

# National Institute of Corrections Training Evaluation Project, 2005-2007

## *Training Results, Activity Level Changes, and Implementation Results*

By James B. Wells, Ph.D.  
Kevin I. Minor, Ph.D.  
and J. Stephen Parson

This is the third in a series of research bulletins on NIC's Training Evaluation Project. The project is being conducted by a team of researchers from Commonwealth Research Consulting (CwRC), in collaboration with NIC's Division of Research and Evaluation. The purpose of the project, and

the bulletin series, is to enhance understanding of training programs, and when appropriate, facilitate program improvements to better serve the field. The scope of the project, initially limited to Academy Division trainings during the pilot phase (2005-2006), was expanded in August 2007 via the current agreement with CwRC.<sup>1</sup> Several Jails Division trainings and Prisons Division trainings are currently being evaluated. Findings from these evaluations will be the subject of future bulletins.

## Highlights

- 4,140<sup>2</sup> surveys completed by 453 training participants in 20 trainings were analyzed for the bulletin. Response rates averaged 96.8% on initial surveys and 73.8% on follow-up surveys (Table 1).
- Three outcome areas were examined:
  1. **Training Results** (progress on training objectives)
  2. **Activity Levels** (pre and post-training behavior)
  3. **Implementation Results** (in the workplace)
- Despite considerable variation in outcomes among training participants (Figures 1b through 6b) findings are generally encouraging in all areas.
- **Training Results:** On average training participants reported moderate to substantial progress on training objectives as a result of attending training, with an overall mean progress rating of 3.76 (SD=.68, N=427) on a scale of 1 to 5 (Figures 1a and 1b).
- **Activity Levels:** Overall participants engaged in 52.8% (SD=.33, N=365) of key training-related behaviors in the 3 or 6 months<sup>3</sup> prior to training (Figures 2a and 2b), and 70.4% (SD=.29, N=255) of these behaviors in the 3 or 6 months after training (Figures 3a and 3b).
- **Implementation Results:** On average training participants reported moderate progress implementing both training-specific objectives (mean=2.86, SD=.74, N=294) and participant-identified objectives (mean=3.05, SD=.81, N=194) in their organizations (Figures 5a, 5b, 6a and 6b).
- Three potential sources or explanations of outcome variation were examined:
  1. **Pre-training measures:** participant demographics, personal qualities, previous training, etc.
  2. **Training measures:** training quality and relevance, trainer effectiveness, etc.
  3. **Post-training measures:** organizational resources and barriers<sup>4</sup> to implementing training objectives; follow-up period; etc.
- A variety of statistical analyses reveal that **pre-training measures**, such as demographic or personal characteristics of training participants, explain the least variation in all outcome areas (Figures 1c through 6c; Table 3).
- **Training measures**, such as overall training quality and relevance, explain significantly more outcome variation than pre-training measures, in all areas tested (Figures 1c, 3c, 5c and 6c; Table 3).
- **Post-training measures**, such as resources and barriers<sup>4</sup> in the organization, were the best predictors of outcome variation in all areas tested (Figures 3c, 5c and 6c; Table 3).

## Training Evaluation Project Primary Staff

### For Commonwealth Research Consulting, Inc:

**James B. Wells, Ph.D.**

President and Chief Research Consultant  
jbwells@cwrc.us

**Kevin I. Minor, Ph.D.**

Senior Research Consultant  
kiminor@cwrc.us

**J. Stephen Parson**

Research Consultant  
jsparson@cwrc.us

### For the National Institute of Corrections:

**Christopher A. Innes, Ph.D.**

Chief, Research and Evaluation  
cinnes@bop.gov

**Dee L. Halley**

Correctional Program Specialist and Project Monitor  
dhalley@bop.gov

## Acknowledgements

The National Institute of Corrections Training Evaluation Project is made possible by the support of NIC via Cooperative Agreements 05A28GJF9, 06PEI01GJM1, and 07PEI12GJQ7.

CwRC staff wish to acknowledge the support and cooperation of the many persons who helped make this project possible. Morris Thigpen, Larry Solomon, Tom Beauclair, Chris Innes, Dee Halley, Virginia Hutchinson, Bob Brown, John Eggers, Leslie LeMaster, Launa Kowalczyk and others at NIC have provided essential support for this project. We also wish to acknowledge our support staff, whose daily efforts further the project in so many ways. Finally, we want to express our appreciation to the growing number of NIC trainers and training participants who have taken time out of their busy schedules to graciously share their insights with us.

Although many persons and organizations contributed to the project described in this bulletin, any errors or omissions are those of the authors alone.

The findings and views presented in this bulletin are those of the authors and do not necessarily reflect the positions or policies of the National Institute of Corrections, or any other organization or individual.

