J Exp Criminol (2009) 5:139–162
DOI 10.1007/s11292-009-9070-z

Unintended consequences: experimental evidence for the criminogenic effect of prison security level placement on post-release recidivism

Gerald G. Gaes · Scott D. Camp

Published online: 18 February 2009

Abstract Most prison systems use quantitative instruments to classify and assign inmates to prison security levels commensurate to their level of risk. Bench and Allen (The Prison Journal 83(4):367-382, 2003) offer evidence that the assignment to higher security prisons produces elevated levels of misconduct independent of the individual's propensity to commit misconduct. Chen and Shapiro (American Law and Economics Review, 2007) demonstrate that assignment to higher security level among inmates with the same classification scores increases post-release recidivism. Underlying both of these claims is the idea that the prison social environment is criminogenic. In this paper we examine the theoretical premises for this claim and present data from the only experiment that has been conducted that randomly assigns inmates to prison security levels and evaluates both prison misconduct and post-release recidivism. The experiment's results show that inmates with a level III security classification who were randomly assigned to a security level III prison in the California prison system had a hazard rate of returning to prison that was 31% higher than that of their randomly selected counterparts who were assigned to a level I prison. Thus, the offenders' classification assignments at admission determined their likelihood of returning to prison. There were no differences in the institutional serious misconduct rates of these same prisoners. These results are contradictory to a specific deterrence prediction and more consistent with peer influence and environmental strain theories. These results also raise important policy implications that challenge the way correctional administrators will have to think about the costs and benefits of separating inmates into homogeneous pools based on classification scores.

G. G. Gaes (✉)
College of Criminology & Criminal Justice, Florida State University, Tallahassee, FL 32306, USA
e-mail: ggaes@comcast.net

S. D. Camp
Federal Bureau of Prisons, Washington, DC, USA
e-mail: scamp@bop.gov

Keywords Classification · Criminogenic · Peer influence · Recidivism · Reentry

One of the pillars of correctional practice is to place inmates with similar security risk scores into the same prison. The security procedures of the prison are designed to be commensurate to the risk posed by the inmate. Classification risk scores are primarily based on past criminal and institutional behavior, so that prisoners with more serious and violent backgrounds are assigned to higher security prisons.[1] There are at least three primary reasons why inmate classification is used (Brennan 1987; Glaser 1987). Higher risk inmates are separated from lower risk inmates so that the former will not extort, assault, or intimidate the latter. Prison officials can use procedural and technological methods to impose greater constraints to suppress violence among the higher risk inmates. Finally, by limiting the extra security methods to only higher risk inmates, the overall system is more efficient and administrators do not need to squander scarce prison resources on inmates who do not need the additional security procedures.

Recently, scholars have questioned whether classification assignments actually meet those goals and whether they have unintended post-release consequences (Bench and Allen 2003; Chen and Shapiro 2007). At a broader level, the concern is that prisons may make individuals more criminal upon release, the so-called criminogenic effect (Vieraitis et al. 2007). These discussions raise important issues about the impact of the deprivations associated with the social environment of prisons, and the influence of high risk peers on those confined to higher security institutions (Austin 2001; Carceral 2004; Clear 1994; Gendreau et al. 1999; Irwin 2005; Petersilia 2003). This paper describes evidence from an experimental study which tested whether assignment to higher security level produced elevated rates of recidivism after release. To our knowledge, this is the only experimental test of security level assignment on recidivism. In this paper we also review the evidence on the relationship between security level assignment and institutional misconduct. The security environment in which an inmate is placed could suppress the misconduct anticipated and predicted by the classification procedure. Bushway and Smith (2007) raise similar concerns in their discussion of sentencing and other criminal justice contexts where high risk offenders have criminality suppressed by the system and low risk offenders have a more liberal environment to express potential criminality. To frame this discussion of the potential criminogenic influence of security level placement, we describe a three-tiered conceptual model that distinguishes among institution regime, inmate culture, and criminal propensity.

Institution regime refers to the formal organization of the prison. Typically, the regime has the primary goal of controlling inmate behavior and even staff behavior in support of institution goals such as safety and security. However, administrators may also emphasize other agency goals such as rehabilitation, retribution, or

[1] Inmate classification has also been designed to estimate the level of escape risk for inmates, but this prediction may not have the same specification as risk of violence and is secondary to the arguments posed in this paper. However, prison systems that try to maximize the prediction of both of these elements, violence and escape risk, in one equation may be introducing error that limits the predictability of each of these elements separately.

repairing harm to the victim. Agency missions and goals directly and indirectly influence the prison social environment (Lin 2000; Riveland 1999). Even the physical features of a prison, such as fences, surveillance cameras, and its architectural design, determine the types of interactions that occur among inmates and between inmates and staff. There are also the bureaucratic rules that codify the expectations for inmate and staff behavior. Prison security level, more than any other prison level variable, identifies variation in all of these dimensions of prison regime. Institution regime is also the dimension that scholars have associated most commonly with prison deprivation (Toch 1977).

Inmate culture is the collective response of inmates to the institution regime that is captured in the values, the informal rules and the behaviors of inmates. Significant differences in opinion exist over the origins and homogeneity of inmate culture, especially whether inmate culture is imported into the prison or whether it develops in response to prison conditions. The inmate culture that surfaces in prison has been the subject of classic volumes by Sykes (1958), Clemmer (1940), Jacobs (1976), Toch (1977), Irwin (1980), and Carrol (1974). In this paper we focus on the influence of the criminal attitudes and criminal lifestyles of the inmate's peers that develop from, and are part of, the inmate culture.

Finally, there is the criminal propensity of individual inmates. There is a widespread acceptance among criminologists of the basic principle that past behavior is the best predictor of future behavior. Inmates are believed to have different criminal histories that precondition their probability of future criminal behavior. Our interest lies in distinguishing a single person's propensity to commit crime relative to the collective propensity of his or her peer group.

To study the impact of security level assignment using this conceptual model, a research design must explicitly manipulate the three levels through experimental design, or control for the three levels using a strong quasi-experimental design. Such a design is required to sort out the three levels of influence on offender behavior, both within prison and after release from prison.

In the remainder of this paper we first review the few studies that have actually addressed how security level placement affects inmates' institution behavior as well as inmates' attitudes; we then turn our attention to post-release indicators of crime. Following this discussion, we consider the literature on peer influence on juveniles. Previous analysts have suggested that one of the criminogenic mechanisms in prison might be the influence of criminal peers within the prison environment. Until recently, there has been no empirical work that we are aware of that directly addresses this issue in the literature on adult prisoners. We reviewed the few studies that have been designed to test peer influences among youth in juvenile justice settings, and we tried to draw from this body of evidence theoretical implications for our own study. We then explain how we were able to capitalize on classification research conducted by Berk and his colleagues (2003) that involved random assignment to prison security level to do the first experimental test of the influence of security level placement on post-release criminality. Following the results, we explore the theoretical and policy implications of our research.

The impact of security level assignment on prisoner behavior and attitudes

There has been an abundance of theoretical and psychometric research on the classification instruments that have been designed to predict and classify risk for

individuals involved in different criminal justice processes, including prison (see volumes by Champion 1994, and Gottfredson and Tonry 1987), but only a few papers have evaluated the impact of security level placement on subsequent inmate behavior (Bench and Allen 2003; Berecochea and Gibbs 1991; Berk and de Leeuw 1999; Chen and Shapiro 2007; Camp and Gaes 2005). Instruments such as the level of supervision inventory (Bonta and Motiuk 1992), the salient factor score (Gottfredson et al. 1978) and the Federal Bureau of Prisons inmate classification system (Kane 1986) have demonstrated that risk assessment instruments can be developed that predict prison misconduct, post-release recidivism, and other criminal justice behaviors. There has also been a great deal of psychometric work to minimize the misclassification of inmates into erroneous risk categories. Part of the analysis challenge in prison systems that use these risk classification devices to assign inmates to prison security levels is that once these instruments have been developed, inmates are assigned to an institution security level that may condition the behavior the instrument is designed to predict (Bushway and Smith 2007).

