prison managers to reduce costs even at the expense of quality. Despite the claims of the different parties, the cost and quality of private prison operations remain empirical questions. As such, this raises the question of identifying the best methods to use in evaluating the quality of operations at any

---

<sup>4</sup>It should be noted that the four comparison prisons differ somewhat in terms of overall size of the inmate population. Almost all of this difference comes from the size, or existence, of the attached prison camps. Taft was the only facility activated with a full prison camp. The differences in size are not of concern to this study as only inmates in the secure, low-security prisons were surveyed. Also, Taft activated about a year later than the other three prisons because of delays created by deliberations about whether to operate Taft as a private or public prison. At the time of the survey, Taft was at its full complement of inmates.



prison, whether public or private. As a point of fact, our own bias is that we do not believe that any one method alone is sufficient, but we argue for the advantages of performance measures generated from inmate survey data to supplement other sources of information. At the very least, inmates can be thought of as an interested party in the use of private prisons.<sup>5</sup> We would not go as far as the writings of constitutive criminologists and suggest that private prisons be constructed using the voices of inmates (Henry & Milovanovic, 1996), but we would agree with their perspective to the extent that inmates voices are one valid source of information.

Existing quality comparisons have primarily relied upon reviews of operational data and/or audits (Archambeault & Deis, 1996; Bowery, 1997; Tennessee Select Oversight Committee on Corrections, 1995; Thomas, 1997). There are some issues that make audits problematic for comparing prisons on a continuing basis. First, audits are expensive and somewhat subjective, although the solution used in Tennessee to use members of both the public and private sectors as audit team members was productive. Another problem with audits is the difficulty of obtaining information on a sufficient number of facilities at approximately the same time. Finally, audits disrupt normal institutional operations at the respective prisons. The best use of audits, in the Tennessee study, found that public and private prisons were comparable in terms of their operations (Tennessee Select Oversight Committee on Corrections, 1995).

The collection of operational data is certainly less disruptive than audits as it is generally part of normal institutional operations, but the use of operational data presents its own challenges. The most serious challenge to operational data is the ease with which it can be

---

<sup>5</sup>In the following discussion, we largely ignore an important assessment of prison quality in the private sector, contract monitoring. While contract monitoring is important in holding private vendors to the conditions of the contract and to minimal standards, usually established by American Correctional Association guidelines, similar monitoring is generally not present in the public sector. As such, contract monitoring is not directly useful for comparing private and public prisons.

manipulated by differences in reporting, correctional philosophy, and operational practice. The operational data that are most commonly used to compare prisons are misconduct data. The usual assumption is that differences in reported misconduct result from “real” differences in the underlying behaviors. While this assumption may be more adequate for the most serious types of misconduct in prison, such as murders or escapes from secure prisons, it is more tenuous for less serious misconduct. If a warden can convey a philosophy of informal resolution of minor disputes, and especially if the warden is fortunate enough to have a seasoned workforce, then many types of misconduct may never appear in official, operational records. As a result, the operational data are “contaminated” in the sense that they reflect differences other than differences in underlying misconduct rates. The contamination of data for less serious misconduct is most unfortunate for research purposes as the more serious misconduct is much less common, especially in lower security level prisons where private operators tend to be concentrated. Most evaluations of private prisons and comparisons with public facilities have relied upon operational data (Archambeault & Deis, 1996; Lanza-Kaduce, Parker, & Thomas, 1999; Thomas, 1997). However, serious methodological problems seriously limit the utility of using these studies for drawing general conclusions about the desirability of using private prisons (Austin & Coventry, 2001; Gaes et al., 1998; General Accounting Office, 1996)

Largely missing from discussions of prison performance are the views of inmates and staff (for limited exceptions see Camp et al., 2002; Logan, 1992; Sechrest & Shichor, 1996). In the United States, inmate surveys are generally used to generate estimates of specific issues such as the socio-demographic or criminal history characteristics of offenders (Beck, Gilliard, & Greenfeld, 1993), program/educational needs (Carlson, 1997; Lariviere & Robinson, 1999; Martin, Zimmerman, & Long, 1993), prison adjustment issues of inmates (VanVoorhis, 1994),

prison victimization (Edgar, 1998; Maitland, 1996; McCorkle, 1993; O'Donnell, 1998), or to examine contemporary issues in criminological theory (Miranne & Geerken, 1991; Tremblay & Morselli, 2000). In other countries, surveys of inmates are more likely to be part of larger efforts by prison administrators to evaluate regime needs and more general inmate concerns (Price Waterhouse, 1996; Robinson, 1996; Walmsley, Howard, & White, 1992; Wozniak & McAllister, 1992). Nonetheless, even in the foreign contexts, the goal is not always to use the data to compare specific prisons.

Survey data collected from inmates have the normal advantages associated with survey data: the data are cheap to collect, they cause little disruption during collection, and they can be administered to many institutions in a relatively short time span. Of course, survey data have the typical disadvantages, the one of greatest concern to many prison administrators is the ability of respondents to “exaggerate” if they are so inclined. To be blunt, prison administrators often fear that respondents lie to make them look bad. However, an analysis of a subset of national inmate survey data collected by the Bureau of Justice Statistics in 1997 demonstrated that inmate responses systematically fluctuated at different U.S. federal prisons, suggesting that the surveys provided something more than a uniform “gripe” forum for the inmates (Camp, 1999). Likewise, Camp and his colleagues have demonstrated that staff responses to certain survey questions at federal, U.S. prisons are influenced by differences between prisons (Camp et al., 1997; Camp et al., 1999).

Given the focus of previous studies using inmate data, analysts have rarely asked what we believe are extremely important and fundamental questions: How were inmate responses conditioned by the institutions in which the inmates were incarcerated? Were the responses affected by operational and other differences between the prisons in which the inmates were

incarcerated? Do inmates provide the same information as staff when asked about prison conditions? Typically, analysts have made the mistake of either assuming that responses are necessarily influenced by the prisons in which inmates are incarcerated (see the critique of Logan in Camp et al., 2002), or they simply never address the issue at all.

Several methodological tasks were addressed to answer the overriding methodological question of this analysis: Do inmates provide information about prison conditions that can be aggregated into institutional-level performance measures? Where possible, inmate responses were triangulated against staff responses.

Given the above objective, four steps were followed. First, we investigated the organizational properties of the data collected from inmates (and staff). As has been demonstrated elsewhere, it is not sufficient to compare prisons on the proportion of inmates and staff who provide favorable evaluations to survey questions (Camp et al., 2002). Instead, it is necessary to use multilevel modeling techniques to control for pertinent factors when constructing organizational measures with individual-level data (Bryk & Raudenbush, 1992). Second, once the appropriate measures were defined, we examined how well inmate and staff evaluations correlated across institutions. That is, if inmates rated institution X as being high on a given measure in comparison to their peers at other prisons, did staff at institution X provide a corresponding evaluation? Both bar graphs and correlations (parametric and nonparametric) were used to assess this component. Third, we examined the coherence of measures within an institution to answer whether inmates at a particular institution tended to rate all aspects of prison operations similarly. A unique graphing package, *Parallel Coordinates Display* was used to generate the graphs for this portion of the study (Howell, 2001). Finally, we used all of the

preceding information to compare the operations at one private prison with operations at selected BOP prisons.