Previous bulletins in the series, produced under cooperative agreement with Eastern Kentucky University's Center for Criminal Justice Education and Research, are available at [www.nicic.org/research](http://www.nicic.org/research).<sup>5</sup> Please refer to Bulletin 1: *Participant Demographics, Overall Evaluation of Training, and Applicability Ratings* (February 2007), for additional information on the Training Evaluation Project and the bulletin series.

The current bulletin presents findings from the analyses of data provided by 453 participants<sup>6</sup> in 20 trainings regarding **training results, activity level changes, and implementation results**. These outcome measures are discussed in greater detail under **Findings** (page 4).

The 20 trainings under evaluation took place during the pilot phase of the evaluation project (2005-2006) and include:

- Critical Elements of Re-entry (CER)
- Correctional Leadership Development-Adult (CLD-A)
- Correctional Leadership Development-Juvenile (CLD-J)
- Management Development for the Future (MDF)
- Training Design and Development (TDD)
- Juvenile Agency Training Directors (JATD)
- Meeting the Needs of Female Offenders (MNFO)
- New Juvenile Facility Directors (NFD)

As indicated in Table 1 (page 3), most of these training programs were evaluated multiple times. A variety of written survey instruments were administered to training participants in the course of each evaluation. Data analyzed for the current bulletin were drawn from portions of the five surveys listed in Table 1.

*Although some findings from these surveys were reported in previous bulletins, the focus was on **process measures**.*

*Here the focus is on **outcomes**, and **testing** the extent to which outcomes can accurately be **predicted** from process (and other) measures.*

Participant demographic forms were administered on the first day of each training. Overall training evaluation forms, trainer evaluation forms,<sup>7</sup> and current application forms<sup>8</sup> were administered on the last day of each training. Response rates for these four surveys averaged 96.8% (Table 1).

The fifth survey, current applications follow-up, was administered via U.S. mail either three or six months<sup>3</sup> after each training. Follow-up data collection concluded in May 2007, after the fourth mailing to participants of the MDF (06-R019) training program. The multi-step follow-up protocol, based on Dillman's Total Design Method,<sup>9</sup>

produced a 73.8% response rate (Table 1). Overall, a total of 4,140<sup>2</sup> completed surveys were analyzed for this bulletin.

### Limitations and Quality Assurance

Although the research team employed a variety of quality assurance measures in all phases of the Training Evaluation Project leading up to the current bulletin, certain limitations remain:

1. The original research design utilized in the pilot phase of the project called for individual training evaluations,<sup>10</sup> i.e., survey instruments and evaluation protocols were customized to each training. For example, the number and type of training objectives varied by training, as did the precise wording of survey items, the subject<sup>11</sup> of the items, response formats, time frames,<sup>3</sup> directions, etc. To the extent these differences could be resolved satisfactorily, data were combined and analyzed across trainings for purposes of the bulletin. Although the merging of divergent data sets is not an exact science, a variety of quality assurance measures revealed no random or systematic data corruption or distortion during the recoding and merging processes. (Non-combinable data either have been or will be analyzed and reported separately.)<sup>10</sup>
2. Some data are not available across all trainings due to: differences between trainings; insufficient time between

project start date and some training dates; participant non-response; survey design error; survey administration error; or failure to pass data validation or other quality control measures. However, the research team found no evidence that the extent of missing or unavailable data compromised the analyses or findings reported in the bulletin.

3. The data analyzed for the current bulletin were all self-reported by training participants. Generally speaking, self-report data are subject to self-censorship, inaccurate recall, inaccurate projection, deception, and other problems. However, the self-report data analyzed for the current bulletin were less likely to be problematic because, for example:
  - The several hundred survey items administered were non-sensitive and non-controversial, with the possible exception of some demographic variables, e.g. race, gender, and salary.
  - Memory recall was limited to specific behaviors, over relatively short periods of time, often marked by memorable events, e.g., the span of time between the training and the present moment.
  - Projections or estimates, e.g., of behaviors, progress, etc., were analyzed in conjunction with data from follow-up reports of actual events.
4. Draft (un-validated) survey instruments were used during the pilot phase of the project. However, findings

from the analyses of data collected with the pilot instruments have been logical, consistent with theory, and consistent with findings published by other researchers in the area. Moreover, post-pilot analyses suggest the draft instruments were adequately valid and reliable.

Limitations such as these are common in survey research in the social sciences, especially among pilot studies. With respect to the current bulletin however, the research team found no evidence that these limitations have manifested at unacceptable levels. With regard to the Training Evaluation Project overall, limitations in general have diminished as the project has evolved beyond the pilot stages. Limitations specific to particular findings are presented in conjunction with those findings, next.