Throughout this paper we distinguish between the security level of an inmate and the security level of the prison. The former is based on the results of the classification system where cutoff scores determine the appropriate level of custody. Thus, we refer to a level III inmate based on the classification score. The security level of a prison is based on the kinds of formal controls used in a prison and other features of an institution that we subsume under prison regime. Level IV prisons are much more secure than lower level prisons. Most prison systems use categories such as minimum, medium, and maximum. Since our data came from the California Department of Corrections and Rehabilitation (CDCR), we adopt their Roman numeral convention. While, at times, it is awkward when we have to refer to level III inmates in level I prisons or similar language, it should be clear that we are distinguishing between the criminal propensity of the offender measured by his or her security classification level and the features of the prison that are designed to exert formal control over the inmate population. The higher the security level of the inmate, the more risk he/she poses to the prison and outside communities. The higher the security level of the prison, the greater the number of procedures and technological constraints prison officials employ to control inmates' misconduct.[2]

In one of the earliest papers that directly tested the assignment of prison security level on serious misconduct committed inside prison, Berecochea and Gibbs (1991) used a logistic regression to evaluate simultaneously the inmate's classification score, the length of risk period, and four levels of prison security. Berecochea and

[2] During the data collection for the original study by Berk et al., approximately 20,000 inmates were initially placed in the following types of California facilities: 2.55% in reception centers, 13.82% in community corrections facilities, 25.54% in level I prisons, 31.92% in level II prisons, 21.42% in level III prisons, 4.46% in level IV prisons and 0.29% in special housing units. These assignments were based on security score thresholds established by the CDCR. For 75% of inmates, initial placement in a facility depended solely on their security score. Inmates with scores from 0 to 18 were placed in level I prisons; inmates with scores 19 to 27 were placed in level II prisons; inmates with scores 28 to 51 were placed in level III prisons; and inmates with scores 52 and above were placed in level IV prisons. For the other 25% of inmates, initial placement was based on administrative rules that were based on sentencing and other characteristics of the inmates deemed important by CDCR officials. These administrative rules 'trumped' the security score decision rules.

Gibbs found that, within the California prison system, only security level IV assignment had any effect, and it suppressed serious misconduct. Berecochea and Gibbs also reported on a natural experiment in which a select group of 'low risk' level IV inmates were placed in level III prisons that had enhanced security procedures. Over a period of time the low risk level IV inmates exhibited the same rate of serious misconduct as did the higher risk level III inmate. However, the natural experiment did not address whether the same level of misconduct for low risk level IV inmates would have been observed if they had remained in level IV prisons. Had some of the low risk level IV inmates remained in level IV prisons and been compared with their level IV counterparts placed in the high supervision level III prisons, the natural experiment would have been more relevant to theory and policy.

Bench and Allen (2003) were able to assign maximum security inmates randomly to both maximum and medium security facilities. This occurred in a double blind experiment where neither the inmate nor the supervising staff knew the actual security risk score. The maximum security inmates placed in maximum security prisons showed the same level of misconduct as did the maximum security inmates placed in medium security prisons. Bench and Allen also found that maximum security inmates assigned to medium security level institutions demonstrated the same level of misconduct as did medium security inmates placed in medium security prisons. Bench and Allen reasoned that their results could be interpreted to indicate that normal increases in misconduct across security levels may be due to a stigma (labeling) effect. This would explain why maximum and medium security level Inmates placed in medium security level prisons showed similar levels of misconduct when staff and inmates were no longer aware of their actual security levels. This would not explain why maximum security inmates placed in maximum security prisons would also have the same level of misconduct as their peers placed in medium level security prisons. One plausible explanation for not finding a difference between the maximum security inmates placed in maximum and medium security prisons is that the maximum security environment suppressed their misconduct. The results may also suggest that the particular instrument used in this jurisdiction has poor discriminatory power and cannot distinguish between the assumed differences in risk for medium and maximum security inmates.

In a prior paper (Camp and Gaes 2005), we examined the influence of security level assignment on prison misconduct. The study reported on experimental data gathered by Berk et al. (2003) to test changes to the inmate classification system used by the California Department of Corrections and Rehabilitation. As part of the original data collection and research design, level III security level inmates were randomly assigned to level I and level III institutions. We found that there were no statistically significant differences in total and serious prison misconduct between the two groups during the follow-up period. Berk et al. (2003) presented strong evidence that the new classification system improved the prediction of misconduct and discriminated quite nicely among the various levels of custody. Therefore, we can rule out that the null findings were due to a classification tool with poor discriminatory power. Level III inmates in level III institutions exhibited higher levels of total and serious misconduct than level I inmates in level I facilities did. The level III inmates who were randomly assigned to level III institutions based on

the new classification scoring system committed misconduct at the same rate as did other level III inmates in level III security institutions. Level III inmates randomly assigned to level I institutions had misconduct rates similar to those of the level III inmates in level III institutions and much higher rates of misconduct than the level I inmates placed in level I institutions had.

Based on these results, at least in California prisons, it does not appear that the prison environment exacerbates prison behavior. Furthermore, inmate disposition, measured by the classification risk score, is a good predictor of institution behavior. As noted, part of the challenge in conducting this research was that once the classification-risk tools had been developed, unlike in the study by Berk et al. (2003), all newly admitted inmates were assigned to a prison security level appropriate to their level of risk. This creates a selection problem in which inmate risk and security level placement are confounded, making it difficult to separate the effect of security placement from the effect of inmate risk. Other than random assignment, a second solution to this problem, demonstrated by Berk and de Leeuw (1999), was to use a regression discontinuity design. By evaluating inmates with similar security classification scores who had been assigned to prisons having different security levels, they were able to examine the impact of security level placement on inmates' misconduct. Berk and de Leeuw found that the security score of inmates placed in California prisons was a strong predictor of misconduct. For every additional 10 points on the risk assessment device, the odds of misconduct by an inmate increased by 1.22. The coefficient for the discontinuity threshold between level IV, the highest security level of California prisons, and all other security level prisons was also significant. The odds multiplier indicated that inmates placed in level IV prisons reduced their misconduct by half when compared with inmates placed in any of the lower security level prisons.

A recent study by Lerman (2009a, b) that was also conducted within the California Department of Corrections and Rehabilitation examined the influence of security level assignment on inmates' criminal psychology and inmates' social networks. Lerman used a regression discontinuity design to examine how security level placement affected an inmate's criminal personality and criminal cognitions. Both these psychological dimensions were derived from the correctional offender management profiling for alternative sanctions (COMPAS) tool developed by Brennan et al. (2006). The scales are based on inmates' responses to a series of questions. According to Lerman, "a high criminal personality is defined as being prone to anger and violence; and having a propensity for boredom and "get[ting] into trouble"; a talent for manipulating others; and a tendency to self-isolate from other people (Lerman, p. 16)." The cognitions scale also describes more than one dimension. One factor is the extent to which inmates justify their actions when they hurt other people. The second dimension describes whether inmates justify their criminal behavior. The assessment of these psychological scales was conducted just prior to each inmate's release from prison. Lerman found that inmates just above the security level II/III threshold had higher scores for criminal personality and criminal cognitions than did inmates just below that threshold *only if they had low prior criminal involvement*.

Lerman also used data from COMPAS to compare inmates with classification scores just above and below the level II/III threshold to evaluate the inmate's self-

reported social network in prison. She found that "...inmates with classification scores just above the cutoff have significantly more friends who have been arrested, friends who have been jailed, and friends involved in gangs. (Lerman 2009b: 19)." While this could be due to affiliations outside prison, Lerman argues it is more likely that these are network ties established within prison. She provides data that indicate that inmates who were unaffiliated to a gang prior to entering prison were more likely to join a gang the higher their security level assignment. What is even more striking is that inmates who were identified at admission as gang members by California authorities were much more willing to self-identify as a gang member when they were assigned to the level III prison as opposed to the level II prison. There were no such differences among inmates who had not been officially designated gang members at admission. She also presents evidence regarding the relationship between criminal cognitions and whether inmates join, renounce, or maintain gang membership, or abstain from gang membership. It appears that security level assignment elevates criminal cognitions for gang joiners and maintainers. She interprets this to mean that higher security level assignment increases the adoption of antisocial norms. The adoption depends on low prior criminal involvement, joining a gang during the current incarceration, or gang affiliation prior to imprisonment. She calls the adoption of antisocial norms the "dark side" of social capital.

Lerman's scholarship is the only work we are aware of that directly addresses the 'black box' of potential attitudinal, personality, and social network mediators of the relationship between security level assignment and subsequent conduct. She is concerned with the normative influence of peer relationships inside prison. As Lerman acknowledges, she could have had an even stronger design if COMPAS had been administered at prison intake. That would have allowed her to specify exactly how attitudes, dispositions, and affiliations had changed as a result of security level assignment. Furthermore, as in our study, her analysis was restricted by data limitations that prevented her from incorporating the riskiest and most violent inmates in the California Department of Corrections and Rehabilitation—those assigned to level IV prisons.

