September 2011

Federal PROBATION

a journal of correctional philosophy and practice

Volume 75 Number 2

SPECIAL ISSUE: Evidence-based Practices in Action

Published by the Administrative Office of the United States Courts www.uscourts.gov
Publishing Information

Volume 75 Number 2

Federal PROBATION
a journal of correctional philosophy and practice

A Random (Almost) Study of Staff Training Aimed at Reducing Re-arrest (STARR): Reducing Recidivism through Intentional Design

Charles R. Robinson[*]
Scott VanBenschoten[**]
Melissa Alexander[†]
Christopher T. Lowenkamp[††]

COMMUNITY SUPERVISION IS one of the most widely imposed court responses, with approximately 5,095,200 or 70 percent of the correctional population being under community supervision (Glaze, 2010). Despite its popularity, researchers have limited insight into whether community supervision is an effective strategy for reducing recidivism. The most recent reviews of the effectiveness of community supervision (Solomon, Kachnowski, & Bhati, 2005; Aos, Miller, & Drake, 2006; Bonta, Rugge, Scott, Bourgon, & Yessine, 2008; Green & Wink, 2010) brought sobering results. More than three decades after Martinson (1974) summarized the findings of his review of rehabilitation efforts by saying "with few and isolated exceptions, the rehabilitative efforts that have been reported so far have had no appreciable effect on recidivism," Bonta and his colleagues (2008) have found that more recent research yields no better results: the impact of community supervision is limited at best and non-existent in the most pessimistic interpretation.

The Urban Institute (Solomon, et al., 2005) reached a similar conclusion after analyzing the impact of post-prison supervision on re-arrest outcomes. The researchers, after comparing mandatory parolees with similar prisoners released without supervision, concluded that "overall, parole supervision has little effect on re-arrest rates of released prisoners." Consistent with the findings of the Urban Institute, Green & Wink (2010) declared "… probation does not alter the probability of recidivism" after tracking more than 1,000 offenders randomly assigned to nine judicial calendars. Simply put, the recidivism rate of those placed on probation was no different from that of those who weren't placed on probation. Taxman (2002) further affirms the notion of ineffectiveness after reviewing studies on intensive supervision and caseload size. Taxman states

that "unless the contacts are more than check-ins it is unlikely that they will impact outcomes."

Collectively, these reviews indicate that community supervision has little to no impact on the likelihood of future crime. However, none of these studies examined exactly what occurs in meetings between officers and those on supervision. To answer this question, researchers and practitioners have started to unpack the "black box" of supervision (Bonta et al., 2008) to determine what might impact client outcomes. Results suggest that the core of community supervision must be built on a foundation that targets those at highest risk of engaging in criminal behavior, that the areas targeted are those closely linked to future criminal behavior, that barriers to treatment must be removed, and that cognitive-behavioral strategies must be utilized (Andrews, Zinger, Hoge, Bonta, Gendreau, & Cullen, 1990). Likewise, research supports the notion that the quality and nature of the relationship between the client and the supervision officer has an impact on outcomes (Skeem et al., 2007; Paparozzi and Gendreau, 2005). Finally, Jalbert and colleagues (2011) recently completed an analysis of the impact of caseload size on supervision outcomes, and concluded that caseload size can improve outcomes only if used in combination with the effective controlling and correctional strategies described above.

Moving from a "check-in" to the use of a core skill set to increase effectiveness is supported by Andrews and Kiessling (1980), Dowden and Andrews (2004), Trotter (1999), and Taxman (2008). Andrews and Kiessling (1980) introduced the five dimensions of effective correctional practice that were designed to enhance the potential of rehabilitation programs for offenders. Dowden and Andrews (2004) provided a meta-analytic review of the core correctional practices indicating that the use of authority, disapproval, reinforcement, modeling, teaching problem solving skills, and structured learning are all related to the effectiveness of correctional services. While much of the research reviewed by Dowden and Andrews focused on treatment programs, the Bonta et al. (2010) findings are consistent with other research focused on testing the use of these skills in a community supervision setting (Trotter, 1996 & 1999; Taxman et al., 2006). Bonta et al. (2010) affirms the relationship between specific core correctional skills and the effectiveness of community supervision officers, noting that those officers trained in the skills utilized the skills more often and clients of officers. trained in the skills had lower recidivism rates than those of untrained officers.