### **Data and Methods**

The Federal Bureau of Prisons routinely administers a survey of staff, known as the Prison Social Climate Survey (PSCS). The PSCS was first administered in 1988, and it has been administered annually since that time (Saylor, 1984). In 1999, the year analyzed here, staff at 98 prisons were surveyed, and 86.9 percent of the 10,710 staff returned useable surveys (see Camp et al., 2002).

In March and April of 2000, about six months after the 1999 staff administration, an inmate version of the PSCS was given to inmates at ten low-security prisons. Inmates were free to choose between taking the survey in English or Spanish. The ten institutions were deliberately chosen to be representative of all low-security prisons in the BOP, although the four comparison prisons discussed above, Taft, Federal Correctional Institution (FCI) Elkton, FCI Forrest City, and FCI Yazoo City were deliberately included. Surveys of inmates occur on an as-needed basis at the BOP, and this survey administration was predicated on the need to obtain additional information about prison operations at the privately operated Taft Correctional Institution. A little over 100 inmates were randomly selected and surveyed at each institution, or 1,080 inmates overall. Completed surveys were obtained from 950 inmates, for a response rate of 88.0 percent.

As can be seen in the Appendix, the inmates who completed the survey appeared to be different from the inmates selected for the survey with the exception of age where the intended and actual sample data agree. Even with the discrepancies noted in Appendix 1, we are confident that there was no systematic bias in the types of respondents who completed the survey because of the control that was exercised in administering the survey. The survey was completed by

calling inmates out to an area of the prison where the surveys were administered. We are confident that the inmates on our survey lists were the inmates who showed up. Almost all inmates who were available for callout completed the survey. Most of the non-completions were the result of inmates not being available (such as inmates in administrative or disciplinary segregation). We believe that reporting differences accounted for the discrepancies between the intended and actual samples reflected in the Appendix. Hispanic inmates who classified their own race, as opposed to the official data where race was coded by staff, were likely to choose their race as other (88 inmates) or leave the race question blank (231 inmates). Clearly, many Hispanic inmates viewed their *race* as being Hispanic, where the survey forced this to be a choice for *ethnicity*.

There are two other caveats about the data worth noting. First, in the random selection of inmates into the sample, inmates with an Inmate Hearing Program (IHP) status were disproportionately sampled in five of the ten prisons.<sup>6</sup> Taft, at the time, held a large number of IHP inmates, and we wanted to collect information from other prisons to test whether IHP status influenced inmate evaluations of prison conditions. Originally, we had planned to match survey responses back to the respective individual-level operational data with a matching program to re-identify the IHP inmates. However, the socio-demographic information collected in the surveys was not rich enough to obtain satisfactory matches. Since we could not identify the IHP inmates with the matching technique, and because we did not ask the respondents directly about their IHP status, we included a variable in all of the models analyzed here for the propensity, or

---

<sup>6</sup>As used here, IHP inmate indicates an inmate who is turned over to the Immigration and Naturalization Service (INS) upon release from federal prison for deportation. IHP actually designates the Inmate Hearing Program. With the IHP program, INS conducts deportation hearings at selected prisons while the inmates are still serving their sentences. Taft is not a designated hearing site, and all IHP inmates at Taft already had their INS hearing.

predicted probability, that the respondents were IHP inmates. The coefficients for predicting the propensity scores were derived from the operational data on the intended sample, where IHP status was known, and the coefficients were then used on the survey responses to calculate the propensity score. In almost all cases, the propensity score did not significantly enter into equations of the outcome variables examined here, leading us to conclude that the sample was not biased in those institutions where IHP inmates were oversampled. The second caveat in the data is the time separation between the inmate and staff surveys. There were no major events in the BOP, such as Bureau-wide inmate disturbances, that would have had an undue, systematic impact upon altering the later-collected inmate views. Nonetheless, it would have been desirable to collect the data contemporaneously.

Missing data in the inmate and staff databases were handled with the data augmentation procedures described by Schafer (1997) and implemented in the experimental SAS Version 8.1 procedure PROC MI. Since PROC MI requires that categorical variables be measured on an ordinal scale, some variables in the inmate and staff databases were recoded into dummy variables that were appropriate. In particular, race in both the inmate and staff databases was recoded into a dummy variable comparing non-minorities (whites) and minorities. Additionally, for the staff database, an eleven-category variable measuring job category was collapsed into correctional officers, non-correctional officers. For each data source, three imputed databases were produced for the analyses that follow.

*Step 1.* The inmate and staff data were modeled with the Hierarchical Linear Model (HLM) package developed by Raudenbush, Bryk, Cheong, and Congdon (2000) as described in Bryk and Raudenbush (1992). While HLM models are estimated in one step, the models can be thought of as existing at different levels. At level-1, the outcomes observed for each individual

are modeled as a function of the covariate pattern for that individual. The following level-1 equation looks very much like an ordinary least squares (OLS) model with some important exceptions. First, error and associated variance have been separated into two components, only one of which is shown in the level-1 equation. The other component of variance, represented by  $u_{0j}$  in the level-2 equation, is associated with the clustering unit, in this case each prison. The  $u_{0j}$  term is used to measure the systematic differences between prisons and their impact upon individual-level evaluations. The error component shown at level-1,  $r_{ij}$ , is the typical error associated with each individual, stripped of the influence of the clustering unit. The  $i$  subscript references the  $i$  individuals within the  $j$  units.

$$\text{Level 1:} \quad Y_{ij} = \beta_{0j} + \sum \beta X + r_{ij}.$$

The second level of an HLM model specifies how the level-1 random coefficients are modeled as a function of the characteristics of the clustering unit. In the models examined here, the only level-1 coefficient that is treated as random is the intercept. As such, the intercept for each institution is modeled by the following level-2 equation. The estimate of  $u_{0j}$  is important for this analysis as it is the estimate of the amount by which the “average” response for the outcome variable is raised or lowered by the respondent being located at institution  $j$ .

$$\text{Level 2:} \quad \beta_{0j} = \gamma_{00} + \sum \gamma W + u_{0j}.$$

The equations presented above were used where the outcome variables were treated as quasi-continuous. While none of the outcome measures we used were truly continuous, previous experience has demonstrated that Likert items with at least seven categories work well when treated as continuous measures. Most of our items met this criteria (see Table 1). Several of the outcome measures, though, were dichotomies where respondents indicated their agreement or not with a survey item. For models of these items, a nonlinear specification was used for the



model. Conceptually, the models still existed at two levels, but level-1 of the model estimated a logit, or the log odds of agreement with the statement. Since the estimate of the variance of a probability or proportion used to compute a logit is a function of the probability/proportion, there is no independent estimate of variance available for the individual-level error term, the  $r_{ij}$ . This means that a measure of the intra-class correlation (ICC) could not be computed. ICC is defined as the amount of individual-level (level 1) variance divided by the total variance at both levels 1 and 2. The ICC was computed for the quasi-continuous measures and can be thought of in either of two ways. First, ICC can be viewed as the amount of variance that exists in the dependent measure at the level of the organization. ICC can also be thought of as the correlation between responses associated with the nesting of respondents within the respective institutions, e.g., inter-rater reliability.