The impact of security level assignment on post-release behavior

A less obvious manifestation of prison security level assignment is its affect on post-release outcomes. While the in-prison misconduct results seem to confirm that prison security level placement does not increase institutional misconduct, and, at least at higher security levels, it dampens such behaviors, a more recent study raises the possibility that security placement has a criminogenic impact on post-release outcomes.

Chen and Shapiro (2007) examined the effect of different security level placements of inmates with similar classification scores on post-release arrests using a release cohort of Federal Bureau of Prisons' inmates. Those authors posed competing explanations for the impact of prison security level assignment. Underlying these explanations is the assumption that higher security level prisons are harsher environments and, as such, represent higher levels of punishment. To

support the contention that higher security level prisons are more harsh, Chen and Shapiro cited evidence from the Bureau of Prisons "Survey of Inmates of Federal Correctional Facilities" (Saylor 1984; U.S. Department of Justice, Bureau of Prisons 1991) to demonstrate that inmates in higher security levels have less contact with the outside world, less freedom, and are exposed to more violence.

Specific deterrence theory implies that a more severe level of punishment will reduce the individual's propensity to commit crime after release from prison. However, higher prison security levels will also expose inmates to cohorts of prisoners who are more disposed to crime, elevating the criminal human capital of each individual inmate. Chen and Shapiro also hypothesized that harsher prison conditions may lead to inferior labor market outcomes (Western et al. 2001), mediating the recidivism outcome. They argued that harsher prison environments were more likely to lead to skill depreciation and labeling (called "signaling" in their paper). However, their analysis of post-release employment outcomes seemed to rule out such an explanation. Inmates having similar security risk scores but who had been placed in different security level prisons had equivalent post-release employment success. Chen and Shapiro also suggested that the harsher prison conditions associated with higher prison security levels lead to what economists call a "taste for violence" (Banister et al. 1973).

To circumvent the selection problem, Chen and Shapiro also used a regression discontinuity design. They tested whether inmates who had similar security classification scores but were placed in higher security prisons had a higher hazard of arrest after their release from prison than inmates placed in the lower security level prisons. They controlled for linear, quadratic, and cubic trends in the security risk score and other covariates in the regression discontinuity analysis. Chen and Shapiro found that, relative to placement in a minimum security facility, assignment to a low security level prison doubled the daily hazard of post-release arrest. Minimum and low security prisons are the two lowest levels of security placement in the Bureau of Prisons, followed by medium and high security. At higher thresholds, such as medium versus high security level prisons, there was no impact of the discontinuity thresholds on the arrest outcomes; however, the number of observations at these discontinuity thresholds was small, and there may not have been sufficient power to test the effects at those levels. The regression discontinuity design tests effects at each discontinuity threshold, not the average effect of security level placement. Other analyses conducted by Chen and Shapiro indicated that security level placement tended to damage family ties and increase the level of post-release arrest for violent crimes. To our knowledge, the Chen and Shapiro study is the only direct test of the effects of adults' prison security level placement on post-release outcomes, other than our study, that also controls for selection artifacts.

The literature on peer influence in juvenile justice settings

While there are no other direct tests of security level placement on recidivism in the literature on adult correction, there are several studies in the literature on juveniles which are analogous and are based on arguments about the impact of peer influence on juvenile behavior. The security classification procedures for adult prisoners create

homogeneous risk pools, fostering the conditions under which peer influence effects could occur. The higher the security level placement, the more likely it is that an inmate will be placed with other inmates who have more criminal dispositions. Similar to this hypothesis is the idea that prisons are "schools of crime" (Letkemann 1973).

The argument of gain in criminal human capital made by Chen and Shapiro is consistent, although not identical, with the arguments made in the literature on juvenile peer influence. Human capital gains would imply some gain in knowledge about delinquency or crime; however, other mechanisms could include the reinforcement of deviant attitudes, and the potential to form partnerships and coalitions among youth to commit deviant and criminal acts [Bayer, P., Pintoff, R., and Pozen, D. E. (2004). Building criminal capital behind bars: social learning in juvenile corrections, unpublished].

Osgood and Briddell (2006) reviewed the literature on peer effects in juvenile justice settings. Their paper was included in a recent book devoted to juvenile peer influence on deviant behavior in many different settings, including: schools, mental health treatment facilities, child welfare programs, and street gangs (Dodge et al. 2006). According to Osgood and Briddell, the mechanisms fostering peer influence are still not completely understood. They explain that, historically, many of the current peer influence themes emanated from work by Clemmer (1940), Sykes (1958), and Goffman (1961), who each articulated different mechanisms about how the prison and other "total institutions" promoted a culture that was in opposition to the authority and the reform-minded intentions of the administrators. The literature on juveniles also includes studies in non-residential settings which are places of congregation for delinquents that seem to produce negative peer influence effects. Osgood and Briddell also point out that, separate from the security procedures employed by the institution, there are other dimensions of the environment that can affect the behavior of the residents. They cite the importance of whether the institution has a correctional or treatment orientation. This latter observation is important, because it focuses attention on the separate contributions of the prison or residential setting and its environmental characteristics, apart from the composition of the inmates or residents. This is consistent with the distinction we make between institution regime and inmate culture that we discussed earlier in this paper.

Osgood and Briddell note that there are not very many studies of peer effects in juvenile justice programs. They review studies by Buehler et al. (1966), Eddy and Chamberlain (2000), Feldman (1992), Feldman et al. (1983), Gold and Osgood (1992), and Bayer et al. (2004; unpublished). Although not reviewed by Osgood and Briddell, the study by Dishion and colleagues (Dishion and Andrews 1995; Dishion et al. 1999; Poulin et al. 2001) examined the impact of peer influence in a treatment program to prevent substance use. The peer influence in this latter study was most pronounced for juveniles who had started with moderate, as opposed to severe, levels of delinquency.

Although the literature on juvenile peer influence is not overly compelling, the weight of the evidence is that there are peer and/or institutional impacts that elevate criminality or delinquency for youth exposed to other delinquent peers. Although the Chen and Shapiro study confounds institution security level with inmate composition, some of the designs of the studies on juvenile peer influence emphasize peer group composition in situations where the institution level is the same, thus

controlling for institutional variation, There are also studies in this literature that vary the type of institution in which juvenile composition is the same, essentially controlling for peer composition. There is no study, either in the literature on juveniles or on adults, including ours, that experimentally varies institution level, peer composition, and individual risk level of the subject.

This study is the only experimental test of the effects of security level assignment on post-release outcomes. This is also the only study in an adult prison setting that tests the impact of institution and peer influences on both *institution* and *post-release* behavior. As noted earlier, in a prior paper using this same exact sample of inmates (Camp and Gaes 2005), we showed that there was no impact of security level placement on serious institutional misconduct among level III inmates who were randomly assigned to level I and level III prisons. Subsequently, with the assistance of the California Department of Corrections and Rehabilitation, we were able to obtain recidivism data on these same individuals, and we report those results in this paper.

Methods

The data for our study were provided by Richard Berk and the California Department of Corrections and Rehabilitation (CDCR). The data were initially collected to evaluate proposed changes to the CDCR inmate classification system. Berk and his colleagues were able to demonstrate that additional factors proposed for classification, such as age and street gang affiliation, added to the predictive power of a model of in-prison misconduct (Berk et al. 2003). For a period between 1 November 1998 and 30 April 1999, adult male inmates were classified under both the proposed new classification system and the old classification system; however, a randomization process was used to determine whether the inmate would be placed under the old or the new classification system. The initial purpose for collecting both the old and new classification scores was to monitor how the old and new systems would affect actual placement. Most of these inmates were assigned to the same security level based on the old and new scoring system. However, there was a subset of individuals whose classification levels were changed, and the majority of these, 561 inmates, were those whose scores required that they be placed in a level III prison rather than the level I prison prescribed under the old scoring system.

The classification randomization process was based on whether the inmate's identification number ended in an even or odd digit, and this determined whether the inmate was designated to a level I or a level III prison. The paper by Berk et al. describes, in some detail, how the randomization process was conducted. As California inmates arrived at the reception center, they were sequentially assigned a unique identification (ID) number. If their ID number ended with an even number, they were assigned to the group placed under the old classification system. If their ID number ended in an odd number, they were placed according to the new classification system. Berk et al. (2003) described how the randomization process was closely monitored and statistical tests of the balancing covariates were conducted to insure conformity to the randomization process. Our data are a subset of the classification data. There were 297 inmates who were classified as risk level III under the new system and were assigned to level I prisons, the lowest security level CDCR has in its

level I to level IV tiers. There were 264 inmates who were classified as level III under the new system and assigned to level III prisons. Recognizing that this procedure allowed us to test the criminogenic hypothesis experimentally, we contacted Richard Berk, who generously provided us with the data.