The existing research (Bonta et al., 2008; Taxman et al., 2006; and Trotter, 1996) is encouraging and points to a need for further research on the training of community supervision officers providing direct service to clients. The current study uses a larger sample size than previous studies to further investigate the application of techniques that influence change and provide a model for targeting dynamic risk factors. The current study also uses random assignment to control for the selection and assignment biases associated with observational studies and the use of volunteer participants.

back to top

Method

This study used an experimental pre-post test design. All officers who volunteered for the training and study were randomly assigned to the experimental (trained) and control (untrained) groups.[1] Random assignment was completed such that 66 percent of the officers were randomly assigned to the experimental condition and the remainder assigned to the control condition. While officers were, for the most part, randomly assigned to the two conditions in this study, the clients were not. However, as is the case with most community supervision agencies, once geography was factored in, the client assignment process was based on rotation, caseload size, or some other factor unrelated to officer status in this study.

Participants

Officers

The study began with 53 officers in the experimental group and 35 in the control group. Immediately following the training, 6 officers were lost from one district (both experimental and control group) due to a lack of desire to participate. This reduced the numbers to 49 and 33. An additional group of officers (8 from the experimental group and 7 from the control group) were lost because some officers received promotions, left their position with the judiciary, dropped

out of the study, or did not have moderate- and high-risk clients both pre- & post-training. This reduced the total number of officers in the experimental group to 41 and the total number of officers in the control group to 26. The attrition rate of 18 percent, while of some concern, retained the original distribution of officers across the two groups.

Clients

Clients were identified for inclusion in this study based on when their period of supervision began. While clients were not randomly assigned to experimental and control group officers, clients were assigned to officers independent of the officers' STARR training status. The pretrial sample is made up of cases that were assigned to pretrial supervision. Pre-training cases were those cases that began pretrial supervision during 2007 and 2008 and terminated supervision before the training event date. Post-training pretrial cases were those cases that were assigned to study officers after May 31, 2009. Post-conviction pre-training cases were identified as those cases that began their supervision between May 31, 2007 and May 31, 2008, as this allowed for a follow-up time of at least 12 months. Post-conviction post-training cases were those cases assigned for supervision after May 31, 2009 up until December 12, 2009.[2]

In all there were 345 pre-training cases assigned to control officers and 446 pre-training cases assigned to the experimental officers. A total of 218 post-training cases were assigned to the control officers and 295 post-training cases were assigned to the experimental officers. A breakdown of these cases by pretrial versus post-conviction supervision is provided in Table 1. As indicated in Table 1, a much larger number of the clients included in this study are those on post-conviction supervision.

The demographic statistics of the clients included in this study are presented in Table 2. Fifteen percent of the sample is female and 57 percent belongs to a minority race or ethnicity. Half of the clients were moderate-risk and half were high-risk according to the RPI.[3] The average age for the clients included in this study is approximately 35. Table 2 also shows the descriptive statistics by group (experimental versus control). None of the observed differences were significant at the $p < .05$ level. Analysis of demographic characteristics by pre- and post-status and pre-post status by group (pre-training control group, post-training control, pre-training experimental, and post-training experimental) revealed no statistically significant differences.

