Throwing Away the Key: Measuring Prison Reform Attitudes

PAUL J. SILVIA

University of North Carolina at Greensboro

The American prison system is larger than ever and ranks among the largest in the world. Yet, prisons have received little research attention relative to other issues in forensic psychology. In an effort to study one facet of the prison system, a scale for measuring attitudes toward prison reform was developed. The 12-item scale has a single factor measuring whether people feel prisons should be tougher or softer on inmates. Several studies with diverse samples found that the scale has a consistent factor structure, good reliability, and a coherent pattern of relationships to other psychological variables. The scale appears to be a promising tool for studying how people want to treat those who have been officially rejected by society.

The American prison system has never been larger and ranks among the largest in the world. More than 5 million Americans are under correctional supervision; over 1.9 million are inmates in state and federal prisons and in local jails (Beck & Harrison, 2001). Prisons reflect society’s struggles with racial inequality, gender relations, and drug policies. At the end of 2000, for example, 9.7% of African American men between the ages of 25 and 29 were in prison, compared to 2.9% of Hispanic males and 1.1% of White males in the same age range (Beck & Harrison, 2001). The number of female inmates doubled during the decade 1990-1999, and the number of inmates in federal prisons for drug offenses increased nearly 125%. Drug offenses alone accounted for over 60% of the federal prison system’s total growth in that decade (Beck & Harrison, 2001).

Yet, Americans are surprisingly unaware of prisons as a social institution. Flanagan (1996) noted in his surveys, “Americans’ opinions about correctional policy and practices were much less fully developed than views about law enforcement or other components of the criminal justice system” (p. 75). This relative ignorance is mirrored by the scientific literature on forensic psychology.

1I would like to thank the students in the Measurement of Attitudes seminar for their comments on an early version of the scale. Larry Wrightsman deserves special thanks for his advice throughout this project.

2Correspondence concerning this article should be addressed to Paul Silvia, Department of Psychology, P. O. Box 26170, University of North Carolina at Greensboro, Greensboro, NC 27402-6170. E-mail: p_silvia@uncg.edu
People who study the legal system recently have called for more research on the legal system’s correctional aspects (e.g., Johnson, 2001). A large literature addresses the psychology of courtroom procedures, eyewitness accuracy, jury decision making, and so forth, but relatively little is known about the psychology of prisons and imprisonment. This is unfortunate, for both practical and conceptual reasons. On the practical side, the enormity of the prison system makes it an important public institution that deserves social scientific attention. More conceptually, interesting psychological issues emerge whenever one person rejects and punishes another (Feather, 1999; Robinson & Darley, 1995; Tsang, 2002).

Developing a Prison Reform Attitude Scale

There is no question that prisons are aversive places, both for the inmates and the employees (Kauffman, 1988; Paulus, 1988; Whitehead, 1989). Yet, just how aversive should they be? The public discourse surrounding prisons as social institutions shows conflicting positions. According to one camp, the public needn’t worry about making prisons “cozy”—inmates should have considered the unpleasantness of prison before committing crimes. Indeed, in this point of view, prisons must be unpleasant to deter crime and to punish offenders properly. This position has recently found its way into policy, as some states have returned to “chain gangs” and gravel smashing. According to the other camp, prisoners are still American citizens with many core civil rights. The public should thus be concerned about prison living conditions, ensuring prisoners’ safety, and maintaining dignity more broadly. This position is usually enacted by the courts, rather than through legislation, such as when prisoners are guaranteed access to law libraries and the services needed to sue prison employees.

One example of debate over tougher and softer changes in the prison system involves “Supermax” prisons (Riveland, 1999). Inmates in super-maximum security prisons spend their time in solitary cells sealed with solid doors, usually without reading material, opportunities for education, and social contact. In some Supermax facilities, inmates spend all but 3 hr a week in their cells. Some politicians and correctional authorities advocate expanding Supermax systems, seeing them as essential deterrents to prisoner misconduct. Others, in contrast, advocate discontinuing Supermax confinement, citing their higher rates of inmate mental illness and declaring that the system reflects “a stunning disregard of the fact that all prisoners . . . are members of the human community” (Human Rights Watch, 2000).