Training, Code, and Number of Participants	Completed Surveys (N = 4,140)						Response Rates	
	Participant Demographic Form	Overall Training Evaluation	Trainer Evaluation	Current Applications	Current Applications Follow-up	Initial	Follow-up	
<b>CER</b> 05D1501	24	24	120	22	21	99.0%	87.5%	
05D1502	27	27	108	25	22	98.9%	81.5%	
06D1501	21	21	105	21	21	100.0%	100.0%	
<b>CLD-A</b> 05M101	22	22	110	22	15	100.0%	68.2%	
05M102	25	25	150	25	23	100.0%	92.0%	
05M103	22	22	191	17	19	94.7%	86.4%	
<b>CLD-J</b> 05D101	27	27	242	27	26	99.7%	96.3%	
06D101	21	21	164	21	19	98.3%	90.5%	
<b>JATD</b> 05D801	21	21	83	20	20	98.6%	95.2%	
06D801	24	c	69	24	22	97.5%	91.7%	
<b>MDF</b> 05R039	34	34	136	33	25	99.6%	73.5%	
06R012	27 <sup>a</sup>	a	78	27	21	97.0%	77.8%	
06R019	27 <sup>a</sup>	a	68	22	14	83.7%	51.9%	
<b>MNFO</b> 05D1001	29	29	99	d	19	87.9%	65.5%	
05D1002	25	25	100	24	16	98.9%	64.0%	
06D1001	18	18	88	18	13	98.6%	72.2%	
06D1002	31	31	154	31	11	94.8%	35.5%	
06D1003	39	39	174	35	12	90.7%	30.8%	
<b>NFD</b> 06D301	33	33	131	33	25	99.5%	75.8%	
<b>TDD</b> 05D902	15	15	60	d	14	100.0%	93.3%	
<b>TOTALS</b>	512 <sup>ab</sup>	458 <sup>b</sup>	427	2430	447	378	96.8%	73.8%

<sup>a</sup> MDF was a three phase training with the same group of participants; demographics were collected only during phase 1.  
<sup>b</sup> Five other participants attended two trainings and completed demographic forms at each training. Of the 512 recorded participants and 458 completed demographic forms, 453 represent distinct participants.  
<sup>c</sup> Portions of these data are invalid due to survey error.  
<sup>d</sup> These data were not collected due to insufficient preparation time between award date and training dates.

# Findings

Findings from the analyses of data from 4,140<sup>2</sup> survey instruments completed by 453 training participants in 2005-2007 (Table 1) are presented in three sections below. The first section, **Training Results**, describes participant progress on training-specific objectives during training. The next section, **Activity Levels**, compares participant involvement levels in key activities (training-related behaviors) before and after training. The final section, **Implementation Results**, describes participant progress on implementing training-specific objectives and participant-identified objectives in their organizations.

In each of these areas, analyses revealed considerable variation in participant outcomes. Therefore, **potential sources of outcome variation** were examined in greater detail. Several hundred variables from the five surveys listed in Table 1 were considered as possible outcome predictors. Based on previous research, an evolving theory of training outcomes, and results from preliminary analyses, approximately 75 of those variables were selected and arranged into three sets or models:

1. **Pre-training measures** include personal or demographic<sup>12</sup> variables such as age, experience, education, prior training, area of employment, gender, salary, number of staff supervised, etc.
2. **Training measures** include several variables of training quality, training relevance, trainer effectiveness, training type, etc.
3. **Post-training measures** consist primarily of organizational variables, such as various resources and barriers<sup>4</sup> to implementing training objectives on the job, but also include length of follow-up period, subsequent training, etc.

**Multiple regression modeling** was then used to test the predictive or explanatory power of the three models. Although multiple regression modeling is an advanced statistical technique, the basics of the procedure are relatively simple: it tests the predictive power of each variable in the model, independently and collectively. In lieu of further explanation or lengthy examples, suffice it to say that **no technical understanding of the procedure is required** to grasp findings as presented in the bulletin narrative. Although multiple regression modeling produces several statistics, only the percentage of explained outcome variation<sup>13</sup> is presented in the bulletin narrative; some of the more technical results are presented in endnotes.<sup>14</sup>

The extent to which each of the three regression models predict or explain variation in **Training Results**, **Activity Levels**, and **Implementation Results**, is discussed in each of those sections to follow. When interpreting the percentage of outcome variation explained by each of the three models, please note the following:

1. **Pre-training measures** such as age, rank, or gender are expected to explain very little outcome variation in relatively uniform or homogenous settings such as training environments, unless there is bias in the selection or treatment of training participants. In more stratified environments, such as the realities of the workplace in correctional organizations, demographic variables may explain somewhat more outcome variation simply because some persons in the organization are more highly placed, or more empowered to act, than others. In either environment, however, pre-training measures are expected to explain relatively little outcome variation compared to training measures and post-training measures.
2. **Training measures**, on the other hand, are expected to explain relatively more outcome variation because, theoretically and practically, one would expect variables such as training quality and relevance to be associated with (predictive of) positive outcomes.
3. **Post-training measures** typically explain relatively large percentages of outcome variation. This is neither good nor bad, per se. It simply reflects the common sense (and research supported) notion that the organization matters, and that training graduates must contend with the complexities and difficulties of post-training environments in order to transfer training substance to the workplace.