back to top

Procedures

This training was intentionally designed to be responsive to the literature on technology transfer[4] and the use of one-on-one officer-client interactions to reduce risk and thereby client recidivism. Officers in the experimental group participated in a 3½ day classroom training that included a discussion of the theory supporting the development of the STARR curriculum, a demonstration of each skill, exercises, and an opportunity for officers to practice each skill and receive feedback. The theory discussion reviewed the risk, need, responsivity model (see Andrews & Bonta, 2003) and the research demonstrating the effectiveness of a skill-focused supervision approach. The STARR skills themselves include specific strategies for Active Listening, Role Clarification, Effective Use of Authority, Effective Disapproval, Effective Reinforcement, Effective Punishment, Problem Solving, and Teaching, Applying, and Reviewing the Cognitive Model. For each strategy skill cards were developed that outline the specific activities officers needed to do to successfully deliver the strategy. A fundamental focus for each skill is the internalization of strategies so that defendants/offenders begin to learn and apply the strategies on their own. In addition to the skill cards, video examples of some skills were presented, while others were demonstrated live. The exercises allowed officers to practice each skill. For example, after listening to a discussion about reinforcement, officers were asked to identify a behavior and a reinforcement strategy for a specific offender, then role play that interaction with another officer. The officers (experimental and control) were asked to send in audiotaped interactions (1 before the training event and up to 30 after the training event) at designated intervals: initial meeting with the client, an interaction with the client 3 months later, and then a third and final taping 3 months after that (6 month interaction). Officers made recordings with up to 10 moderate- or high-risk clients. The audiotapes were used to gain a better understanding of skill development and provide feedback to the officers. Four "booster"

trainings[5] were held over the next year to provide officers with additional training on skill deficits identified on the tapes. Booster trainings were delivered by phone and included discussion of specific skills, audiotape examples of the skill, and individual feedback and coaching.

back to top

Measures

Intermediate officer measures

Officer's use of the skills taught during the training was measured by reviewing audiotape recordings of interactions with clients recruited for the study. In all 731 audio recordings were submitted for review. This included 491 recordings from the experimental group and 240 from the control group. The audiotapes were coded by trained raters who focused primarily on behaviors consistent with the skills introduced during the training. For example, with reinforcement or disapproval, raters coded whether the officer identified the specific behavior and whether the officer had the offender explore the short- and long-term consequences of the behavior. For the cognitive model, raters coded whether the model was taught, applied, or reviewed, and coded for discussions of internal cues, consequences of internal cues, and identification of counter thoughts. Finally, raters coded what topics were discussed in the interaction.

back to top

Client Outcomes

Two different outcome measures were used in this study. For pretrial clients, the outcome measure was failure on supervision, as evidenced by failure to appear in court, supervision being revoked, or being arrested for a new criminal charge while on pretrial supervision. The data for this measure was taken from PACTS.[6] The outcome measure for the post-conviction cases was arrest for new criminal behavior, as identified in the NCIC or ATLAS databases.[7]

For clients on post-conviction supervision, the follow-up time was standardized to 12 months. For pretrial clients the time period was limited to the time they were on pretrial release. For the pre-training group, the average time on pretrial release was 229 days and the post-training average time on pretrial was 185 days (observed differences between experimental and control groups and across time periods within groups were not significant).

Analysis

Bivariate analyses were used to assess the change in officer behavior from pre- to post-training and across the experimental and control groups and to assess the impact of the training on client outcomes. Multivariate analyses were used to determine the interaction between individual client characteristics and officer training and their impact on client outcomes.

back to top

Results and Discussion

Intermediate Outcomes

The 88[8] officers participating in the study submitted 598 audio recordings for review. This included 400 from the experimental group and 198 from the control group. The audiotapes were coded by trained raters using a structured guide primarily focused on behaviors supported by core correctional practices. This analysis focuses on three intermediate variables: a) the officer's use of reinforcement and disapproval, b) interactions where cognitions, peers, or coping skills were discussed, and c) the officer's use of cognitive techniques during interactions with clients. Analysis of pre-training audiotapes showed no difference between the experimental and control groups in the use of these skills.

As seen in Table 3, officers in the experimental group used reinforcement and disapproval at nearly twice the rate of untrained officers. This suggests that trained officers were almost twice

as likely to capitalize on opportunities to use behavioral strategies that help shape client behavior. The finding provides cause for optimism because of the demonstrated impact of operant conditioning techniques like reinforcement and disapproval (see Dowden and Andrews 2004).