The matrix of individual-level covariates,  $\mathbf{X}$ , differed in the respective analyses for inmates and staff. Different information was collected for each group. For inmates, the individual-level controls included dummy variables indicating Hispanic ethnicity, race (minorities coded 1), and whether the survey was taken in Spanish. Sex was not included as all inmates surveyed were male. The continuous variables used for inmates consisted of age, time in prison, time left to serve on sentence, time at current prison, number of federal prisons at which incarcerated, number of state prisons at which incarcerated, time spent in current housing unit, and the propensity score for being an IHP inmate. For staff, the individual-level controls included dummy variables for sex (females coded 1), Hispanic ethnicity, race (minorities coded 1), supervisory status (supervisors coded 1), college degree, being a correctional officer, working on the day shift, and not working at the main facility (such as at a prison camp). The only continuously measured variables for staff were age and years of BOP tenure. For parsimony, the

individual-level controls were treated as nuisance variables impacting upon the  $u_{0j}$  estimates of institutional differences and are not presented for discussion.

The matrix of organizational-level covariates,  $\mathbf{W}$ , also differed for inmates and staff. Because information on inmates was only collected at ten prisons, there were not sufficient degrees of freedom to model the random  $\beta_{0j}$  coefficients in the inmate models. As such, the level-2 model for inmates was simply  $\beta_{0j} = \gamma_{00} + u_{0j}$ . For staff evaluations, surveys were collected at 98 institutions, and the  $\mathbf{W}$  matrix included controls for the security-level of the prison as well as region of the BOP. Region of the BOP corresponds roughly to U.S. geographical regions. With the exception of the effect of region, though, the inmate and staff models did not differ significantly because all of the inmate prisons surveyed were low-security prisons.

Because of the inability to model the random  $\beta_{0j}$  for inmate responses, partial results for staff data are presented where no level-2 covariates were included. Further, the limited models for staff data were restricted to the ten institutions for which inmate data were collected. This allows for at least an intuition about how inmates results would look if the surveys were distributed to more institutions.

*Step 2.* Once measures were identified with desirable organizational properties, those measures were examined to see if there was correspondence between the organizational measures developed from inmate and staff data, the  $u_{0j}$  terms from the HLM models. The  $u_{0j}$  terms that were generated from staff data for the ten prisons examined in this step were taken from models of all 98 prisons. First, correlations were computed for the respective  $u_{0j}$  terms. For example, there were two measures of problems of sanitation in the dining hall (DIRTDINE) for each prison, one calculated from inmate data and one from staff data. The correlations assessed

the congruence between measures developed from the different sources. The actual metric of the  $u_{0j}$  terms were compared with a Pearson correlation coefficient, and the rankings of the institutions produced by the  $u_{0j}$  terms were assessed with the nonparametric Spearman's rho coefficient.

*Step 3.* William Saylor and Schoeneck Howell of the Office of Research and Evaluation have worked over the past decade to develop better means of presenting information on measures of prison performance. They have consistently tried to simplify the presentation of information so that more data can be grasped at one time by BOP managers. One graphical method they have developed, built upon the work of Hartigan (1975), Inselburg (1990), and Wegman (1990), for such a display of information is the technique called parallel coordinates display. The idea is rather straightforward. For a given number of institutions defined by the user, create horizontal axes for each measure selected with the value zero at the midpoint of the axes. Then, plot the values recorded for each prison on the axes and join the values for each prison on adjacent axes with a vertical line. A score above zero on the plot indicates positive performance, and a score below zero negative performance. That is, a positive value for an institution indicates performance about that expected on the measure in question, and a negative value suggests poorer performance than expected. Thus, the parallel coordinates displays allow interpretations to be made about how an individual institution's performance is better or worse than expected and how that institution compares to other institutions on similar performance indicators. The result is a plot that easily demonstrates whether the prisons as a group or separately are ranked consistently by the measures. As mentioned, the plots demonstrate whether the prisons are doing better or worse than other prisons of interest (indicated by being further to the right or left, respectively, on the axes). If the lines on the plot are generally parallel to one

another, it is an indication that the prisons are consistently ranked on the measures in question. On the other hand, if the lines crisscross one another, this is an indication of inconsistency. The software developed to realize this notion, *Parallel Coordinates Display*, was used to examine the  $u_{0j}$  values produced by the HLM runs in step 1 (Howell, 2001).

*Step 4.* The information gleaned from the previous steps was used to compare the operations at one private prison, Taft Correctional Institution (TCI), with the operations of the 9 BOP prisons at which inmate surveys were gathered. In particular, though, TCI was compared with three other BOP prisons, FCI Elkton, FCI Forrest City, and FCI Yazoo City. As noted previously, the comparison prisons were built upon identical architectural footprints and activated at about the same time as Taft Correctional Institution. Primarily, these comparisons were generated from information produced in steps 2 and 3 above.

## **Results**

The results of the analyses of the organizational properties of the different outcome variables, step 1 of the analysis, are presented in Table 2. For all measures, the reliability of the ranking of the institutions that was based on the deviations from the intercept or the typical response, the  $u_{0j}$ , is presented. While there is no technical cut-off value, most analysts consider as acceptable a measure with a reliability of at least .70. For those outcome variables that were measured with more than two categories (quasi-continuous variables), it was also possible to calculate the ICC. Obviously, the most desirable measures for making institution comparisons are those with high values for ICC and reliability.

### *Step 1. Identify Organizational Properties of Measures*

The results presented in Table 2 strongly suggest that there is merit to using inmates to evaluate conditions at prisons. A striking finding is that if we had relied only upon staff working

at the ten prisons where inmate surveys were given, then we would have not been able to conclude that any of the measures had desirable organizational properties. None of the reliability scores for measures generated with staff data from the limited number of prisons exceeded 0.7. In fact, almost all of the reliability scores were well below that cutoff. We also encountered difficulty in generating estimates for several of the nonlinear models for these data. At least partially due to the low reliability in measuring the random component associated with the intercept, the  $u_{0j}$ , many of the nonlinear models for the limited staff data failed to converge in 3,000 iterations for at least one of the multiply imputed databases. Since the results from the other databases clearly suggested low levels of reliability for the measures, and because similarly low levels of reliability were noted for the measures when using staff data from all institutions, the problematic models were not “fixed” in any sense, and the results were reported in Table 2 for those models where there was convergence to a solution. A fix could have been attempted by producing another multiple imputation and re-estimating the models. Quite frankly, this was not deemed to be necessary as the results of the models for the databases where convergence did occur clearly showed that meaningful results were unlikely.

Likewise, the reliability values for the measures computed from staff data for all 98 prisons were generally low. There were only two measures, the measure of sanitation in the dining halls (DIRTDINE) and the measure of sanitation in the housing unit (DIRTUNIT), that had sufficient levels of reliability to be used as organizational-level measure. Even for these two measures, the reliability values for the measures derived from inmate data were higher, as were the ICC values. Nonetheless, since these two measures could be reliably used when generated from either inmate or staff data, they were used in the next step of the analysis (presented below)

to assess whether there was general congruence between inmate and staff evaluation of sanitation.

The reliability values for measures computed from inmate survey data were typically acceptable. Of the twenty-four outcome measures examined here, a reliability of at least .7 was noted for sixteen of the measures. In fact, most of these measures had reliability values in excess of .8, with some going as high as .97. Given the fact that there were only ten institutions, this finding is remarkable, especially in comparison to the general failure of any of the measures computed with staff data to exceed a reliability of .7 for these same ten prisons. All of the inmate measures that did not meet the general cutoff for reliability were items that asked about security and safety conditions at the prison. With perfect hindsight, it would be easy to speculate about why these measures were not reliable, but we resist the temptation. Instead, we note that there were five security and safety condition questions that were reliable.