## Training Results

### **Participant Progress on Training-specific Objectives during Training**

On the last day of each training, participants were asked to rate their progress on a series of training objectives as a result of participating in the training. Although objectives varied by training,<sup>15</sup> a consistent response format was used on all surveys, i.e., participants rated their progress on each training objective as: none (1); slight (2); moderate (3); substantial (4); or exceptional (5). These ratings were then summed and divided by the number of objectives to arrive at an average (mean) progress score for each participant.

Overall participants reported an average progress rating of 3.76 (N=427, SD=0.68), which indicates moderate to substantial progress on training objectives (Figures 1a and 1b). Larger subgroups of participants varied little in mean progress ratings. For example, the mean progress ratings for participants in juvenile and adult training programs were

3.72 (N=277, SD=0.68) and 3.85 (N=150, SD=0.66), respectively. Likewise, mean progress ratings for participants in 2005 and 2006 training programs were 3.72 (N=264, SD=0.70) and 3.83 (N=163, SD=0.64) respectively.

However, mean progress ratings varied somewhat more for smaller subgroups of participants, such as individual trainings. For example, mean progress on training objectives varied from a low of 3.35 (N=27, SD=.81) for participants in 05-D1502 (Critical Elements of Re-entry/Continuing Care Systems) to a high of 4.26 (N=22, SD=.54) for participants in 05-M101 (Correctional Leadership Development – Adult) as illustrated in Figure 1a.

Outcome variation is most evident in the mean progress ratings of individual participants, who ranged from almost no progress (1.12) to exceptional progress (5.00) on training objectives, as illustrated in Figure 1b. Each bar in the figure represents the number of participants (vertical axis) with mean progress ratings between the numbers on the horizontal axis. For example, the two tallest bars in the center of the figure each indicate that about 125 participants reported mean progress ratings between 3.50 and 4.00, and between 4.00 and 4.50.

### Potential Sources of Outcome Variation

Clearly, most participants feel they made at least moderate progress on training objectives during their training (85% of participants reported mean progress scores of 3.0 or higher). Nonetheless, some participants reported very low progress ratings, and the variation in progress ratings overall is considerable (Figure 1b). Multiple regression analysis was utilized to test two models or potential sources of this variation in training results, as illustrated in Figure 1c:

- **Pre-Training Measures** (participant demographic and personal variables) explained or predicted only about 4% of the variation<sup>16</sup> in progress ratings on training objectives.

- **Training Measures** (training quality and relevance; trainer effectiveness, etc.) explained or predicted over 30% of the variation<sup>17</sup> in progress ratings.

- **Post-Training Measures** (resources and barriers<sup>4</sup> to implementing training objectives on the job, follow-up period, subsequent training, etc.) were not tested as possible predictors because they take place after training results are determined.

Note that explained variation from different models can not be summed because it is possible for different models to explain or predict some portion of the same variation.<sup>18</sup> However, models can legitimately be compared for relative explanatory or predictive power. In this case, for example, measures of training quality and relevance are clearly more predictive of training results (progress on training-specific objectives during training) than are pre-training and personal measures.

### Interpretation and Practical Significance of Findings

Although reported progress on training-specific objectives varied among participants (Figure 1b), results are broadly positive. Even among trainings with the lowest mean (average) ratings, about 80% of participants reported at least moderate progress.

Moreover, multiple regression results are encouraging (Figure 1c). Findings from the pre-training model are especially positive, with only 4% of the variation in training results explained by personal or demographic characteristics of the participants. (Or perhaps more intuitively, 96% of the variation in training results is explained by factors **other than** the participant demographic variables reflected in the pre-training model.) This finding suggests that both the participant selection process and participant treatment during training were unbiased with respect to demographic variables, e.g., age, gender, rank, etc.