Cognitions, peers, and impulsivity empirically represent some of the strongest predictors of future criminal behavior. A primary focus of STARR is addressing dynamic risk factors using a structured cognitive-behavioral approach. Table 4 shows post-training interactions where cognitions, peers, and impulsivity were discussed. As seen in Table 4, discussions about cognitions, peers, and impulsivity were significantly more likely to occur among officers in the experimental group than among officers of the control group (44 percent vs. 30 percent). This represents a significant difference in how often primary risk factors are targeted.

In addition to targeting dynamic risk factors and using operant conditioning techniques, analysis suggests significant differences in officers' use of the cognitive model. As seen in Table 5, control group officers used the cognitive techniques in 1 percent of interactions where the skills were applicable compared to 17 percent by experimental group officers. This suggests that experimental group officers were significantly more likely to use the cognitive techniques to teach offenders the link between thinking and behavior.

back to top

Client Outcome

Our first analysis aimed at identifying the difference in failure rates for those clients assigned to two groups of officers prior to the training. The first panel of Table 6 indicates that there was no difference in client failure rates between the groups prior to STARR training; control group officers had a failure rate of 38 percent, whereas clients assigned to the experimental officers had a 39 percent failure rate. The second panel of Table 6 displays the post-training failure rates of the clients based on group assignment. The control group cases had a post-training failure rate of 34 percent, which did not significantly differ from the pre-training failure rate. The post-training failure rate for the clients assigned to experimental group officers was 26 percent, which is significantly lower than the pre-training experimental failure rate, and more importantly, than the post-training failure rate for the control group (see note 2 of Table 6).

Taken together, these results indicate that training can significantly impact strategies used by officers during supervision, and that these strategies lead to lower failure rates. The overall difference in failure rates between the trained and untrained groups is nine percentage points, which equates to a reduction in relative-risk of approximately 25 percent and is quite remarkable, as these differences were achieved with approximately 40 total hours of training, no reductions in caseloads, and no additional work hours from officers.

The next set of analyses focused on investigating the impacts of group membership on client outcomes by risk, specifically whether changes and differences in failure rates differed drastically between moderate- and high-risk clients. Tables 7 and 8 present the failure rates by group for moderate- and high-risk clients.

Table 7 presents the failure rates by group for the moderate-risk clients. In the first panel note that the failure rates for the control and experimental groups pre-training were again statistically indistinguishable from one another (31 percent and 32 percent respectively). The failure rates for the control group pre- and post-training and the experimental group pre-training failure rate also do not differ significantly from one another. Post-training failure rates between the groups indicate a very different trend. The post-training failure rate for the control group is 32 percent, while the post-training failure rate for the experimental group is 16 percent. This is an absolute reduction of 16 percent and a relative risk reduction of 50 percent. This again is noteworthy given the amount of training the officers were given and the fact that the experimental group officers had no additional resources or reduced caseloads.

The failure rates by group for the high-risk clients are contained in Table 8. The first panel of Table 8 indicates, once again, that the pre-training failure rates across the two groups did not differ significantly from one another. The second panel of Table 8, which displays the post-training failure rates across the two groups, also indicates no difference. An important trend that

occurs for both groups is the decrease in failure rates from pre- to post-training. This is likely the result of other efforts that had been ongoing in the districts selected for this study.[9] At any rate, it doesn't appear that the STARR skills, in this context, produced any beneficial results over and above the targeted efforts of the officers in the study.

The bivariate analyses indicated that the two groups of clients did not differ in terms of race, age, gender, or time at risk. Intra-group pre-post differences in race, age, gender, and time at risk were determined to be statistically not significant. Even so, since clients are not necessarily randomly assigned to officers, we felt it was important to construct and estimate a series of multivariate logistic regression models predicting client failure. The three models are presented in Table 9. The only difference across the three models is the addition of interaction terms. More specifically, Model 2 includes an interaction term between group membership and pre-post time period which isolates the effects for clients assigned to experimental officers after the training. Model 3 includes an interaction term between group membership, pre-post time period, and high-risk status. Thus the additional interaction term allows us to better understand the effects for moderate- and high-risk clients net the effects of other variables in the model.