### *Step 2. Congruence between Inmate and Staff Evaluations*

The results in Table 3 addressed the next issue of this analysis, whether inmates and staff provided congruent responses when evaluating prison conditions. The obvious caveat to this portion of the analysis was that we only found two questions with which to compare measures generated from inmate and staff data. Nonetheless, the results presented in Table 3 suggest that inmates and staff were highly congruent in their evaluations. From Table 3, we see that the correlations, both the parametric, Pearson and the non-parametric, Spearman correlation coefficients suggested a close match between inmate and staff evaluations. For the sanitation in the dining hall question, the respective coefficients were .741 and .661. Both coefficients were statistically significant even with scores on only ten prisons. For the item about sanitation in the

housing unit, the correlations were even stronger, .900 and .855. Again, both correlations were statistically significant.

### *Step 3. Assessing Congruence of Subject Area Measures*

The evaluation of whether the different measures were congruent for the topical areas noted in Tables 1 and 2 was accomplished with Figures 1 through 4. As noted above, the graphs presented information on the systematic influence of the institution on evaluations of the different measures, the  $u_{0j}$  terms in the HLM models. Generally speaking, the different measures produced consistent information (e.g., the parallel lines) for the individual prisons within each substantive area. Figure 1 presents the results for the three measures of gang activity. The plot demonstrates that prisons high on one measure tended to be high on the others and vice versa.

One thing to keep in mind when looking at the plots is that the dichotomous and quasi-continuous variables (variables with 7 choices on a Likert scale) are both plotted on the same graph. Dichotomous outcomes are easy to distinguish as the measures always range between -0.5 and 0.5. The values were calculated by taking the predicted value for an institution and subtracting the median for the measure. This centered each measure around zero, and those institutions with poorer performance for the measure had negative values, and those institutions with more positive performance had positive values for the measure. The quasi-continuous measures also took on both positive and negative values and were centered around zero. In the case of these measures, the value is the amount above or below the “average” value for all institutions that a particular institution fell. Again, positive performance was indicated by values greater than zero, and poor performance was noted by negative scores.

Figure 2 shows the same general consistency for the security and safety measures. Figure 3 seems to show the least consistency of the subject area measures examined so far. But this first perception is somewhat misleading. There were two factors underlying the measures plotted in Figure 3. The first two measures were clearly sanitation measures, and the institutions held their



respective rankings well for these two measures. The other factor was noise at the institution. If we look at the lines connecting the “quite during evening hours” and “quite during sleeping hours” measures, it can be seen that these measures also were fairly congruent. The mismatch occurred because there was no necessary correspondence between sanitation and noise at all respective prisons.

Figure 4 shows that there was stability in the rankings of food service. Not only was there stability, there was greater separation (distance) between highly and lowly ranked prisons for these measures than for the other topical areas. The clear separation of these institutions was also captured in the reliability scores. The reliability values for the group of food service measures were the highest noted in this analysis.

#### *Step 4. Comparing Taft and the BOP Prisons*

The graphs produced for this analysis make it easy to compare the different prisons. In Figures 1 through 4, Taft and the three BOP comparison prisons were highlighted.<sup>7</sup> In Figure 1, it is easy to identify that all of the comparison prisons and Taft (TAF) fell in the middle of the pack when ranking gang-related aspects of prison operations. Taft was generally ranked higher than Forrest City (FOR), Elkton (ELK), and Yazoo (YAZ). The only exception to this pattern was that Elkton rated slightly higher than Taft on the safety for gang members measure.

Figure 2 shows that Taft ranked near the middle of all of the prisons with respect to safety and security concerns. The three BOP comparison prisons, on the other hand, generally had the lowest ratings for these measures. Inmates at these facilities rated their institutions as

---

<sup>7</sup>Since the other BOP prisons were not part of the formal comparison and served as context for comparing Taft to the three BOP comparison prisons, there was no perceived benefit to identifying the prisons explicitly. In fact, we would have preferred not to identify any of the prisons, but this was not feasible. It is widely known that Taft is being operated as a demonstration project for the BOP. As such, we decided to identify the BOP institutions as well in the interest of parity.

lower than the other prisons both in terms of their perceived safety from being hit or assaulted, and these same comparisons facilities tended to have the lowest ratings for having sufficient numbers of staff available at the institution to protect inmates.

Figure 3 demonstrates that Taft experienced some problems with sanitation. Taft was lower than all but one BOP prison on the measures of sanitation in the dining hall. For housing unit sanitation, Taft fell in the bottom half of the distribution as only three BOP facilities rated lower. The three BOP comparison prisons fell in the middle of the distribution for these same measures. For noise at the institution, on the other hand, Taft ranked very well with the highest score for quiet during the evening hours and the next to highest rating for quiet during sleeping hours. The BOP prisons produced mixed results, but noise was clearly more of an issue at these prisons. In fact, Elkton had the lowest score for quiet during sleeping hours. Yazoo was not much higher than Elkton, and Forrest City fell in the middle of the prisons for these measures.

Figure 4 demonstrates that Taft clearly had problems with food services. The rankings produced with the inmate data demonstrate that Taft rated less favorably than the BOP prisons as they had the lowest ranking for three of the four items, namely the quality of food, the variety of food, and amount of food being served. For the fourth item, the appearance of the food, Taft had the second lowest score, again indicating a problem. The BOP comparison prisons scored better than Taft, but they tended to be on the lower side of the scores or in the middle of the rankings of the prisons.

## **Discussion and Conclusion**

Given the results presented here, measures derived from inmate data deserve better treatment than the distrust they often receive from prison administrators. Inmate data can be used to generate reliable measures that differentiate performance at prisons. Not only were the inmate

measures much more reliable than measures developed from staff data, the inmate-derived measures correlated well with staff evaluations for the two measures of sanitation where comparisons were possible. The correspondence between inmate and staff evaluations on the sanitation items demonstrated two facts. First, where items of common interest to staff and inmates are used, inmates and staff appear to be affected by similar situational factors in providing evaluations of institutional operations. Second, it is important to identify and develop survey items that are clearly relevant to and appropriate for the intended sample.

The parallel coordinates graphs demonstrated the ease by which multiple measures generated with sophisticated statistical techniques could be examined at once. Even though the models used to construct the graphs were beyond the understanding of the typical correctional administrator, the results of the models are easily explained to nontechnical audiences. By dividing the 24 measures into subject areas, and limiting the graphs to subject areas, we were able to quickly examine how Taft compared to the other nine BOP prisons, but especially how Taft compared to the three comparison institutions that were very similar with respect to physical plant.

The findings suggested that Taft had problems in maintaining quality in some areas, and the BOP comparison prisons had problems as well. As noted previously, the comparison institutions were very similar to Taft with respect to size of the facility, architectural design, security-level, activation date, and even costs. In several areas, Taft performed at an average level in comparison to the BOP prisons, most notably on the gang measures and the safety and security concern measures. In the other areas examined here, Taft had problematic findings. There were problems noted for the measures of sanitation in both the dining hall and housing units, and there were more serious problems noted for the food service measures where Taft had

the most problematic rankings for three of the four measures. For the fourth measure of food services, Taft had the next to worst rating.

We would have preferred a simple solution to the question of whether private prisons produce better quality on the limited number of measures we examined, but our results suggested that the world is more complex. We would argue that private prisons need the same level of monitoring as public prisons to insure adequate performance.