The data from the Berk study were supplemented with data on post-release recidivism provided by the CDCR. All the inmates in this classification study had been released on parole by the time CDCR provided the data. Recidivism was defined as recommitment to the California prison system. Recommitment to CDCR could occur because of a conviction for a new offense or a parole revocation. The data were primarily analyzed with Cox proportional hazards models, allowing us to control for time at risk as well as to test other covariates that might have mediated recidivism. Observations were censored at the end of the follow-up period, 21 September 2006, or if the inmate had died during the release period. The average total risk period for this sample was 5.9 years (range 192 to 2,832 days). This was defined as the amount of time from first release to the end of the follow-up period. Of course, many inmates were recommitted prior to the end of the follow-up period.

Results

Descriptive statistics and randomization tests

In their study Berk et al. reported on whether the individual variables that make up the composite security classification score were statistically balanced. They provided data to demonstrate that the experimental and control groups were balanced on all the classification variables including age, gang membership, criminal history, and history of violence. Because we had a subset of the original data, we asked CDCR to give us additional variables, and these are described in Table 1. We used these additional data to test balance in our subset of the original classification data set.

First, there was no significant difference in the probability of an individual's being assigned to the level I versus the level III prison. While we report the percentage of inmates committing serious misconduct in Table 1 and the rate of serious misconduct per months of time served, we do this purely for descriptive purposes. Strictly speaking, inmate misconduct occurred after the security level placement and was endogenous. It is really an outcome rather than a balancing variable. Despite the random assignment, there were differences between the two groups in time served and age at release. This is to be expected as a chance event, even in a random assignment study (Rosenbaum 2002). Level I inmates served 23 months, on average, compared with the level III inmates, who served only 16.3 months. The medians were respectively 19.4 months and 13.8 months, suggesting that some of the discrepancy was due to outliers. Both groups were relatively young when they entered prison. They were, on average, slightly over 22 years of age, and the test statistic shows their admission ages were equivalent, but the longer time served by the level I inmates made them a little older upon release. Since age at release is dependent on time served, one could argue that, substantively, there is only one difference in the group of covariates. Furthermore, age is one of the strongest inverse predictors of crime, and the somewhat older inmates released from

Table 1 Statistics testing balance data. There were 297 inmates assigned to level I prisons and 264 to level III prisons. Pearson's χ^2 was used to test for differences in categorical variables. Student *t*-test was used to compare continuously measured variables

Variable	Level I	Level III	Test Statistic	*P*
Assigned security level	52.94%	47.06	1.941	0.164
Rate of serious misconduct	0.055	0.070	1.51	0.133
Serious misconduct[a]	49.5%	46.6%	0.472	0.790
First arrested as juvenile[a]	77.1%	82.6%	2.584	0.275
Person crime[a]	43.8%	40.5%	0.602	0.740
Drug crime[a]	28.3%	26.9%	0.135	0.935
Other crime[a]	7.1%	9.8%	1.405	0.495
Property crime[a]	20.9%	22.7%	0.282	0.867
Time served (months)	23.0	16.3	6.571	0.000
Age at release	24.8 years	23.8 years	2.190	0.029
White[a]	14.1%	9.5%	2.901	0.235
Mexican[a]	29.3%	32.6%	0.706	0.702
Hispanic[a]	19.9%	25.8%	2.771	0.250
Other ethnic identity[a]	2.7%	w3.4%	0.243	0.885
Black[a]	34.0%	28.8%	1.763	0.414
Age at incarceration	22.8 years	22.5 years	0.976	0.329

[a] Categorical variables

security level III prisons should, if based on age alone, recidivate at a lower rate than their slightly younger counterparts released from security level I prisons. The groups were balanced on the other covariates (type of crime, instances of serious prison misconduct, whether first arrested as a juvenile, and ethnicity).

Analysis of recidivism

On average, 77.7% of level III inmates recidivated, while 68.7% of the inmates placed in level I prisons were returned to prison within the risk period. On average, level III inmates spent fewer days in the community after release than did the level I inmates, 740 days versus 839 days. Because inmates were at risk for different periods of time, the data were analyzed by hazard rate analyses.

Our first set of analyses was developed to show the shape and timing of the post-release hazard rate of recommitment. We did this to understand how the hazard rate changed over time and the period when inmates were most likely to be returned to prison. To recover the shape of the hazard rate, we used the SAS procedure LIFEREG and methods that tested twice the difference in the likelihood functions for nested models as described by Allison (1995: 88–89) to find the appropriate distributional form for the accelerated failure time model. The generalized gamma distribution seemed to fit the data best (scale parameter = 1.51; shape parameter = −1.02). The lognormal distribution was the closest alternative. The shapes of the hazard rates were the same for these two choices; however, the peak of the hazard for the gamma distribution was a hazard rate of 0.0024, while it was 0.0019 for the lognormal distribution.

We then used an SAS macro developed by Allison based on a completely specified accelerated failure time regression model to generate a graph of the hazard function. This appears in Fig. 1. The hazard rate increased for approximately the first 125 days after release from the CDCR and then declined. At its peak, the hazard rate was about 0.0024. Since the time interval of measurement was 1 day, this meant that an individual could be expected to be returned to the CDCR 0.0024 times in a 1 day interval approximately 125 days after release. Before running the inferential models, we used the SAS procedure Lifetest to graph the raw survival functions for the two groups of inmates. These survival graphs appear in Fig. 2. There is a clear difference in survival between the two groups starting from the very beginning of release to the community. Prisoners who had been assigned to a level III prison began failing at a higher rate upon release and continued to do so until about a 1,000 days after release from prison. At that point, the survival functions flatten and remain equidistant from each other. The majority of censoring occurs at 1,500 days and beyond.

Although we ran the accelerated failure time models to recover the shape of the hazard function, we estimated the impact of security level placement and other covariates using a Cox proportional hazards model. It is unnecessary to specify the distributional form of survival for this model. The results of the proportional hazards model are presented in Table 2. The only covariate included in this analysis was a dummy variable indicating that the respondent had been initially assigned to a level III prison as opposed to a level I prison. The variable is positive and statistically significant. Assignment to a level III prison increased the hazard of recidivating by 31.1%.

One of the conditions of Cox hazard models is that the hazard rates are proportional for the groups under study (Hosmer and Lemeshow 1999; Allison 1995). There is one simple test of the proportional hazards assumption. By entering