The three multivariate models presented in Table 9 all predict the same outcome and use the same set of control variables. The difference between the three models is the addition of the interaction terms as described above. In each of the three models the parameter estimate for minority status is not significant, meaning that once the other factors are controlled for minority status is not associated with failure. In all three models female is inversely associated with failure, age is inversely associated with failure, and high-risk status is positively associated with failure. In each of the three models the parameter estimate for group, which captured whether an offender was assigned to a control or experimental group officer, was not significant.

The parameter estimate for the pre-post period (coded as 1 for post-training period) is significant and negative in Model 1. Once, however, the interaction term between group membership and pre-post period is introduced, the parameter estimate for the pre-post variable is no longer significant. This indicates that what was driving the effect of the pre-post period in Model 1 is attributable to the reductions in the post-training experimental group. Model 3 introduces an interaction term that quantifies the impact of high-risk clients in the post-training experimental group. As indicated, compared to moderate-risk clients, high-risk clients in the post-training experimental group are slightly more likely to fail; however, this parameter estimate does not quite reach the typically accepted probability values associated with statistically significant findings.

In converting the log-odds ratios into probabilities, we are able to develop a better understanding of a particular variable while holding the other factors constant. In doing so, we estimated the probability of failure for a 35-year-old, minority, moderate-risk male who was on post-conviction supervision and in the control group prior to training to be at .35. A 35-year-old, minority, moderate-risk male that was in the experimental group prior to training has an estimated probability of failure at .36. A case with those same characteristics that was on an experimental officer's caseload after training had a probability of failure at .18, while a case with the same characteristics on a control officer's caseload after training had a .31 probability of failure. One can quickly see from these calculated numbers that the cases on the experimental officers' caseload had about half the failure rate. This decrease is not as large for high-risk cases; however, it is still in the expected direction and the parameter estimate that captures the post-training difference in effect for the high-risk cases was positive (indicating an increase in the probability of failure and thereby a decrease in effect) but not statistically significant. A visual display of the predicted probabilities of re-arrest by risk and group assignment is provided in Figure 1 below.

back to top

CONCLUSION

This study used an experimental design to assign officers to an experimental and control condition focused on training officers in specific strategies to use during direct supervision of defendants/offenders. Clients supervised by these officers both before and after the training were used to assess the impact of the training on officer behaviors and client outcomes. Results

indicate that officers utilized effective strategies more often post training, and that client outcomes were impacted. Clients supervised by the experimental group of officers after the training had far superior outcomes, even after controlling for individual client level characteristics. A sub-analysis indicated that the effects of the experimental condition were not present for high-risk clients. Nonetheless, this study, a fairly rigorous test, indicated that training in STARR is associated with 50 percent reductions in 12-month failure rates for moderate-risk clients.

There are a number of limitations to this study. First, the clients were not randomly assigned to the officers. The various districts involved in this study confirmed that client assignment was based on a combination of geography, case specialty (substance abuse specialists received a higher percentage of these types of clients), and attempts to balance caseload sizes. However, while the clients do not appear to differ in terms of individual level characteristics across the groups or pre-post time periods, it could be the case that some unmeasured bias in assignment exists. Second, a full analysis of the intermediate measures and their relation to client outcomes has not yet been conducted. These analyses might lead to important adaptations in the training and implementation. Third, while there was an attempt to minimize attrition, we did end up losing just under 20 percent of the officers that started this study. Fourth, the study was conducted by those that developed the training program and trained the officers. Given previous findings on the relationship between demonstration studies and real-world applications, independent and external replications of this study will be important.