The BOP comparison prisons also had some problematic findings that may have been related to the competition between them and Taft. In particular, the finding that the institutions generally rated worst in terms of providing an environment in which inmates felt safe from being hit or assaulted was noteworthy, especially since these institutions also rated lowest in terms of inmates believing that there were enough staff during the different shifts to provide for their safety. This finding was interesting because Nelson found that the BOP comparison facilities operated with fewer staff than Taft during this time period (Nelson, 1999). The BOP comparison prisons also tended to have bad ratings for the food service measures, although clearly Taft was more problematic in this area. The findings about food at Taft and the three BOP comparison prisons were significant as the money spent on food is more easily manipulated than other parts of a prison budget. The overall portion of a prison budget spent on food services, though, is small. In estimates of BOP spending at comparison prisons in 1999, Nelson found that the BOP spent \$2.63 on food for each inmate per day out of a total per diem cost of \$37.14.

Taft and the BOP comparison prisons did differ from one another on the information presented here. Without complete consensus among all measures used, we were not able to conclude that on the whole Taft did better or worse than the BOP comparison prisons. Instead, the most general conclusion is that Taft did worse in some areas but not others. We would never

argue for sole reliance upon survey data to measure prison performance or prison quality, whether collected from inmates or staff. But given the relative ease and convenience of data collection, the low expense associated with data collection, and the favorable measurement properties associated with inmate survey data, we do argue for using performance measures generated from inmate survey data as part of a more comprehensive strategy for comparing prisons.

**Table 1. Questions Analyzed**

Variable Name	Wording of Question	Staff Version	Differences	Cate-gories*
<i>Gang-Related Issues</i>				
GANGACTV	Has there been any gang activity in this prison during the past 6 months?	Yes	No Knowledge Choice for Staff	2/3
GANGSAFE	How safe or dangerous do you think it has been in this prison for inmates who <i>are</i> members of a gang?	Yes	None	6
NGANGSAF	How safe or dangerous do you think it has been in this prison for inmates who <i>are not</i> members of a gang?	Yes	None	6
<i>Security/Safety Concerns</i>				
ASS_UNIT	How likely is it that an inmate would be assaulted in his/her living unit?	Yes	Slight Wording	4
SAFE_HIT	How safe do you feel from being hit, punched, or assaulted by other inmates?	No		5
SAFE_PRP	How safe do you feel your property has been?	No		5
YUASSLTD	Have you been physically assaulted by an <i>inmate</i> within the last 6 months?	No		2
STOPSEX	Staff have prevented forced sex among inmates.	No		6
ENUFSTFD	Do you think there have been enough staff here to provide for the safety and security of <i>inmates</i> : During the <i>day</i> (8 am - 4 pm) shift?	Yes	No Knowledge Choice for Staff	2/3
ENUFSTFN	Do you think there have been enough staff here to provide for the safety and security of <i>inmates</i> : During the <i>night</i> (4 pm - midnight) shift?	Yes	No Knowledge Choice for Staff	2/3
ENUFSTFM	Do you think there have been enough staff here to provide for the safety and security of <i>inmates</i> : During the <i>morning</i> (midnight - 8:00 am) shift?	Yes	No Knowledge Choice for Staff	2/3
SAFSTFMA	How safe do you think it is for <i>male</i> staff members who have frequent contact with inmates in this prison?	Yes	Slight Wording	6
SAFSTFFE	How safe do you think it is for <i>female</i> staff members who have frequent contact with inmates in this prison?	Yes	Slight Wording	6

**Table 1. Questions Analyzed, Continued**

Variable Name	Wording of Question	Staff Version	Differences	Categories*
SEARCHES	How often have there been shakedowns in this institution during the past 6 months?	Yes	Slight Wording, No Knowledge Choice for Staff	7/8
STRPSRCH	How often have you been strip or pat searched at this institution during the past 6 months ( <i>not</i> including those required for visits)?	No		7
SEARFREQ	Are the shakedowns done frequently enough?	Yes	None	2
<i>Overall Sanitation/Conditions</i>				
DIRTUNIT	How often have insects, rodents, dirt, or litter been a problem in the housing units?	Yes	No Knowledge Choice for Staff	7/8
DIRTDINE	How often have insects, rodents, dirt, or litter been a problem in the dining hall?	Yes	No Knowledge Choice for Staff	7/8
EVENOISE	How noisy has it been in your housing unit during the evening hours?	Yes	Wording, question and choices	5
SLPNOISE	How noisy has it been in your housing unit during sleeping hours?	Yes	Wording, question and choices	5
<i>Food Service Delivery</i>				
QUALFOOD	The quality of food at this prison has been: (poor, fair, good)	No		3
VARIFOOD	The variety of food at this prison has been: (poor, fair, good)	No		3
AMTFOOD	The amount of food served for main courses has been (Not enough, Enough)	No		2
APPFOOD	The appearance of the food at this prison has been: (Unappealing, Appealing)	No		2

\* For items so marked, inmates had only two answer choices, yes and no. Staff had a third choice, as noted in the column "Differences." Staff could mark that they possessed no knowledge about the question of interest.

**Table 2. Results for Inmate and Staff Respondents for the Prison Social Climate Survey (PSCS) in 2000**

Variable Name	Inmate PSCS <sup>1</sup>		Staff PSCS (Limited) <sup>2</sup>		Staff PSCS (All, Full) <sup>3</sup>	
	Reliabilit	ICC <sup>4</sup>	Reliabilit	ICC	Reliabilit	ICC
<i>Gang-Related Issues</i>						
GANGACTV (Gang Activity)	.939	ND <sup>5</sup>	.614 <sup>6</sup>	ND	.583	ND
GANGSAFE (Gang members safe)	.936	.134	.018	.000	.299	.005
NGANGSAF (Non-gang safe)	.895	.084	.052	.001	.284	.004
<i>Security/Safety Concerns</i>						
ASS_UNIT (Assault likely, housing)	.848	.056	.125	.002	.411	.008
SAFE_HIT (Safe from being hit)	.706	.025	NA <sup>8</sup>	NA	NA	NA
SAFE_PRP (Property safe)	.409	.008	NA	NA	NA	NA
YUASSLTD (Were you assaulted)	.443	ND	NA	NA	NA	NA
STOPSEX (Prevent forced sex)	.672	.023	NA	NA	NA	NA
ENUFSTFD (Enough staff, day)	.764	ND	.363 <sup>7</sup>	ND	.157	ND
ENUFSTFN (Enough staff, night)	.721	ND	.116 <sup>7</sup>	ND	.437	ND
ENUFSTFM (Enough staff, morning)	.812	ND	.357 <sup>6</sup>	ND	.344	ND
SAFSTFMA (Male staff safe)	.676	.022	.011	.000	.367	.006
SAFSTFFE (Female staff safe)	.618	.017	.015	.000	.361	.006
SEARCHES (Periodicity of searches)	.662	.021	.026	.000	.233	.003
STRPSRCH (Were you strip-searched)	.596	.016	NA	NA	NA	NA
SEARFREQ (Searches frequent)	.089 <sup>5</sup>	ND	.259	ND	.270	ND
<i>Overall Sanitation/Conditions</i>						
DIRTUNIT (Sanitation, housing unit)	.919	.107	.593	.017	.802	.042
DIRTDINE (Sanitation, dining hall)	.930	.123	.314	.006	.803	.042
EVENOISE (Evening noise)	.761	.033	.540	.013	.648	.020
SLPNOISE (Sleep noise)	.852	.057	.659	.024	.567	.014
<i>Food Service Delivery</i>						
QUALFOOD (Quality of food)	.978	.316	NA	NA	NA	NA
VARIFOOD (Variety of food)	.971	.262	NA	NA	NA	NA
AMTFOOD (Amount of food)	.952	ND	NA	NA	NA	NA
APPFOOD (Appearance of food)	.831	ND	NA	NA	NA	NA

## TABLE 2 NOTES:

1. N=950
2. N=916
3. N=9,491
4. Intra-Class Correlation (ICC)
5. Not defined. ICC cannot be computed for binomial models where there is no estimate of individual error,  $r_{ij}$ .
6. Estimate based on results for only one of the three multiply imputed databases. After 3,000 iterations, HLM was unable to derive a solution for the other two databases. This result should be treated with caution.
7. Estimate based on results for two of the three multiply imputed databases. After 3,000 iterations, HLM was unable to derive a solution for the third database.
8. Not available. The question was not asked of staff respondents.