Likewise, findings from the training model are somewhat

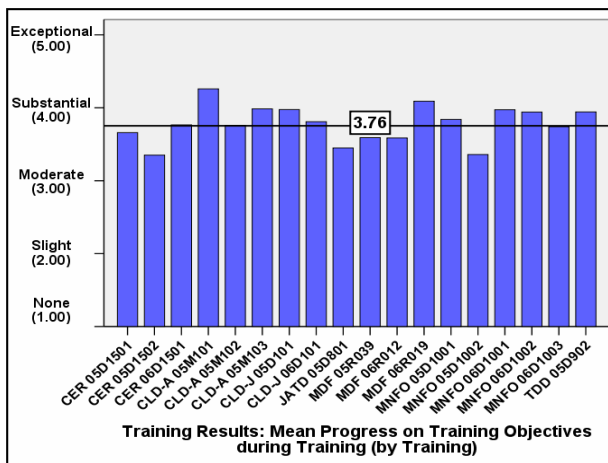


Figure 1a

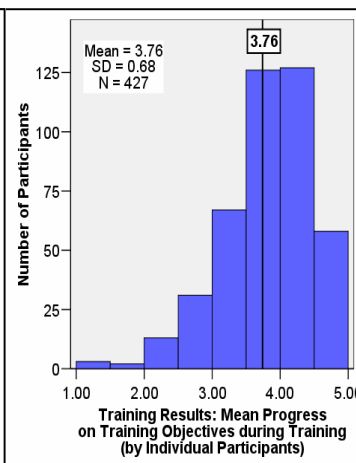


Figure 1b

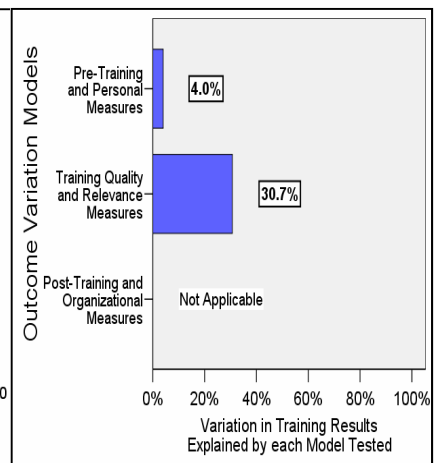


Figure 1c



positive, explaining 30.7% of participant variation in training results; a reasonably high percentage by social science standards. Thus training quality and relevance, trainer effectiveness, and other measures included in the model are moderately predictive of training results. This provides some evidence to support the common sense notion that improvements in training quality may improve outcomes among training participants. However, an alternative view of this statistic is that

*nearly 70% of the variation in training results could not be explained by measures of training quality.*

This finding cautions that further efforts to improve trainings that are already moderately to highly rated may meet with diminishing returns more quickly than expected.

Improved return on investment, for NIC and its constituent individuals and organizations, may be achieved by placing greater emphasis on the organizational environments to which participants return after training. Several approaches are possible, e.g., direct efforts to improve the balance of organizational resources and barriers to the transfer of training, or the selection of training participants from organizations already exhibiting a favorable balance. While this is not a novel proposition,<sup>19</sup> findings from the Training Evaluation Project, especially those presented in the remainder of this bulletin, offer considerable support for such an approach.

## Activity Levels

### Participant Involvement in Key Activities Before and After Training

At each training, participants were asked to indicate whether they had engaged in certain key activities (behaviors) related to training objectives in the three (or six) months<sup>3</sup> prior to training. Follow-up surveys mailed three (or six) months after training asked participants to indicate whether or not they had engaged in these same training-related activities **since completing the training**. Note that activities,<sup>20</sup> time frames,<sup>3</sup> and the precise wording<sup>11</sup> of the questions varied by training.

Overall participants reported engaging in 52.8% (N=365, SD=.33) of these behaviors prior to training (Figures 2a and 2b) and 70.4% (N=255, SD=.29) of the behaviors after training (Figures 3a and 3b). Likewise, increased post-training activity levels were evident in all subgroups examined, as summarized in Table 2 (page 7). For example, mean activity levels for participants in:

- Juvenile training programs increased from 48% to 64%.
- Adult training programs increased from 77% to 93%.
- 2005 training programs increased from 62% to 76%.
- 2006 training participants increased from 44% to 65%.
- Trainings with three month time frames<sup>3</sup> increased from 43% to 54%.
- Trainings with six month time frames<sup>3</sup> increased from 75% to 94%.

At the level of individual trainings, 12 of the 13 trainings with valid<sup>21</sup> pre/post data displayed mean post-training activity level increases, as illustrated in Figure 4a. MNFO 06-D1002, on the other hand, displayed a 3.0% decrease in post-training activity levels. However, this finding should be interpreted cautiously as only 11 of the 31 participants provided follow-up data (Table 1). Such a low response rate can produce mixed or unrepresentative findings.<sup>22</sup> The data displayed in Figures 4a and 4b are especially