In both sides we see an emphasis on reform in its broad, literal sense—prisons need to be changed from their present inadequate form to some new and better state. They need to be more punishing or more humanizing, tougher or softer, worse or better. A scale was developed to capture this aspect of prison
attitudes. Public opinion surveys often include single items relevant to prisons, such as attitudes toward mandatory minimum sentences or attitudes toward building new prisons. I was unable to find, however, a psychometrically sound scale devoted to attitudes toward how prisons ought to treat prisoners. Such a scale would be useful in researching how prison reform attitudes contribute to policy attitudes, public discourse, and ultimately policy enactment.

Study 1

Scale Development

An initial pool of 50 items was developed. Nearly all of the items were adapted from newspaper articles on prison policies and Internet sites devoted to prisoner advocacy. This ecological approach to item generation increases the chance that the scale’s items accurately reflect the universe of attitude content (DeVellis, 1991). Items were discarded based on qualitative and quantitative considerations.

The first step in scale development was a group item-evaluation session. As part of a seminar on attitude measurement, graduate students and the instructor evaluated the items for clarity, potential bias, and relevance to the construct of prison reform attitudes. This process led to discarding 20 items. In the second step, 132 undergraduates at the University of Kansas (KU) completed the 30-item version of the scale. Ten items were excluded because of skewed distributions, low factor loadings, and poor contributions to internal consistency. In the third step, 400 undergraduates completed the 20-item version of the scale. Eight items were rejected because of factor loadings below .40 on the first factor, loadings on factors defined by a single item, and poor contributions to internal consistency.

A final 12-item scale resulted, as displayed in Table 1. The scale is scored in the liberal reform direction: Higher scores reflect positive attitudes toward progressive reforms.

Factor Structure and Reliability

Sample 1: College students. To test the final scale’s factor structure, 561 undergraduate students (288 women, 273 men) at KU completed the Prison Reform Attitudes (PRA) scale during a large mass-testing session. Each item was answered using a Likert-type format with a 7-point scale ranging from 1 to 7. Alpha for the 12 items was .83; no item would have increased alpha upon exclusion.

The items were averaged to yield a single attitude score, with an overall mean of 4.04, a standard deviation of 1.01, and a range from 1 to 7. The distribution is
There was a slight but significant gender difference in PRA scores, $F(1, 559) = 3.75, p < .053$. Men were more punitive in their PRA scores ($M = 3.96, SD = 1.02$) than were women ($M = 4.12, SD = 1.00$), although the effect size was trivial ($d = .014$; Cortina & Nouri, 2000).

A principal axis factor analysis was conducted to see if the expected single-factor solution emerged. The criteria for retaining factors were eigenvalues greater than 1, the scree test, and factor loadings greater than .40. Two factors with eigenvalues greater than 1 appeared. The first factor had an eigenvalue of 

<table>
<thead>
<tr>
<th>Item</th>
<th>College sample</th>
<th>Internet sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Prisoners should be paid at least minimum wage for their labor.</td>
<td>.46</td>
<td>.70</td>
</tr>
<tr>
<td>2. Prisons should be much tougher for inmates. (R)</td>
<td>.43</td>
<td>.71</td>
</tr>
<tr>
<td>3. More money should be spent on alternatives to prisons.</td>
<td>.51</td>
<td>.65</td>
</tr>
<tr>
<td>4. Society has things to do that are much more important than guarding prisoners’ civil rights. (R)</td>
<td>.49</td>
<td>.77</td>
</tr>
<tr>
<td>5. Prisons should return to “chain gangs” and other forms of hard labor. (R)</td>
<td>.67</td>
<td>.75</td>
</tr>
<tr>
<td>6. Prisoners should have better access to mental health care.</td>
<td>.63</td>
<td>.67</td>
</tr>
<tr>
<td>7. Prisoners’ leisure activities should be severely regulated.</td>
<td>.43</td>
<td>.70</td>
</tr>
<tr>
<td>(R)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Prisons and prison guards should be closely monitored by civil rights organizations.</td>
<td>.57</td>
<td>.61</td>
</tr>
<tr>
<td>9. Prisons should have better libraries.</td>
<td>.52</td>
<td>.59</td>
</tr>
<tr>
<td>10. The public should be more concerned with what happens to prisoners.</td>
<td>.55</td>
<td>.81</td>
</tr>
<tr>
<td>11. Prisoners should be forced to smash rocks and wear striped uniforms like in the old days. (R)</td>
<td>.54</td>
<td>.65</td>
</tr>
<tr>
<td>12. Society needs to remember that prisoners also have civil rights.</td>
<td>.77</td>
<td>.80</td>
</tr>
</tbody>
</table>