**Table 3. Correlations between Inmate- and Staff- Based Evaluation of Prison Operations  
Where both Inmate and Staff Data Produce Reliable Measures**

Variable Name	Pearson (metric of $u_{ij}$ )		Spearman rho (rankings of $u_{ij}$ )	
DIRTINE (Sanitation, dining hall)	.741	p=.014	.661	p=.038
DIRTUNIT (Sanitation, housing unit)	.900	p=.000	.855	p=.002

**Appendix. Representativeness of Survey Respondents**

Socio-Demographic Characteristic	Intended Sample	Actual Sample	Difference
Race			
Minority	348 (32.2%)	395 (56.9%)	-24.7
Non-minority	732 (67.8%)	299 (43.1%)	24.7
Missing		256	
Ethnicity			
Hispanic	452 (41.9%)	455 (52.5%)	-10.6
Non-Hispanic	628 (58.1%)	411 (47.5%)	10.6
Missing		84	
Age			
Mean	36.5 years	36.8 years	-0.3
Median	35 years	35 years	0
Missing		43	
Years at Current Prison			
Mean	0.7 years	2.2 years	1.5
Median	0.8 years	2.2 years	
Missing		403	

Overall Response Rate: 87.96 percent

## REFERENCES

- Archambeault, W. G., & Deis, D. R. (1996). *Cost effectiveness comparisons of private versus public prisons in Louisiana: A comprehensive analysis of Allen, Avoyelles, and Winn Correctional Centers*. Baton Rouge: Louisiana State University.
- Austin, J., & Coventry, G. (2001). *Emerging issues on privatized prisons*. Washington, D.C.: Bureau of Justice Assistance.
- Beck, A., Gilliard, D., & Greenfeld, L. (1993). *Survey of state prison inmates, 1991*. Washington, D.C.: Bureau of Justice Statistics.
- Bowery, M. (1997). *Private prisons in NSW: June - year three (38)*. Sydney, Australia: NSW Department of Corrective Services.
- Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods*. Newbury Park, CA: Sage.
- Camp, S. D. (1999). Do inmate survey data reflect prison conditions? Using surveys to assess prison conditions of confinement. *The Prison Journal*, 79(2), 250-268.
- Camp, S. D., & Gaes, G. G. (2000a). Private adult prisons: What do we really know and why don't we know more? In D. Shichor & M. Gilbert (Eds.), *Privatization of criminal justice: Past, present and future* (pp. 283-298). Cincinnati, OH: Anderson.
- Camp, S. D., & Gaes, G. G. (2000b). *Private prisons in the United States, 1999: An assessment of growth, performance, standards, and training requirements*. Washington, D.C.: Federal Bureau of Prisons.
- Camp, S. D., & Gaes, G. G. (2002). Growth and quality of U.S. private prisons: Evidence from a national survey. *Criminology & Public Policy*, forthcoming.

- Camp, S. D., Gaes, G. G., & Saylor, W. G. (2002). Quality of prison operations in the U.S. federal sector: A comparison with a private prison. *Punishment & Society*, 4(1), 27-53.
- Camp, S. D., Saylor, W. G., & Harer, M. D. (1997). Aggregating individual-level evaluations of the organizational social climate: A multilevel investigation of the work environment at the Federal Bureau of Prisons. *Justice Quarterly*, 14(4), 739-761.
- Camp, S. D., Saylor, W. G., & Wright, K. N. (1999). Creating performance measures from survey data: A practical discussion. *Corrections Management Quarterly*, 3(1), 71-80.
- Carlson, J. R. (1997). Differential perceptions of women inmates, correctional staff, and correctional officers regarding rehabilitation program modalities. *Journal of Offender Rehabilitation*, 25(3/4), 159-170.
- Edgar, I. O. D. a. K. (1998). Routine victimization in prisons. *The Howard Journal*, 37(3), 266-279.
- Gaes, G. G., Camp, S. D., & Saylor, W. G. (1998). Appendix 2: Comparing the quality of publicly and privately operated prisons: A Review. In D. McDonald & E. Fournier & M. Russell-Einhorn & S. Crawford (Eds.), *Private prisons in the United States: An assessment of current practice* (pp. 1-38). Boston, MA: Abt Associates Inc.
- General Accounting Office. (1996). *Private and public prisons: Studies comparing operational costs and/or quality of service*. Washington, D.C.: U.S. General Accounting Office.
- Harding, R. (2001). Private prisons. In M. Tonry (Ed.), *Crime and justice: A review of research* (Vol. 28, pp. 265-346). Chicago: University of Chicago Press.
- Hart, O., Shleifer, A., & Vishny, R. W. (1997). The proper scope of government: Theory and an application to prisons. *The Quarterly Journal of Economics*, 112(4), 1127-1161.
- Hartigan, J. A. (1975). *Clustering algorithms*. New York: John Wiley & Sons.

- Henry, S., & Milovanovic, D. (1996). *Constitutive criminology: Beyond postmodernism*. London: Sage.
- Howell, S. (2001). Parallel coordinates display (Version 2.0). Washington, D.C.: Federal Bureau of Prisons.
- Inselburg, A., & Dimsdale, B. (1990). *Parallel coordinates: A tool for visualizing multi-dimensional geometry*. Paper presented at the First IEEE Conference on Visualization.
- Lanza-Kaduce, L., Parker, K. F., & Thomas, C. W. (1999). A comparative recidivism analysis of releasees from private and public prisons. *Crime and Delinquency*, 45(1), 28-47.
- Lariviere, M., & Robinson, D. (1999). The prevalence and determinants of HIV risk behaviors and perceived threat of infection in Canadian federal penitentiaries: Results from a national survey of male inmates. *Journal of Correctional Health Care*, 6(1), 7-26.
- Logan, C. (1990). *Private prisons: Cons and pros*. New York: Oxford University Press.
- Logan, C. H. (1992). Well kept: Comparing quality of confinement in private and public prisons. *The Journal of Criminal Law and Criminology*, 83(3), 577-613.
- Maitland, A. S. a. S., R.D. (1996). Victimization in prisons: A study of factors related to the general well-being of youthful offenders. *Federal Probation*, 60(2), 24-31.
- Martin, R., Zimmerman, S., & Long, B. (1993). AIDS education in U.S. prisons: A survey of inmate programs. *The Prison Journal*, 73(1), 103-129.
- McCorkle, R. C. (1993). Fear of victimization and symptoms of psychopathology among prison inmates. *Journal of Offender Rehabilitation*, 19(1/2), 27-41.
- McDonald, D. C., Fournier, E., Russell-Einhorn, M., & Crawford, S. (1998). *Private prisons in the United States: An assessment of current practice*. Boston, MA: Abt Associates Inc.
- Miranne, A. C., & Geerken, M. R. (1991). The New Orleans inmate survey: A test of Greenwood's predictive scale. *Criminology*, 29(3), 497-518.