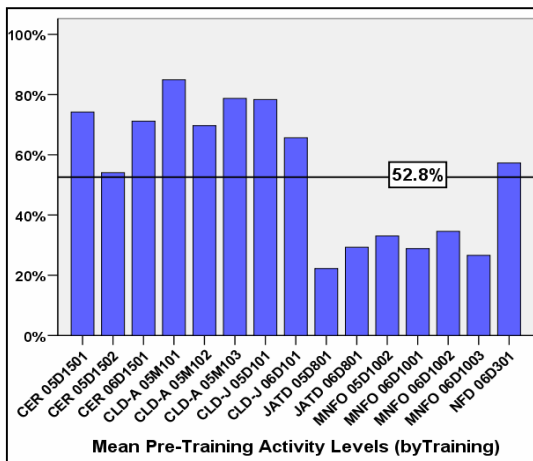


Figure 2a

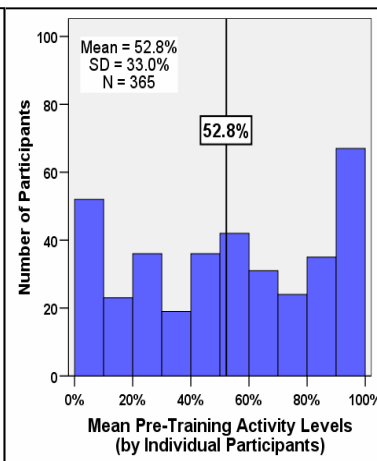


Figure 2b

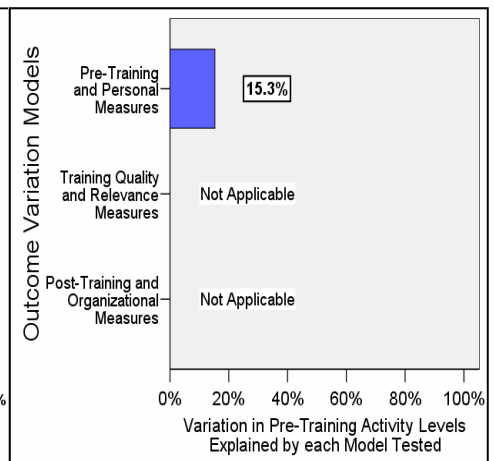


Figure 2c

sensitive to response rates because change scores can not be established for individual participants without both pre and post data. Thus Figures 4a and 4b reflect pre/post change scores only for the 235 participants who provided both pre and post-training activity level data. Figures 2a and 2b, on the other hand, reflect all available pre-training activity level data, i.e., for the 365 participants who answered those questions during training. Likewise, Figures 3a and 3b reflect all available post-training activity level data, i.e., for the 255 participants who answered those questions in the follow-up.

At the level of individual participants, analysis of follow-up data revealed post-training activity levels:

- increased for 161 of 235 participants, averaging 29.1% (SD=17.6%)
- did not change for 30 participants, and
- decreased for 44 participants, averaging -17.2% (SD=16.0%)

Overall, the 235 participants who provided both pre and post data averaged 16.7% increases in activity levels after training (Figure 4b). Note that references to percentage increases (or decreases) in this section should be interpreted in the common sense (no pun intended) rather than in a strict mathematical sense. For example, a participant who reported engaging in 30% of key activities prior to training and 60% of key activities after training would be referred to as a 30% increase, not a 100% increase.

Although these post-training activity level changes are quite positive overall, outcomes varied substantially among individual participants (Figure 4b) and among the various subgroups examined. Almost a third of the 235 participants who provided both pre and post data reported no increases in post-training activity levels. Moreover, several subgroups, though displaying post-training activity level increases, nonetheless failed to rise to the pre-training levels of their counterparts, e.g., juvenile trainings compared to adult trainings, and trainings with three month time frames compared to those with six month time frames (Table 2).

Possible explanations for the variation evident in pre-training and post-training activity levels are discussed next.

### Potential Sources of Outcome Variation

#### Pre-Training Activity Levels

As illustrated in Figure 2b, the 365 training participants who provided data varied considerably in their pre-training activity levels, ranging from 0 to 100% (SD=33.0%). Although pre-training activity levels are included here primarily as a baseline, they can also be viewed as potential outcomes of certain **pre-training measures**, such as those included in the first regression model, e.g., age, gender, salary, etc. In fact, analyses revealed that about **15%** of the variation<sup>23</sup> in pre-training activity levels can be explained or predicted by pre-training measures such as participant demographic and personal variables (Figure 2c). (Note that training measures and post-training measures were not tested as possible predictors because pre-training activity levels precede such measures in time.)

#### Post-Training Activity Levels

Although the variation in participants' post-training activity levels (Figure 3b) was somewhat less than the variation in pre-training activity levels (Figure 2b), it was nonetheless substantial, again ranging from 0 to 100% (SD=28.9%). Multiple regression analysis was utilized to test three models (Figure 3c) or potential sources of the variation in post-training activity levels:

	Before Training			After Training		
	Mean	N	SD	Mean	N	SD
Juvenile	47.6%	301	33.1%	63.8%	198	29.2%
Adult	77.3%	64	17.6%	93.3%	57	10.0%
2005	62.0%	182	33.1%	75.7%	132	28.6%
2006	43.7%	183	30.2%	64.7%	123	28.3%
3 months	42.8%	253	33.0%	54.4%	153	26.5%
6 months	75.4%	112	18.8%	94.3%	102	9.0%