*Note.* (R) = reverse-scored item. Higher scores reflect positive attitudes toward prison reform. The college sample was composed of 561 students from the University of Kansas; the Internet sample was composed of 142 respondents to an Internet-based survey.
4.3, which accounted for 35.8% of the variance. All 12 items loaded greater than .40 on this factor. The second factor had an eigenvalue of 1.3 and consisted of a single item (Item 5), which had a loading of .46. This item loaded higher on the first factor, and the scree test suggested the second factor was unimportant, so only the first factor was retained in subsequent analyses. The factor loadings are shown in Table 1.

Sample 2: College student replication sample. To see if the factor structure would replicate, a second sample of 228 KU undergraduates (103 men, 125 women) completed the PRA scale during a large mass-testing session. A principal axis factor analysis found only a single factor with an eigenvalue of 4.09 accounting for 34.1% of the variance. Alpha for the 12 items was .82; no item would have increased alpha upon exclusion. The items were averaged to yield a single attitude score, with an overall mean of 3.93, a standard deviation of 0.97, and a range from 1 to 6.25. A significant gender difference was found, \( F(1, 266) = 7.84, p < .006 \); the effect size was small \((d = .05)\). Women had more progressive prison reform attitudes \((M = 4.09, SD = 0.88)\) than did men \((M = 3.73, SD = 1.05)\).

Sample 3: General population. To extend the validation process beyond college populations, 142 adults (102 women, 40 men) completed the PRA scale
A principal axis factor analysis found two factors with eigenvalues over 1. The first factor had an eigenvalue of 6.37 and explained 53.1% of the variance; the second factor had an eigenvalue of 1.06 and explained 8.8% of the variance. The second factor, however, clearly failed the scree test and was defined by a single item (Item 5). Because this item had a higher loading on the first factor, the second factor was ignored. Factor loadings for this sample are shown in Table 1. Alpha for the 12 items was .92; no item would have increased alpha upon exclusion. The average of the 12 items had an overall mean of 4.46 ($SD = 1.45$, $Mdn = 4.38$) and a range from 1.5 to 7. No gender difference appeared in this sample (Mann-Whitney $Z < 1$).

Convergent, Discriminant, and Known-Groups Validity

The PRA scale appears to have a reliable factor structure and strong internal consistency. Its correlations with other variables were thus explored.

Political ideology. A new sample of 64 adults completed the PRA scale and several measures of political ideology through an Internet-based survey. People first rated their political beliefs on two separate scales: “How conservative are your political beliefs?” answered on a 7-point scale ranging from 1 (not at all conservative) to 7 (definitely conservative); and “How liberal are your political beliefs?” answered on a 7-point scale ranging from 1 (not at all liberal) to 7 (definitely liberal). Then, people noted their political party affiliation using the options Democrat, Republican, or Other.

As Table 2 shows, prison reform attitudes were positively correlated with liberal political beliefs ($r = .39$, $p < .001$) and negatively correlated with conservative political beliefs ($r = -.26$, $p < .04$). A Kruskal-Wallis test found that people with different party affiliations had significantly different prison reform attitudes, $\chi^2(2, N = 64) = 5.47$, $p < .06$. Democrats had more positive attitudes than did Republicans (Mann-Whitney $Z = 2.03$, $p < .04$); persons checking “Other” also had more positive attitudes than did Republicans (Mann-Whitney $Z = 2.07$, $p < .04$). Persons marking “Democrat” and “Other” did not differ in their attitudes ($Z < 1$).