Notwithstanding, there are some very important policy implications to take from this study. First, it appears that officers can be trained in behaviorally based skills and they can use those in their one-on-one interactions with offenders. The data presented on the intermediate measures provides support for this conclusion. More important, the bivariate and multivariate analyses of outcome measures supports the conclusion that training probation officers in behavioral strategies might have a profound effect on failure rates of clients in the correctional system. Subsequent research should focus on continuing to test this conclusion.

back to top

Volume 75 Number 2

Federal PROBATION
a journal of correctional philosophy and practice

A Random (Almost) Study of Staff Training Aimed at Reducing Re-arrest (STARR): Reducing Recidivism through Intentional Design

Tables

Table 1.

Distribution of pre- and post-training cases by supervision type

	Pre Training		Post Training	
	Experimental	Control	Experimental	Control
Pretrial	52	56	30	26
Post conviction	394	289	265	192

Table 2.

Descriptive statistics for demographic characteristics of clients

	All (1304)	Experimental (747)	Control (547)
Variable	Percent or Mean	Percent or Mean	Percent or Mean
Female	15%	15%	15%
Minority	57%	57%	57%
Percent moderate-risk	50%	48%	53%
Percent high-risk	50%	52%	47%
Age	34.78	34.95	34.55
RPI Score	5.61	5.69	5.49

Table 3.

Post training use of reinforcement and disapproval by group

	Did Not Use Skill		Used Skill	
	N	%	N	%
Experimental	230	66%	117	34%
Control	139	83%	29	17%

$\chi 2(1) = 15.090$; $p = 0.05$; $\eta = 0.171$

Table 4.

Post-training percent of interactions where cognitions, peers, or impulsivity were discussed

	Not Discussed		Discussed	
	N	%	N	%
Experimental	196	56%	151	44%
Control	117	70%	51	30%

$\chi 2(1) = 8.222$; $p \leq 0.05$; $\eta = 0.126$

Table 5.

Post-training percent of interactions where cognitive model was used

	Not Discussed		Discussed	
	N	%	N	%
Experimental	289	83%	58	17%
Control	167	99%	1	1%

$\chi 2(1) = 28.995$; $p \leq 0.05$; $\eta = 0.237$

Table 6.

Failure rates for moderate and high-risk clients by group membership

	Failure Rate	
	Pre-training[1]	**Post-training[2]**
Control Officers	130/345 (38%)	75/218 (34%)
Experimental Officers	175/446 (39%)	78/295 (26%)

[1] χ^2 (1)= 0.199; $p > 0.05$
[2] χ^2 (1)= 3.798; $p \leq 0.05$

Table 7.

Failure rates for moderate-risk clients by group membership

	Failure Rate	
	Pre-training[1]	**Post-training[2]**
Control Group Officers	58/188 (31%)	36/113 (32%)
Experimental Group Officers	71/221 (32%)	22/135 (16%)

[1] χ^2 (1)= 0.782; $p > 0.05$
[2] χ^2 (1)= 8.314; $p \leq 0.05$

Table 8.

Failure rates for high-risk clients by group membership

	Failure Rate	
	Pre-training[1]	**Post-training[2]**
Control Group Officers	72/157 (46%)	39/105 (37%)
Experimental Group Officers	104/225 (46%)	56/160 (35%)

[1] χ^2 (1)= 0.005; $p > 0.05$
[2] χ^2(1)= 0.127; $p > 0.05$

Table 9.

Multivariate analyses of client outcomes

	Model 1			Model 2			Model 3		
Variable	**B**	**SE**	**P**	**B**	**SE**	**P**	**B**	**SE**	**P**
Minority	.088	.126	.485	.085	.127	.501	.088	.127	.486
Female	-.435	.182	.017	-.441	.182	.015	-.438	.182	.016
Age	-.040	.007	.000	-.040	.007	.000	-.040	.007	.000
High-risk	.553	.124	.000	.556	.125	.000	.467	.137	.001
Pre-post Period	-.423	.125	.001	-.142	.187	.448	-.139	.187	.455
Group	-.110	.121	.365	.760	.153	.620	.080	.152	.601
Pretrial	-.156	.192	.415	-.149	.192	.603	-.157	.192	.413
Group x Pre-Post	—	—	—	-.505	.251	.044	-.810	.326	.013
Group x Pre-Post X High Risk	—	—	—	—	—	—	.486	.319	.129
Constant	.719	.276	.009	.628	.281	.025	.672	.282	.017