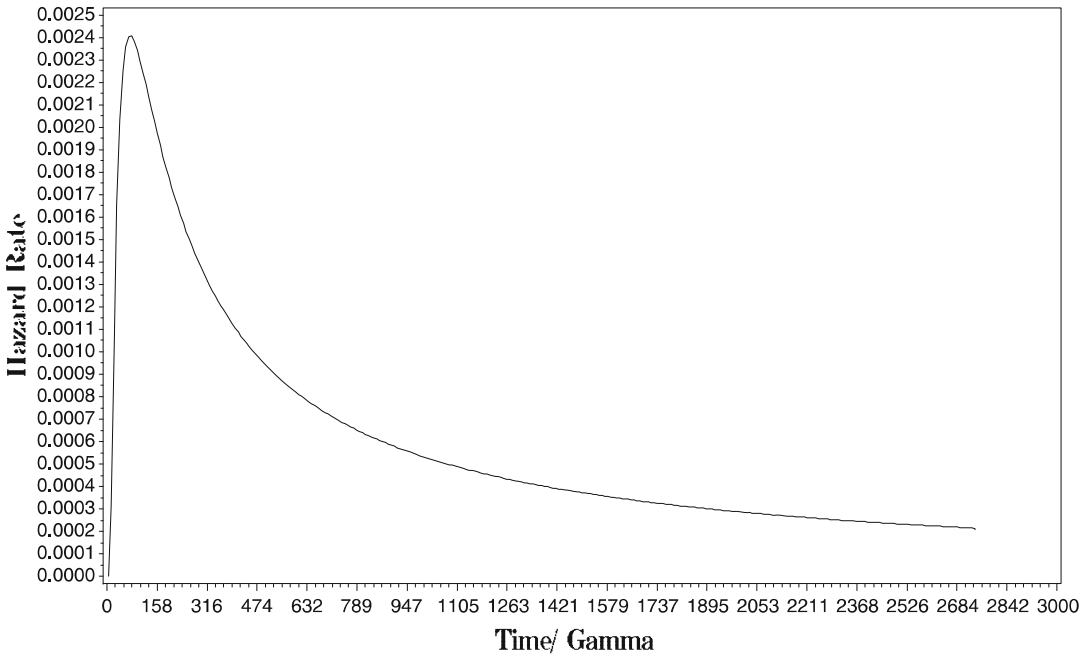

Fig. 1 Fitted hazard rate for return to prison

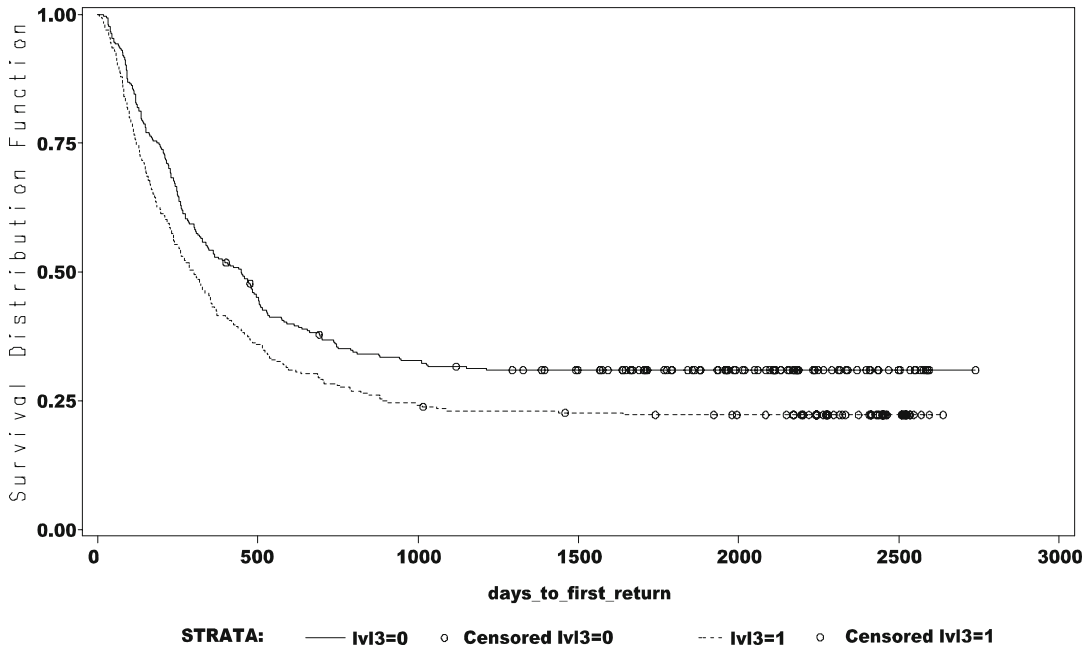

Fig. 2 Survival functions and censored observations for inmates released from level III (lvl3 = 1) and level I (lvl3 = 0) prisons

an interaction between the grouping variable and time, it is possible to see if the groups have similar hazard rates across the time units of the study. In this analysis the interaction between time to recidivism and assignment to level III facilities was not statistically significant, suggesting that the proportional hazards assumption had been met. We also ran a multivariate hazards model that included the covariates reported in Table 1. We did this to see if 'controlling' for other variables increased the efficiency of our estimates. Since we essentially recovered the same effect of security level placement and there were no surprises in interpreting the other covariates, we have not reported the results here. We also estimated

Table 2 Cox proportional hazards model for time to recidivism, prison assignment only (*DF* degrees of freedom)

Analysis of Maximum Likelihood Estimates						
Variable	DF	Parameter Estimate	Standard Error	Ch Squared	Pr > Chi Squared	Hazard Ratio
Assigned to Level III	1	0.27116	0.09896	7.5082	0.0061	1.311
Summary of the Number of Event and Censored Values						
Total	Event	Censored	Percent Censored			
561	409	152	27.09			
Testing Global Null Hypothesis: Beta = 0						
Test	Chi square	DF	Pr > Chi Squared			
Likelihood ratio test	7.4871	1	0.0062			
Score test	7.5542	1	0.0060			
Wald test	7.5082	1	0.0061			

multivariate models that also included dummy variables for year of release, but these had no impact on the model fit or the coefficients in the model. We also experimented with different parametric frailty models to see if heterogeneity in times-to-failure might impact upon our substantive conclusions (Gutierrez 2002). Although the frailty parameter was significant, it did not change the impact of the security level assignment on the hazard of recommitment. The results of all of these analyses are not presented here to conserve space, but they can be obtained from the first author.

Discussion and conclusions

The results of this study demonstrated that level III inmates placed in level III prisons were more likely to be readmitted to prison than were the level III inmates who were randomly assigned to level I prisons. By separating inmates into homogeneous risk pools, prison administrators are inadvertently increasing the likelihood that inmates will be recommitted to prison. These results comport with the findings of Chen and Shapiro (2007). However, instead of testing a one-level difference in security level placement, our study tested a two-level difference. This is a stronger test of the combination of environmental and peer influences than the Chen and Shapiro study was. This was not intended. It was a serendipitous result of the changes in classification procedures that CDCR introduced. Given the research design, we cannot assess whether it was a peer or environmental influence that caused the post-release effects, since these factors were confounded in this study as they were in the studies by Chen and Shapiro and Lerman.

Other than this and the study by Chen and Shapiro, the only other study we are aware of that has attempted to measure the effects of prison characteristics on post-release outcomes for adults was conducted by Drago et al. (2007) and was based on 13,000 prisoners released from Italian prisons. Drago et al. found no effect of the two institutional indicators of deprivation and crowding, and a proxy measure for prison health care on their measure of recidivism.

In the remainder of this "Discussions and conclusions" section, we discuss theoretical explanations for these results, policy implications, methodological and design considerations, and the relationship of this research to prior prisoner classification research.

Theoretical explanations for the experimental results

Knowing why we observed an effect of security level placement is just as important as the effect itself. Based on our prior results from this sample, we know that these two groups of inmates exhibited equivalent levels of total and serious misconduct during their institutional confinement, reinforcing the notion that their behavior in prison was consistent with their classification score. A possible explanation for the discontinuity in behavior between prison and the community is that security level assignment would exacerbate misconduct inside of prison if it were not for the dampening impact of higher security level features. Prior results by Berk and de

Leeuw (1999) in the same prison system demonstrated that security level placement in level IV prisons suppresses institution misconduct. Combining these behavioral results of suppression within prison and expression after prison, we draw the inference that, whatever the criminogenic impact is, it does not express itself until the post-release period. However, what if prison does increase the propensity of an inmate to commit crime over time? Then inmates whose classification scores indicate they are the same at the beginning of the assignment may, in fact, have different propensities after spending time in a more secure facility, and it is possible that the reclassification score based on behavior is not sensitive enough to measure this change. Lerman's study at least indicated that attitudes and dispositions may become more criminal for those with low prior criminal activity who are placed in level III prisons and that placement increases criminogenic affiliations. Perhaps the security features are just enough in a level III prison to suppress misconduct even though an inmate's disposition to commit such misconduct has increased over time. There may also be post-release environmental influences that might explain the apparent inconsistency in behavior between in-prison and post-release behavior.

If post-release supervision resources were based on security level assignment, rather than, or in addition to, the risk inmates pose based on their criminal history, this could be a reason why there would be an increase in the hazard rate for level III inmates placed in level III institutions. Supervising parole officers would more closely monitor the inmate released from level III prisons, and heightened supervision has been shown to elevate recidivism (Petersilia and Turner 1993). We discussed this possibility with researchers who are conducting extensive studies of parole supervision in California (Grattet and Rudes, personal communication; Grattet et al. 2008). This team of researchers is focusing on analysis of the parole supervision data that include extensive interviews with parole officers. At the time of this study there was no objective classification tool for parole risk in California.[3] According to these researchers, parole agents used the criminal history contained in "rap sheets" as their primary source to gauge the supervision levels of released offenders. The classification tool developed by Berk et al. (2003) for the California Department of Corrections and Rehabilitation explicitly incorporates criminal history in the classification score. Security level III inmates, whether placed in a higher or lower security prison, would have had comparable criminal histories. Therefore, we can probably rule out that California parole agents would have placed inmates released from level III prisons on higher levels of supervision than their level I counterparts, since all these inmates had equivalent criminal histories.