- Moore, A. T. (1998). *Private prisons: Quality corrections at a lower cost*. Los Angeles, CA: Reason Public Policy Institute.
- Nelson, J. (1998). Comparing public and private prison costs. In D. McDonald & E. Fournier & M. Russell-Einhorn & S. Crawford (Eds.), *Private prisons in the United States: An assessment of current practice*. Boston: Abt Associates, Inc.
- Nelson, J. (1999). *Taft prison facility: Cost scenarios*. Washington, D.C.: National Institute of Corrections.
- Nelson, J. (2001). *Cost comparison of Taft with BOP low-security prisons*. Unpublished manuscript, Washington, D.C.
- O'Donnell, K. E. a. I. (1998). Assault in prison: The victim's contribution. *British Journal of Criminology*, 38(4 Autumn), 635-650.
- OMB. (1983). *Circular A-76: Performance of commercial activities*. Washington, D.C.: Office of Management and Budget.
- Price Waterhouse. (1996). *1995 National inmate survey: Final report*. Ottawa: Price Waterhouse.
- Raudenbush, S., Bryk, A., Cheong, Y. F., & Congdon, R. (2000). HLM 5: hierarchical linear and nonlinear modeling (Version 5.03). Lincolnwood, IL: Scientific Software International, Inc.
- Robinson, M. L. a. D. (1996). *Attitudes of correctional officers towards offenders* (Executive Summary). Ontario: Price Waterhouse.
- Saylor, W. G. (1984). *Surveying prison environments*. Washington, D.C.: Federal Bureau of Prisons.
- Schafer, J. L. (1997). *Analysis of incomplete multivariate data*. New York: Chapman & Hall.

- Sechrest, D. K., & Shichor, D. (1996). Comparing public and private correctional facilities in California: An exploratory study. In G. L. Mays & T. Gray (Eds.), *Privatization and the provision of correctional services* (pp. 133-151). Cincinnati, OH: Anderson Publishing.
- Segal, G. F., & Moore, A. T. (2002). *Weighing the watchmen: Evaluating the costs and benefits of outsourcing correctional services*. Los Angeles: Reason Public Policy Institute.
- Shichor, D. (1995). *Punishment for profit: Private prisons/public concerns*. Thousand Oaks, CA: Sage.
- Stolz, B. A. (1997). Privatizing corrections: Changing the corrections policy-making subgovernment. *The Prison Journal*, 77(1), 92-111.
- Tennessee Select Oversight Committee on Corrections. (1995). *Comparative evaluation of privately-managed Corrections Corporation of America prison (South Central Correctional Center) and state-managed prototypical prisons (Northeast Correctional Center, Northwest Correctional Center)*. Nashville, TN: Tennessee Select Oversight Committee on Corrections.
- Thomas, C. W. (1997). *Comparing the cost and performance of public and private prisons in Arizona*. Gainesville, FL: Center for Studies in Criminology & Law, University of Florida.
- Tremblay, P., & Morselli, C. (2000). Patterns in criminal achievement: Wilson and Abrahams revisited. *Criminology*, 38(2), 633-659.
- VanVoorhis, P. (1994). Measuring prison disciplinary problems: A multiple indicators approach to understanding prison adjustment. *Justice Quarterly*, 11(4), 679-709.
- Walmsley, R., Howard, L., & White, S. (1992). *The national prison survey 1991: Main findings* (Home Office Research Study No. 128). London: Her Majesty's Stationery Office.
- Wegman, E. J. (1990). Hyperdimensional data analysis using parallel coordinates. *Journal of the American Statistical Association*, 85(411), 664-675.

Wozniak, E., & McAllister, D. (1992). *The prison survey*. Edinburgh, Scotland: Central Research Unit, Scottish Prison Service.