A sample of 63 KU undergraduates (28 women, 35 men) completed Altemeyer’s (1981) Right-Wing Authoritarianism scale and the PRA scale. As

3 The Internet-based samples were gathered through a combination of convenience sampling and snowball sampling. The survey was advertised through a link on the American Psychological Society’s web page of links to Internet studies (http://psych.hanover.edu/aps/exponnet.html). People involved in prisoner advocacy were also informed of the survey and were asked to inform others. Several prisoner advocacy websites placed links to the survey. The data were collected and screened following Birnbaum’s (2001) recommendations. The date, time, and remote address of the data submission were collected along with the survey data, enabling a check of possible multiple submissions. All responses from persons under the age of 18 and from persons residing outside the United States were excluded.
expected, authoritarianism was negatively correlated with prison reform attitudes ($r = -.39, p < .002$).

*Perceived similarity to prisoners.* Interpersonal similarity is a powerful predictor of liking (Byrne, 1971). Perceived similarity to prisoners should thus predict positive attitudes toward prison reform. A 3-item scale was created to measure perceived similarity to prisoners: “I probably have a lot in common with prisoners,” “Prisoners are basically different from me,” (reverse scored); and “Prisoners probably like many of the same things that I like.”

People responded to each item on a 7-point scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). A new sample of 41 adults (28 women, 13 men) completed the PRA scale and the similarity scale over the Internet. Scores from the three-item similarity scale had a reliability of .79 ($M = 3.92, SD = 1.52$). As expected, prison reform attitudes and perceived similarity to prisoners were highly correlated ($r = .67, p < .001$).

*Other variables.* A sample of 220 KU undergraduates (Sample 2 described earlier) completed brief neuroticism and extraversion scales (Eysenck & Eysenck, 1964) and a measure of political trust. The Incumbent-Based Trust scale (Craig, Niemi, & Silver, 1990) is a four-item Likert-type scale measuring trust in politicians and elected officials (e.g., “You can generally trust the people who run our government to do what is right”). This scale was included to see if the PRA scale tapped general political values or trust in government. As Table 2 shows, PRA scores did not correlate with any of the three variables.

People in the large general sample (Sample 3, described earlier) were asked to indicate their age and their educational level. Age ranged from 18 to 57 years

---

**Table 2**

*Correlations Between the Prison Reform Attitudes Scale and Other Variables*

<table>
<thead>
<tr>
<th>Variable</th>
<th>$r$</th>
<th>$p$</th>
<th>$n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceived similarity to prisoners</td>
<td>.67</td>
<td>.001</td>
<td>41</td>
</tr>
<tr>
<td>Political liberalism</td>
<td>.39</td>
<td>.001</td>
<td>64</td>
</tr>
<tr>
<td>Political conservatism</td>
<td>-.26</td>
<td>.042</td>
<td>64</td>
</tr>
<tr>
<td>Right-wing authoritarianism</td>
<td>-.39</td>
<td>.002</td>
<td>63</td>
</tr>
<tr>
<td>Educational level</td>
<td>.19</td>
<td>.022</td>
<td>142</td>
</tr>
<tr>
<td>Age</td>
<td>.17</td>
<td>.051</td>
<td>142</td>
</tr>
<tr>
<td>Neuroticism</td>
<td>.04</td>
<td>ns</td>
<td>220</td>
</tr>
<tr>
<td>Extraversion</td>
<td>-.01</td>
<td>ns</td>
<td>220</td>
</tr>
<tr>
<td>Trust in politicians</td>
<td>-.08</td>
<td>ns</td>
<td>220</td>
</tr>
</tbody>
</table>
(\(M = 29.0\) years, \(SD = 10.5, Mdn = 26.5\)), while educational level ranged from high school diploma and fewer years to doctorate and professional degrees (\(M = 14.9\) years, \(SD = 2.4, Mdn = 14.0\)). Age significantly predicted prison reform attitudes (\(r = .17, p < .05\)), and educational level predicted positive reform attitudes as well (\(r = .19, p < .02\)).