Is there a possible labeling explanation for these results? Labeling may be more important in the post-release environment than it is inside prisons. There are two

[3] One of the anonymous reviewers of this paper indicated that California had adopted COMPAS for post-release risk assessment and supervision and that we should acknowledge that in our paper. At the time of this study, all the offenders, with the possible, but unlikely, exception of four, had been released to supervision prior to the field testing conducted with COMPAS. During the period of our study, CDCR parole agents used characteristics of the instant offense and the pattern of criminal history to assess risk, but there was no tool to scale this information as is currently being done with COMPAS. This is documented on page 49 of the report by Grattet et al. (2008). Since the prison classification system is also based on criminal history, the level III inmates released from the level III and level I prisons would have had, on average, the same levels of supervision risk.

potential labeling mechanisms (Chiricos et al. 2007; Lofland 1969; Sherman et al. 1992). The first is based on an identity transformation, and the second on structural impediments in conventional life (Sampson and Laub 1997). For the first effect to occur, an ex-offender would have to integrate some aspect of the security level placement into his or her post-release self-identity. This seems unlikely, since the primary identity is still one of an ex-felon. For the second mechanism to work, the security level assignment would have to have some meaningful signal to the members of the community. However, it is unlikely that the security level assignment of the ex-felon would be known to the community or that it, alone, would somehow elevate the community's concern above what would be anticipated knowing only that the person is an ex-felon.

The mechanisms of behavioral change suggested by Chen and Shapiro (2007) were the harshness of the prison environment (primarily an environmental effect), increasing the human capital of the offender (social learning/peer effect), increasing a taste for violence (either peer or environmental effect), and inferior labor market outcomes due to labeling (primarily an environmental effect). All these mechanisms may have an impact on criminal dispositions but do not express themselves until release.

Routine activity theory (Cohen and Felson 1979) does make the pieces of the puzzle fit. All the mechanisms suggested by theorists that could affect the increased criminality in prison resulting from higher security level placement cannot express themselves until the offender is released to a community where there is more opportunity to commit crime. In prison, one of the three conditions posited in routine activity theory will occur if prison security levels enhance the disposition toward criminality. Offender motivation increases. However, there may be just enough formal control differences between level I and level III prison environments to suppress misconduct. Consistent with routine activity theory is the fact that prison is not a target rich environment. Once an inmate has been released, however, there is a target rich environment, fewer suitable guardians, and fewer of the formal controls found in prison. The peer literature, according to Osgood and Briddell, is faced with a similar quandary. While there is evidence for a peer 'contagion' effect, it is not clear what is causing the contagion. This is primarily because most studies do not measure intervening processes such as the interaction patterns that would provide insight into the contagion mechanism. This is a weakness in the literature on adults as well, although Lerman's research is a first glimpse into potential mediating mechanisms for adults.

Policy implications

As we noted in the introduction, inmate classification is designed to separate violent inmates from potential inmate victims and efficiently control inmate violence by increasing the security procedures at the higher security level prisons. Those internal management goals seem to have been met in the studies reviewed in this paper. Contrast that with the unintended consequences of increasing the hazard of recommitment once an inmate has been released. The public policy choices are limited. It would be unwise to assign everyone to the same level of custody regardless of their risk. This would provide more targets of low risk offenders for

violent high risk offenders. Furthermore, to maintain security among both high and low security inmates placed together, the security level procedures of the prison would have to be inordinately high for low security risk inmates in order to maintain control of the high security inmates. Prison administrators could experiment with different thresholds so that more prisoners might be pushed down to lower security levels. There is a monetary incentive to raise thresholds so that more inmates are placed in lower security levels that cost much less money per capita. The social cost, however, is that this might increase victimization. We are not aware of any study that considers what the classification score thresholds should be. This is surprising, given the importance they have in the operating and managing of prisons. Lerman (2009a) nicely captures the tension between lower and higher thresholds of formal control as the balance between trust and risk. As prison and community supervision officials lower the level of supervision, putting more trust in the offender, they expose themselves and the community to a greater degree of risk. As any parole or prison official will testify, the costs of reduced thresholds can be quite high, especially if there is shocking victimization. However, if peer pressures and harsher prison environments elevate criminality, the higher thresholds are also costly.

Another policy alternative is to address the reasons for elevated recidivism while offenders are under post-release community supervision. The parametric hazard model demonstrated that the risk increases post-release up to approximately 125 days then declines quite dramatically. Perhaps post-release supervision could address the rising hazard for these first 4 months after release. Practitioners could also devote more resources to higher risk inmates, consistent with the "risk principle" (Andrews et al. 1990; Dowden and Andrews 2000). The difference between this suggestion and the suggestions of prior researchers is that security level placement will have to be integrated into the post-release risk classification. Another way to address this problem is to offer more intensive rehabilitative programming while inmates are in prison, especially at the higher security levels. Yet another possibility is to use COMPAS or other post-release supervision risk classification instruments in assigning post-release supervision levels, provided they are reliable and valid tools. Researchers will have to be mindful of the same selection issues we have discussed in this paper when the validation is done on supervision populations. Unfortunately, this is a solution based on acceptance of the premise that harsher prisons create more criminals—a very unsatisfactory policy choice.

Chen and Shapiro (2007) also discuss the relevance of their findings in relation to the results of a study by Katz et al. (2003) that demonstrated that prisons with harsher conditions have a large contemporaneous deterrent effect. Katz et al. used in-prison mortality rates as their measure of prison conditions. Chen and Shapiro used prison mortality rates to simulate the net increase in re-arrest rates of assigning inmates to higher security levels relative to the general deterrent impact found by Katz and colleagues. While our results support the findings of Chen and Shapiro, we do not try to estimate the societal impact of elevated levels of crime that may be caused by higher security levels, but, instead, we focus on potential ways to ameliorate the problem. A note of caution, however, is in order. There are really only three studies that have directly tested the criminogenic effect of security level placement. Replications in other systems need to be conducted before any serious policy changes are undertaken.

Methodological and design considerations

One of the contributions of this study is that it eliminates identification problems that could be present in regression discontinuity (RD) designs. While RD designs have become very popular in many settings, they are not without some limitations. RD designs have been characterized as sharp or fuzzy (Morgan and Winship 2007; Trochim 2006). In the former there is a clear and sharp distinction between the scores of subjects just below and just above the discontinuity threshold. This is unlikely in a study which uses inmate classification scores as the underlying continuous variable. Staff are often allowed to use 'overrides' that allow them to place inmates in a prison below or above the security level indicated by their classification score, because of some subjective hunch or because there is supplemental information that suggests a security level assignment contrary to the objective score. Therefore, studies that use the classification score are more likely to involve fuzzy regression discontinuity in which there is error in the assignment process (Morgan and Winship 2007). There are instrumental variable (IV) approaches to this problem (Angrist and Lavy 1999), but IV designs are not without controversy either. Even if one could assume that the classification score and security level placement represent a sharp regression discontinuity design, the analyst must model the continuous score with the correct polynomial, otherwise the model could be misspecified, leading to a biased estimate of the security level assignment (Trochim 2006). Chen and Shapiro (2007) provided a great deal of supplemental analysis and chose a higher order polynomial to analyze their data, so it is unlikely that their estimate of the effect of security level placement was biased. Nevertheless, the experimental results give us more confidence in those findings. The experimental evidence presented here not only confirms the prior regression discontinuity result but does so in a different jurisdiction.

One of the limitations of this and prior studies has been the failure to test whether *high* security level placement has a post-release criminogenic effect. The work by Chen and Shapiro was limited by sample size at the higher security level prisons, and our study was limited because there were no random assignments of high security individuals to high or medium security prisons. Most likely such a study will have to be conducted using a regression discontinuity or some other strong quasi-experimental design. Secondly, there has been no study that has experimentally varied or has used an analytic approach to test the independent influences of peer composition and institutional characteristics. The criminogenic nature of prison security level assignment should be clarified by these design improvements. A stronger research design would eliminate some of the uncertainty in explaining these results. There are many difficulties in designing such an experiment. If you assign too many individuals of a given risk level to a particular prison, you change the peer composition of the design. Future studies should also incorporate pre- and post-test risk assessments that are sensitive to possible changes in criminal dispositions. That would allow researchers to monitor both dispositional and behavioral changes.

Another limitation of this study is that there were no assignments of lower security level inmates to higher security level prisons. Would we have observed even more effects of peer and deprivation influences on the lowest risk inmates placed in higher security levels? While prison officials can probably be persuaded to place

some higher risk inmates in lower security prisons, it would be difficult to persuade them to assign low security inmates to high security prisons, and, more importantly, such a design, in our opinion, would be unethical.