Figure 1. Gang-Related Activity

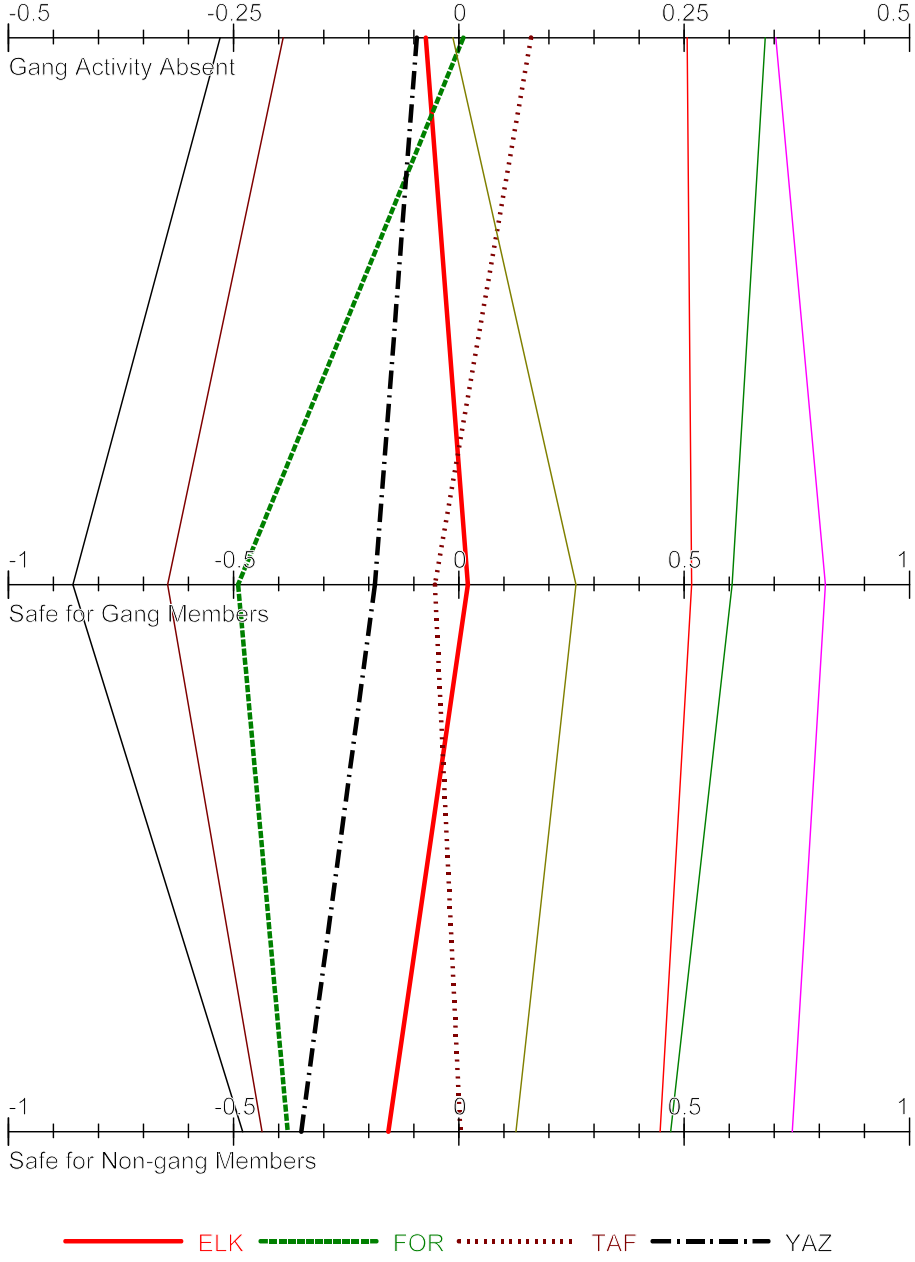


Figure 2. Security/Safety Concerns

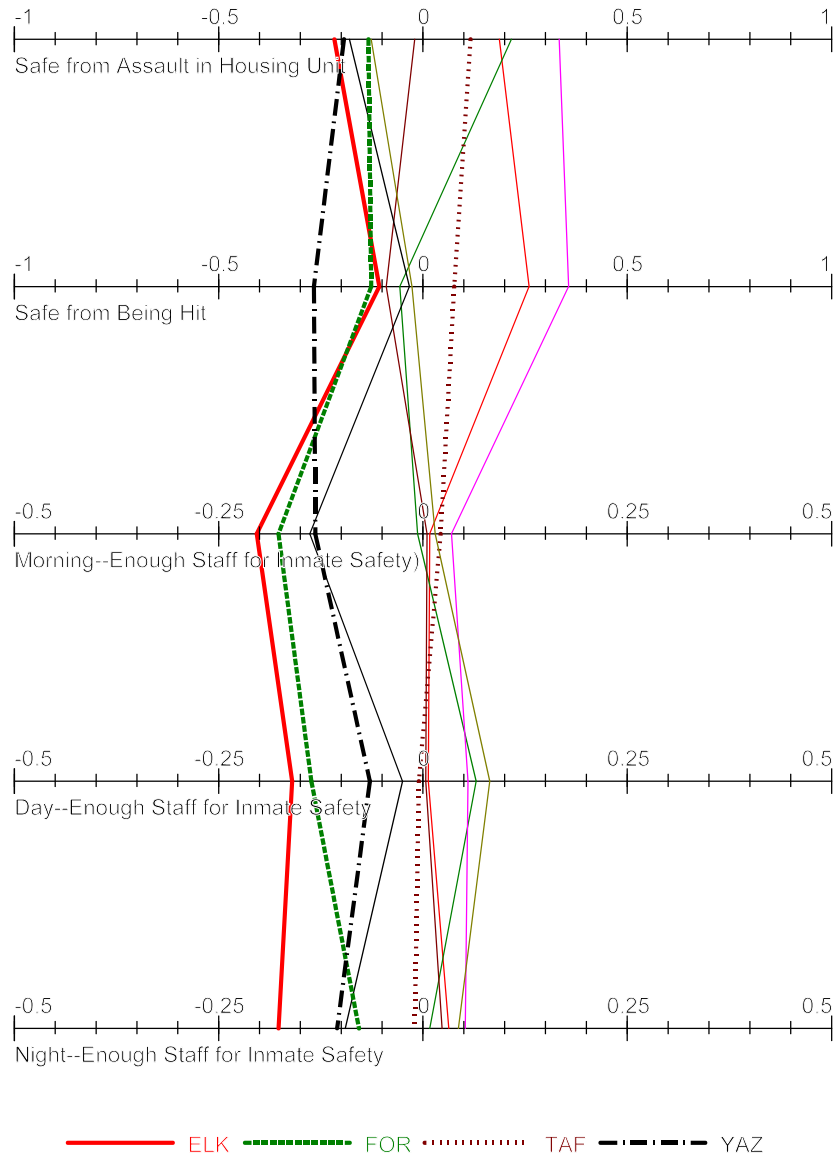


Figure 3. Overall Sanitation/Conditions

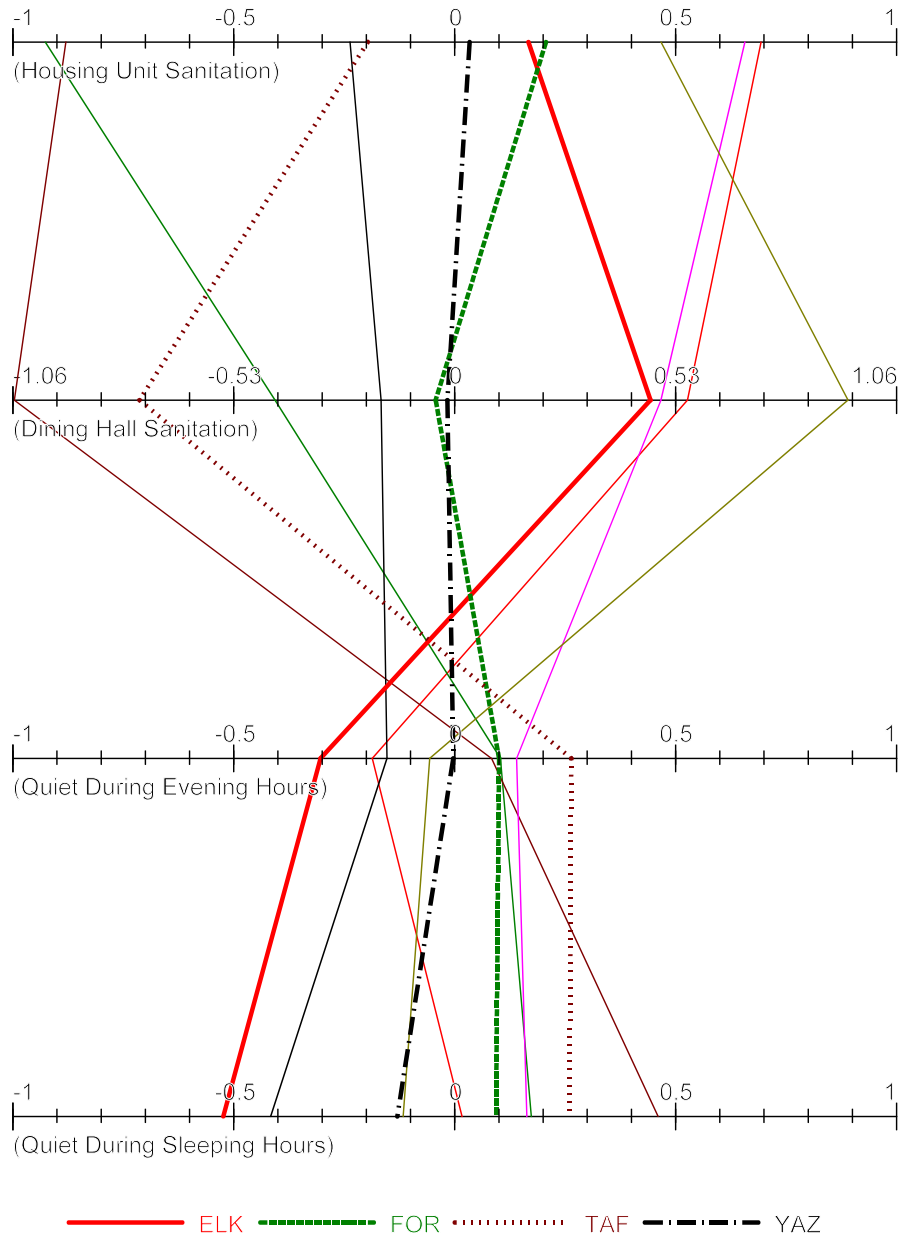


Figure 4. Food Service Delivery