**Known-groups validity.** To pursue known-groups validity (DeVellis, 1991), I compared the attitudes of members of the general population with attitudes of persons involved in prisoner advocacy activities. People working to bring about progressive reforms to the prison system comprise a known group with positive reform attitudes. Data from the large Internet sample (Sample 3) were sorted based on whether people indicated any involvement in any “prisoner rights or prisoner advocacy organizations.” Eleven people indicated involvement. This sample had highly positive attitudes: The median attitude score was 6.83 out of a possible 7. A nonparametric rank test (which is more robust to unequal cell ns; Zimmerman & Zumbo, 1993) found that this group had significantly more positive attitudes toward prison reform than did the rest of the sample (Mann-Whitney \(Z = 3.89, p < .001\)).

A similar analysis was conducted for 21 people who responded “Yes” to the question “Have you ever spent time in jail or prison?” Former prisoners did not have significantly different attitudes from the rest of the sample (Mann-Whitney \(Z = 1.18, p < .24\)). When the persons in prisoner advocacy groups (i.e., the people at the ceiling of the general sample) were excluded, then the difference between former prisoners and nonprisoners became significant (Mann-Whitney \(Z = 1.94, p < .052\)).

**Discussion**

The PRA scale appears to have solid convergent and discriminant validity. The scale correlates in predicted and coherent ways with political ideology, authoritarianism, and perceived similarity to prisoners. Known-group differences were found: Democrats had more positive attitudes than did Republicans, former prisoners had more positive attitudes than did nonprisoners, and persons involved in prisoner advocacy had more positive attitudes than did persons not involved. Age and educational level were both associated with positive prison reform attitudes. The scale did not relate to broad personality dimensions (neuroticism and extraversion) or to general trust in government. Gender differences in prison reform attitudes were found in the college student samples, although the effect sizes were tiny. No gender differences appeared in any of the Internet samples.

**Study 2**

The PRA scale appears to have sound psychometric properties. To explore construct validity further, I examined how prison reform attitudes participate in
sentencing recommendations. People with positive prison reform attitudes should be more receptive to alternatives to prisons and more reluctant to sentence someone to a long prison term. In Study 2, people read about a crime and recommended a sentence for the convicted criminal. If the PRA scale adequately measures what it intends to measure, then (a) people with negative prison reform attitudes should recommend longer prison sentences than people with positive attitudes; and (b) people with negative reform attitudes should be less likely to recommend alternative, nonprison punishments.

Participants

A total of 49 undergraduates (27 men, 22 women) enrolled in Introductory Psychology at KU participated and received partial credit toward a research option. Participants were run in groups of 3 to 8.

Procedure

Participants entered the lab and completed a “background questionnaire” containing demographic items, the PRA scale, and some filler scales intended to distract attention from the PRA scale (ad hoc scales about attitudes toward university policies and ostensible personality items about dreams and creativity). Then they learned that the researchers were conducting a series of surveys on jury decisions and judgments, and that participants would read about a crime and give their impressions and recommendations. The crime involved Edward C., who had recently been convicted of two armed robberies in which no one had been injured. A brief description of the crimes and the trial procedures was provided.

After reading about the crimes and the conviction, participants completed a “recommendations sheet.” First, the sheet described the different sentencing options. The first option was a prison sentence, in which the convict would be required to spend a number of years in prison. The second option was community sentencing, in which the convict would be required to spend a number of years gainfully employed while making reparations to the crime victims, completing community service, receiving mental health care, and refraining from illegal activity (Davies, 1993). The first recommendation question asked “If you had a choice between sentencing Edward C. to a ‘prison sentence’ or sentencing him to a ‘community sentence,’ which would you recommend?” Participants then circled one of the options to express their recommendation. The second question asked “If the law required that Edward C. be sentenced to a prison term, how many years of prison would you recommend, based on your knowledge of the crime?” Participants wrote a number in a blank next to the question.