One of the problems with this setting and many applied settings is that while offenders are initially assigned to the 'treatment' they may, in fact, cross over to the alternative treatment or control condition. In the context of this study, inmates assigned to level III prisons may 'earn' their way into level I prisons by their good behavior. Inmates assigned to level I prisons may require closer supervision and move up in security level because of their misbehavior. One of the reviewers of this study was concerned about these crossover effects and its impact on the outcomes, and this is an important consideration. This study can be considered as having an intent-to-treat design. Fleiss et al. (2003) discuss this problem and note that intent-to-treat designs have two principles. First, all the subjects who are randomly assigned are analyzed; second, the subjects are analyzed "…as members of the group to which they were initially randomized, irrespective of any non-adherence, treatment crossovers, behavior, or anything else" (Fleiss et al., p. 166). By following an intent-to-treat design, it is possible to avoid potential selection artifacts that arise when the researcher analyzes only subjects who complete treatment, who comply with treatment, or who may end up with a treatment different from that intended (as treated clients). The effect of the intent-to-treat design is that it reduces statistical power while preserving the integrity of randomization. In clinical applications, researchers may want to know the impact of a procedure or drug when everyone complies with the intended treatment. In a real world setting such as prisons, it is unlikely that we will ever be able to achieve a policy impact much beyond our intention to influence people for whom the policy is intended. In some applied experimental settings, one can use the random assignment process as an instrumental variable to uncover the 'as treated' effects (Angrist et al. 1996), but the estimate of the instrument depends on untested assumptions (Berk 2004: 226–230).

If crossover is extremely high, this can vitiate the intent-to-treat design. However, that would be an unlikely situation in the current context. The data on whether inmates were transferred to lower or higher security levels after assignment were not collected in the original study, and we could not obtain those data from CDCR. The only insight we have is from Berk et al. (2003), who reported on reclassification results in their original study. The CDCR has a mandatory yearly review of classification. Essentially, the same data used in the original classification instrument are assessed each year, taking into account the inmate's prison conduct. When regressing the reclassification score on the original classification score, Berk et al. found that the coefficient was 1.0, for both the experimental and control classification groups in the original study, and that the intercepts indicated that the classification scores had declined by approximately 2 points for the control group and 4 points for the experimental groups. While this is only a very indirect assessment of the consistency in the classification scores, it does imply that the original score maps extremely close to the reclassification score and that, on average, inmates' 'risk' within prison declines by about 10–20% over the first year. The data do not suggest that there is a lot of security level crossover. In fact, improvements to the classification that were achieved when Berk et. al. redesigned the system should have reduced the level of subsequent reclassification crossovers.

Classification research

In a recent special volume of *Crime and Delinquency* devoted to risk and needs assessment, Gottfredson and Moriarty (2006) reviewed the methodological issues in developing and applying classification instruments. Those authors conclude that risk assessment has been fraught with poorly designed tools, problems in implementation, and failures to meet basic methodological requirements. Part of the problem, according to Gottfredson and Moriarty, is that tools developed to predict events such as misconduct or crime are also used to treat offenders in an effort to reduce their criminal propensity. Gottfredson and Moriarty point out that risk assessment and treatment are two separate objectives, and treatment requires a different set of assessment tools and even a different set of criteria. As we have shown in this paper, even when risk assessment is done properly, the public policy result may have two consequences. One is intended and is consistent with the controlling of serious misconduct and crime inside prison. The other is unintended and is consistent with the promotion of factors that elevate the likelihood an offender will be recommitted the higher the security level placement.

Acknowledgments We are indebted to Richard A. Berk, University of Pennsylvania, and Chris Hawes, California Department of Corrections and Rehabilitation (CDCR), for their assistance in providing data for this study. Richard Berk was especially gracious in taking the time to insure we obtained and understood the classification data. Chris Hawes went out of his way to tutor us on the way in which the recidivism data in CDCR are collected and the meaning of the sundry codes that make any administrative dataset a challenge to analyze. We are grateful to Ryken Grattet, Jesse Shapiro and Bill Rhodes for extensive feedback, insight, and advice on various topics related to this manuscript. David Weisburd, the editor, and two anonymous reviewers also provided very helpful comments. The opinions expressed in this paper are those of the authors and do not necessarily represent the opinions of the Federal Bureau of Prisons, or the U. S. Department of Justice.

References

Allison, P. (1995). *Survival analysis using SAS: A practical guide*. Cary, N.C.: SAS Institute.

Andrews, D. A., Bonta, J., & Hoge, R. D. (1990). Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behavior*, *17*, 19–52.

Angrist, J. D., & Lavy, V. (1999). Using Maimonides' rule to estimate the effect of class size on scholastic achievement. *Quarterly Journal of Economics*, *114*, 533–575.

Angrist, J., Imbens, G. W., & Rubin, D. W. (1996). Identification of causal effects using instrumental variables (with discussion). *Journal of the American Statistical Association*, *91*, 441–472.

Austin, J. (2001). Prisoner reentry: current trends, practices, and issues. *Crime and Delinquency*, *47*, 314–334.

Banister, P. A., Smith, F. V., Heskin, K. J., & Bolton, N. (1973). Psychological correlates of long-term imprisonment. *British Journal of Criminology*, *13*(4), 312–330.

Bench, L. L., & Allen, T. D. (2003). Investigating the stigma of prison classification: an experimental design. *The Prison Journal*, *83*(4), 367–382.

Berecochea, J. E., & Gibbs, J. B. (1991). Inmate classification: a correctional program that works? *Evaluation Review*, *15*(3), 333–363.

Berk, R. (2004). *Regression analysis: A constructive critique*. Thousand Oaks, CA: Sage.

Berk, R. A., & de Leeuw, J. (1999). Evaluation of California's inmate classification system using a generalized discontinuity design. *Journal of the American Statistical Association*, *94*(448), 1045–1052.

Berk, R. A., Ladd, H., Graziano, H., & Baek, J. H. (2003). Randomized experiment testing inmate classification systems. *Criminology and Public Policy*, *2*(2), 215–242.

Bonta, J., & Motiuk, L. L. (1992). Inmate classification. *Journal of Criminal Justice*, *20*, 343–353.

Brennan, T. (1987). Classification: An overview of selected methodological issues. In D. M. Gottfredson, & M. Tonry (Eds.), *Prediction and classification: Criminal justice decision making, crime and justice, a review of research, vol. 9.* pp. 201–248. Chicago: University of Chicago Press.

Brennan, T., Dieterich W., & Oliver, W. (2006). California Department of Corrections, Parole, and Community Services Division: COMPAS Pilot Psychometric Report. Traverse City, Michigan, Northpointe Institute for Public Management.

Buehler, R. E., Patterson, G. R., & Furniss, J. M. (1966). The reinforcement of behavior in institutional settings. *Behavior Research and Therapy*, *4*, 157–167.

Bushway, S., & Smith, J. (2007). Sentencing using statistical treatment rules: what we don't know can hurt us. *Journal of Quantitative Criminology*, *23*, 377–387.

Camp, S. D., & Gaes, G. G. (2005). Criminogenic effects of the prison environment on inmate behavior: Some experimental evidence. *Crime and Delinquency*, *51*(3), 425–442.

Carceral, K. C. (2004). *Behind a convicts eyes: Doing time in modern prison*. Belmont CA.: Wadsworth.

Carrol, L. (1974). *Hacks, blacks, and cons: Race relations in a maximum security prison*. Lexington, MA: Lexington Books.

Champion, D. C. (1994). *Measuring offender risk: A criminal justice sourcebook.* Westport, CT: Greenwood Press.

Chen, M. K., & Shapiro, J. M. (2007). Do harsher prison conditions reduce recidivism? A Discontinuity-based approach, *American Law and Economics Review*, Advance Access, published June 12, 2007.

Chiricos, T., Barrick, K., Bales, W., & Bontrager, S. (2007). The labeling of convicted felons and its consequences for recidivism. *Criminology*, *45*(3), 547–582.

Clear, T. R. (1994). *Harm in American penology: Offenders, victims, and their communities*. Albany, NY.: University of New York Press.

Clemmer, D. (1940). *The prison community.* New York: Rinehart.

Cohen, L. E., & Felson, M. (1979). Social change and crime rate trends: a routine activity approach. *American Sociological Review*, *44*, 588–608.

Dishion, T. J., & Andrews, D. W. (1995). preventing escalation in problem behaviors with high-risk young adolescents: immediate and 1-year outcomes. *Journal of Consulting and Clinical Psychology*, *63*, 538–548.