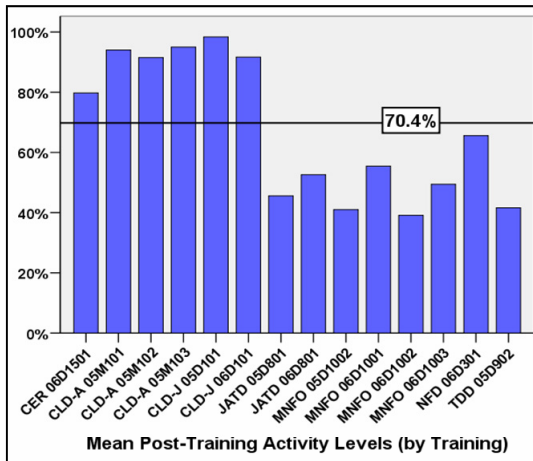


Figure 3a

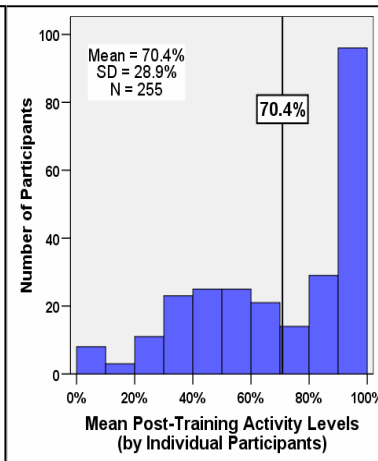


Figure 3b

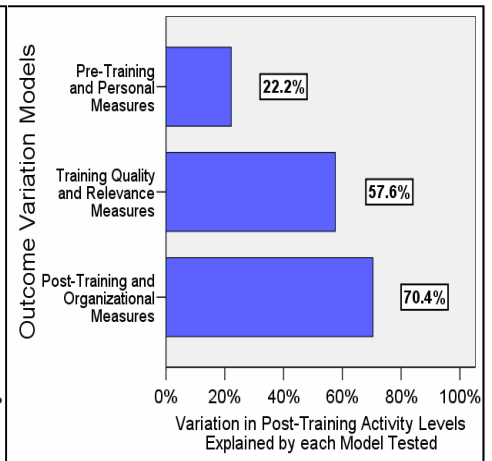


Figure 3c

- **Pre-Training Measures** (participant demographic or personal variables) explained or predicted about **22%** of the variation.<sup>24</sup>
- **Training Measures** (training quality and relevance; trainer effectiveness) explained or predicted about **57%** of the variation.<sup>25</sup>
- **Post-Training Measures** (resources and barriers<sup>4</sup> to implementing training objectives on the job, follow-up period, etc.) explained or predicted over **70%** of the variation<sup>26</sup> in post-training activity levels.

Note that explained variation from different models can not be summed.<sup>19</sup> However, models can legitimately be compared for relative explanatory or predictive power. In this case, for example, post-training measures such as organizational resources and barriers are more predictive of post-training activity levels, than either pre-training demographic measures, or training quality and relevance measures.

Although the amount of outcome variation explained by each model can not be summed in the mathematical sense of adding percentages, the variables in the three models may nonetheless interact and reinforce (or counteract) each other. In the case of variables from several models reinforcing each other, or being predictive of similar outcomes, the cumulative effect can be striking. For example, the presence of two very distinct groups of trainings in Figure 3a, a “**high group**” of six trainings on the left, and a “**low group**” of eight trainings on the right, is explained or predicted by a combination of all three models. In terms of **pre-training measures**, participants in “high group” trainings were older, more experienced, higher ranked, higher paid, personnel from larger organizations, on average, than those from the “low group”. Additionally, in terms of **training measures**, “high group” trainings were ranked significantly higher by participants on measures of training quality, training relevance, and training results. In terms of **post-training measures**, five of the six “high group” trainings were followed-up after six months, while all eight of the “low group” trainings were followed-up after only three months. Moreover, participants in the “high

group” trainings on average reported greater resources and fewer barriers to training transfer in their organizations, than did participants in “low group” trainings.

### Activity Level Changes

Although the multiple regression models used in this bulletin are appropriate for testing predictors of pre-training activity levels and post-training activity levels, they are not appropriate for directly testing predictors of pre/post changes. The regression models each focus on a single span of time, either before, during, or after training, while change scores encompass all three. Insufficient follow-up data were available to support larger models necessary for direct testing of change score predictors. Nonetheless, the identified predictors of post-training activity levels (Figure 3c) serve much the same purpose.

### Interpretation and Practical Significance of Findings

Although pre-training activity levels (Figure 2b), post-training activity levels (Figure 3b), and activity level changes (Figure 4b) all varied considerably among training participants, results are quite positive overall. For example, over **68%** of the 235 participants who provided both pre and post data reported activity level increases after training, compared to less than **19%** who reported decreased activity levels. Superficially these and other activity level findings may not appear especially positive or significant. However...