After completing the questionnaire, the experimenter described the study’s purposes in more detail, talked about past work on prison reform attitudes, and
gave participants a resource sheet with facts about the American correctional system and modern prison policies.

Results and Discussion

A logistic regression analysis was conducted to see if PRA attitudes predicted recommending a prison sentence versus community sentence. This analysis was significant ($b = 1.67, p < .002$), reflecting that community sentences were recommended more often as PRA attitudes became more positive. When people were asked to assume that a prison sentence was required, PRA scores significantly predicted the length of the recommended prison sentence. People with positive PRA attitudes recommended shorter sentences ($\beta = -.55, p < .001$). This effect remained significant after controlling for effects of gender ($\beta = -.57, p < .001$). This study thus offers additional construct validity for the PRA scale. Scores on this scale predicted sentencing recommendations in coherent ways, suggesting that the scale adequately measures attitudes regarding how the prison system should treat prisoners.

General Discussion

Even the vigorous field of forensic psychology has not kept up with the growth of the American prison system. To make inroads into understanding some aspects of prison psychology, I developed a scale to measure attitudes regarding how prisons should treat prisoners. Analyses of several samples demonstrate that the scale has a single factor and good internal consistency. Positive attitudes toward prison reform were associated with political liberalism, self-identifying as a Democrat, more years of education, feeling similar to prisoners, involvement in prisoner advocacy groups, and spending time in jails or prisons. Negative attitudes were associated with political conservatism, self-identifying as a Republican, and authoritarianism. Small gender differences were found only in college student samples; no gender differences appeared in Web-based surveys of the general public.

The PRA scale measures attitudes toward making prisons tougher or softer. It thus assesses the desired severity of the prison experience, and not the reasons for desiring incarceration. People wish to punish others for many reasons: retribution, principled notions of justice, deterring future transgressions, and so forth (Robinson & Darley, 1995). Specific to prisons, the public wants to incarcerate criminals for purposes of rehabilitation, isolation from the community, retribution, and deterring potential criminals (Flanagan, 1996). The PRA scale does not try to measure the goal or motivation behind incarceration. How different goals for imprisonment relate to the desire to make prisons tougher or softer is an empirical issue. To avoid building in such relationships, the scale excludes items
related to the reasons for punishment, the efficacy of rehabilitation efforts, and whether prisoners (versus society) are responsible for their crimes. Indeed, exploring lay concepts of the goals and functions of imprisonment is an important direction for future research on prison reform attitudes.

Prison reform attitudes might offer a glimpse into aspects of prejudice that are difficult to study. Research shows that much prejudice is “underground”: Social norms have made expressing certain prejudices unacceptable. While one might privately dislike Jews, African Americans, or persons with disabilities, it is no longer “okay” to voice such attitudes publicly. These strong social desirability forces have hindered studying certain facets of prejudice (Biernat & Crandall, 1999). Some researchers have thus turned to other prejudices that have yet to be socially circumscribed, such as prejudice based on weight and appearance. Prisoners are a good group for these purposes because it is perfectly acceptable to express negative feelings toward prisoners. As the members of society who have failed to meet society’s rules, prisoners are rejected to the furthest fringes, much as prisons themselves are relegated to desolate areas far from the general population. And if prejudice indeed hinges on whether a group appears deserving of poor treatment, prisoners are an interesting population because, unlike minority groups or persons with different physical appearances, their differential treatment is presumably based on their intentional actions (Feather, 1999).

This scale could enable researchers to assess more precisely the public’s attitudes toward prison reform. Flanagan (1996) noted that politicians typically assume a fearful, vindictive public that wants its elected representatives to crack down on “country-club prisons.” This motivates political leaders to push for policies that reform prisons in tougher directions, such as forced labor and Supermax prisons. Yet, if the public is not so punitive, as Flanagan’s research suggested, then there may be no political necessity behind these measures. This could prevent the implementation of occasionally extreme practices and, at the very least, promote scientifically informed policy decisions.

References