Dishion, T. J., McCord, T. J., & Poulin, F. (1999). When interventions harm: peer groups and problem behavior. *American Psychologist*, *54*, 755–764.

Dodge, K. A., Dishion, T. J., & Langsford, J. E. (2006). *Deviant influences in programs for youth*. New York: The Guilford Press.

Dowden, C., & Andrews, D. A. (2000). Effective correctional treatment and violent reoffending: A meta-analysis. *Canadian Journal of Criminology*, 449–467.

Drago, F., Galbiati, R., & Vertova, P. (2007). Prison conditions and recidivism, Centre for Economic Policy Research papers, 6401. http://ideas/repec/.org/e/pga171.html

Eddy, M., & Chamberlain, P. (2000). Family management and deviant peer association as mediators of the impact of treatment condition on youth anti-social behavior. *Journal of Consulting and Clinical Psychology*, *68*, 857–863.

Feldman, R. A. (1992). The St Louis experiment: effective treatment of anti-social youths in prosocial peer groups. In J. McCord, & R. E. Tremblay (Eds.), *Preventing antisocial behavior: Interventions from birth to adolescence* (pp. 233–252). New York: Guilford Press.

Feldman, R. A., Caplinger, T. E., & Wodarski, J. S. (1983). *The St. Louis conundrum: The effective treatment of anti-social youth*. Englewood Cliffs, N.J.: Prentice-Hall.

Fleiss, J. L., Levin, B., & Paik, M. C. (2003). *Statistical methods for rates and proportions* (3rd ed.). Hoboken, New Jersey: Wiley.

Gendreau, P., Goggin, C., & Cullen, F. T. (1999). The effects of prison sentences on recidivism: Report to the Department of the Solicitor General Canada. Available online: http://www.sgc.gc.ca.

Glaser, D. (1987). Classification for risk. In D. M. Gottfredson, & M. Tonry (Eds.), *Prediction and classification: Criminal justice decision making, crime and justice, a review of research, vol. 9.* pp. 249–292. Chicago: University of Chicago Press.

Goffman, E. (1961). *Asylums: Essays on the social situation of mental patients and other inmates*. Chicago: Aldine.

Gold, M., & Osgood, D. W. (1992). *Personality and peer influence in juvenile corrections*. Westport, CT: Greenwood Press.

Gottfredson, S. D., & Moriarty, L. J. (2006). Statistical risk assessment: old problems and new applications. *Crime and Delinquency*, *52*, 178–200.
Gottfredson, D. M., & Tonry, M. (1987). *Prediction and classification: Criminal justice decision making, crime and justice, a review of research, volume 9*. Chicago: University of Chicago Press.
Gottfredson, D. M., Wilkins, L. T., & Hoffman, P. B. (1978). *Guidelines for parole and sentencing*. Lexington, MA.: Heath.
Grattet, R., Petersilia, J., & Lin, J. (2008). *Parole violations and revocations in California*. University of California Davis and University of California Irvine: Center for Evidence-Based Corrections.
Gutierrez, R. G. (2002). Parametric frailty models. *The Stata Journal*, *2*(1), 22–44.
Hosmer, D. W., & Lemeshow, S. (1999). *Applied survival analysis: Regression modeling of time to event data*. New York: Wiley.
Irwin, J. (1980). *Prisons in turmoil*. Boston: Little Brown.
Irwin, J. (2005). *The warehouse prison: Disposal of the new dangerous class*. Los Angeles, CA.: Roxbury.
Jacobs, J. (1976). *Stateville: The penitentiary in mass society*. Chicago, Il: University of Chicago Press.
Kane, T. R. (1986). The validity of prison classification: an introduction to practical considerations and research issues. *Crime and Delinquency*, *32*(3), 367–390.
Katz, L., Levitt, S. D., & Shustorovich, E. (2003). Prison conditions, capital punishment, and deterrence. *American Law and Economic Review*, *5*(2), 318–343.
Lerman, A. (2009a). The people prisons make: effects of incarceration on criminal psychology. In S. Raphael, & M. A. Stoll (Eds.), *Do prisons make us safer? The benefits and costs of the prison boom*. New York: Russel Sage (in press).
Lerman, A. (2009b). Bowling alone (with my own ball and chain): effects of incarceration and the dark side of social capital, Princeton University.
Letkemann, P. (1973). *Crime as work*. Englewood Cliffs, N.J.: Prentice Hall.
Lin, A. C. (2000). *Reform in the making: The implementation of social policy in prison*. Princeton, N.J.: Princeton University Press.
Lofland, J. (1969). *Deviance and identity*. Englewood Cliffs, NJ.: Prentice-Hall.
Morgan, S. L., & Winship, C. (2007). *Counterfactuals and causal inference: methods and principles of social research*. New York, NY: Cambridge University Press.
Osgood, D. W., & O'Neill Briddell, L (2006).. Peer effects in juvenile justice, In K. A. Dodge, T. J. Dishion, & J. E. Lansford (Eds.), *Deviant peer influences in programs for youth* (pp. 141–161). New York: The Guilford Press.
Petersilia, J. (2003). *When prisoners come home*: *Parole and prisoner reentry*. Oxford, U.K.: Oxford University Press.
Petersilia, J., & Turner, S. (1993). Intensive probation and parole. In M. Tonry (Ed.), *Crime and justice: An annual review of research* (pp. 281–335). Chicago, Ill.: University of Chicago Press.
Poulin, F., Dishion, T. J., & Burraston, B. (2001). 3-Year iatrogenic effects associated with aggregating high-risk adolescents in cognitive-behavioral preventive intervention. *Applied Developmental Science*, *5*, 214–224.
Riveland, C. (1999). Prison management trends, 1975–2025. In M. Tonry, & J. Petersilia (Eds.), *Crime and justice: A review of research*, vol. 26 (pp. 163–204). Chicago Il: University of Chicago Press.
Rosenbaum, P. R. (2002). *Observational studies* (2nd ed.). New York: Springer.
Sampson, R. J., & Laub, J. H. (1997). A life course theory of cumulative disadvantage and the stability of delinquency. In T. P. Thornberry (Ed.), *Developmental theories of crime and delinquency*. New Brunswick, NJ: Transaction Publishers.
Saylor, W. G. (1984). *Surveying prison environments*. Washington, DC: Federal Bureau of Prisons.
Sherman, L. W., Smith, D. A., Schmidt, J. D., & Rogan, D. P. (1992). Crime, punishment and stake in conformity: legal and informal control of domestic violence. *American Sociological Review*, *57*, 680–690.
Sykes, G. (1958). *Society of captives*. Princeton, N.J.: Princeton University Press.
Toch, H. (1977). *Living in prison*. New York: Free Press.
Trochim, W. M. (2006). *The research methods knowledge base* (2nd edn.). Internet WWW page at URL: <http://www.socialresearchmethods.ner/kb/> (version current as of October 20, 2006).
U.S. Dept. of Justice, Bureau of Prisons. (1991). Survey of Inmates of Federal Correctional Facilities, 1991 (computer file). ICPSR version. Washington, DC: U.S. Dept. of Commerce, Bureau of the Census (producer), 1991. Ann Arbor, MI: Inter-university Consortium for Political and Social Research (distributor), 2004.
Vieraitis, L. M., Kovandzic, T. V., & Marvell, T. B. (2007). The criminogenic effect of imprisonment: Evidence from state panel data, 1974–2002. *Criminology and Public Policy*, *6*(3), 589–622.

Western, B., Kling, J. R., & Weiman, D. F. (2001). The labor market consequences of incarceration. *Crime and Delinquency*, *47*(3), 410–427.

Gerald G. Gaes is a criminal justice consultant and Visiting Faculty at Florida State University in the College of Criminology and Criminal Justice in the USA. He was a Visiting Scientist for the National Institute of Justice, where he was senior advisor on criminal justice research, funded by that agency. He was also Director of Research for the Federal Bureau of Prisons and retired from government service in 2002. His current research interests include prison sexual victimization, spatial data analysis of crime, cost benefit analysis of inmate programs, the impact of prison security assignment on post-release outcomes, prison privatization, evaluation methodology, inmate gangs, simulating criminal justice processes, prison crowding, prison violence, electronic monitoring of community supervision cases, and the effectiveness of prison program interventions on post-release outcomes.

Scott D. Camp is a Senior Social Science Analyst at the Federal Bureau of Prisons in the USA. He joined the office in 1992 after completing his Ph.D. in Sociology at The Pennsylvania State University, USA. Much of his current research focuses on performance measurement and program evaluations. He also publishes on prison privatization, diversity issues, and inmate misconduct.