Volume 75 Number 2

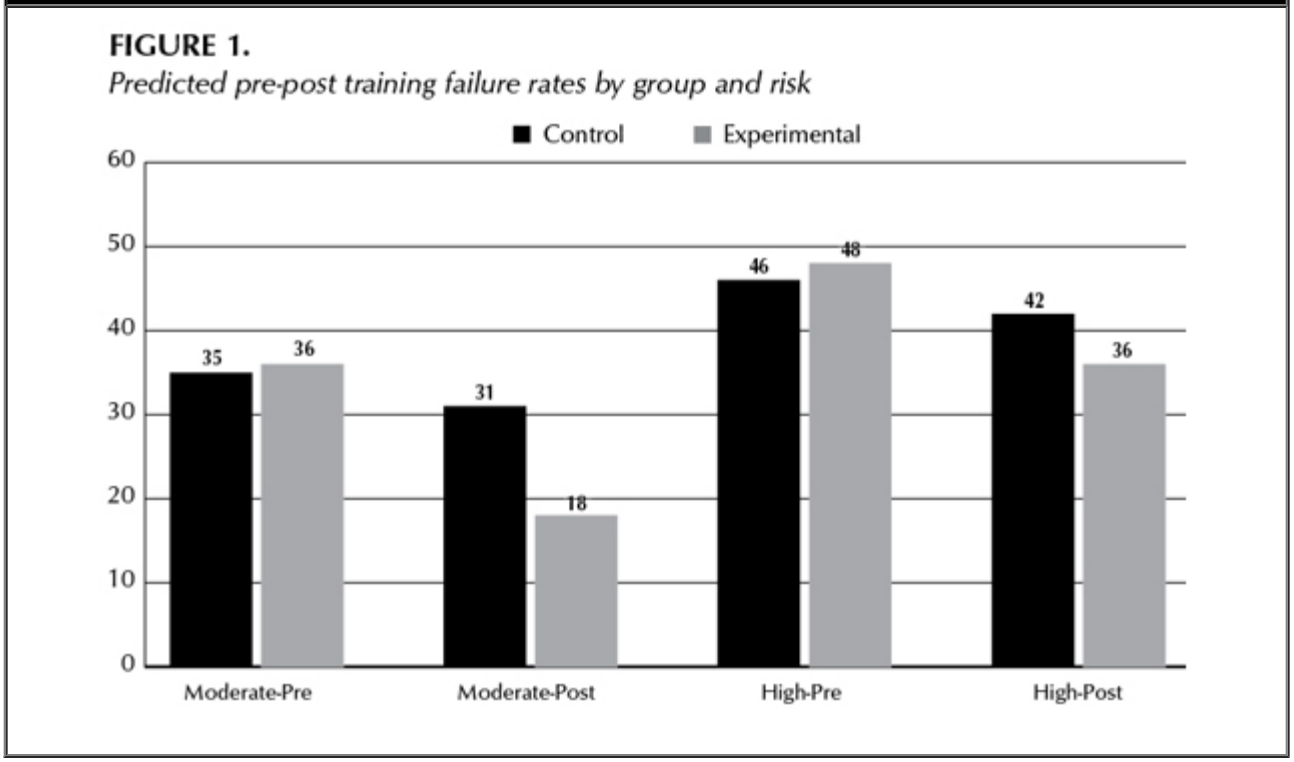
Federal PROBATION
a journal of correctional philosophy and practice

A Random (Almost) Study of Staff Training Aimed at Reducing Re-arrest (STARR): Reducing Recidivism through Intentional Design

Figures

Figure 1

Figure 1.



back to top