*These findings are based on specific training-related behaviors executed in complex organizations despite hindrances and distractions...*

*...an entirely different matter than findings based on intentions or opinions.*

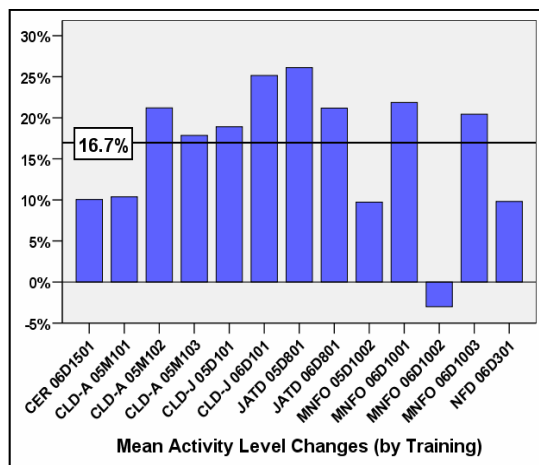


Figure 4a

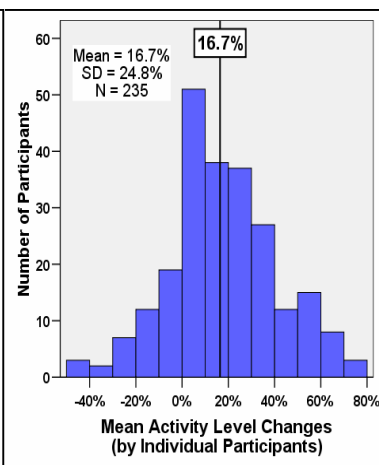


Figure 4b

Moreover, multiple regression results regarding potential sources of outcome variation among participants are enlightening and generally positive. **Pre-training measures** (participant demographic and personal characteristics) were slightly predictive of activity levels, explaining about **15%** of variation in activity levels before training (Figure 2c) and about **22%** after training (Figure 3c). In a more homogenous or uniform setting, such as a training environment, this much outcome variation explained by demographic variables may indicate some bias, either intentional or unintentional, in the selection or treatment of training participants. However, in more stratified environments such as the



everyday reality of correctional organizations in which these behaviors took place, such findings typically indicate that some persons are simply more empowered to take action than others. In this case, for example, male participants who were older, more experienced, higher ranked, higher paid, and worked in larger organizations, were **slightly more likely** to engage in higher percentages of training-related activities (behaviors) in their organizations.

*Note that this is a good example of potential overlap between models,<sup>19</sup> i.e., the primary reason why the percentage of explained outcome variation from different models can not legitimately be summed. In this case, although these variables are considered personal or demographic characteristics of participants prior to attending training (**pre-training measures**) they are also lodged in the organizational context (**post-training or organizational measures**).*

**Training measures**, on the other hand, were much more predictive of post-training activity levels, explaining **57%** of the outcome variation among participants (Figure 3c). Thus measures regarding the individual training, its quality and relevance, trainer effectiveness, etc., are strongly associated with later training-related behaviors in the organization. This provides good evidence to support the common sense notion that improvements in training quality may improve outcomes among training participants. Nonetheless...

### **Post-training and organizational measures**

*proved to be the **most predictive** of training-related **behaviors** in the workplace, explaining over **70%** of the variation among participants after training.*

While this finding does not diminish the importance of training quality, it does suggest that **post-training and organizational issues are major considerations in the transfer of training to the workplace** (Figure 3c). Specific predictors of higher post-training activity levels include longer follow-up period, decreased organizational barriers (especially workload and inadequate funding), increased organizational resources, and increased progress implementing training objectives in the organization. The final predictor, implementation progress, suggests that participants are more likely to initiate, sustain, or increase training-related behaviors if the behaviors appear to be fruitful.

## **Implementation Results**

### **Progress Implementing Training-specific Objectives and Participant-identified Objectives on the Job**

Either three or six months<sup>3</sup> after each training, participants were asked to rate their progress, since completing training, on implementing two types of objectives in their organizations:

- **Training-specific objectives**, which varied by training, but were common to all participants within each training.<sup>15</sup>
- **Participant-identified objectives**, which each participant had previously identified during training in response to open-ended questions regarding:  
“...current strategies you intend to change...”  
“...behaviors you intend to improve...”  
“...problems or issues in your organization you intend to address...”

Participants rated their progress on each objective as: none (1); slight (2); moderate (3); substantial (4); or exceptional (5). These ratings were then summed and divided by the number of objectives to arrive at a mean progress score for each participant.

### **Progress Implementing Training-specific Objectives**

Overall participants reported a mean or average post-training progress of 2.86 (N=294, SD=.74), or approximately moderate progress implementing training-specific objectives in their organizations (Figure 5b). Larger subgroups of participants varied slightly in their implementation results. For example, participants' mean implementation progress was:

- 2.78 (N=238, SD=.75) in juvenile training programs
- 3.16 (N=56, SD=.59) in adult training programs
- 2.67 (N=193, SD=.76) in three month follow-up programs
- 3.22 (N=101, SD=.54) in six month follow-up programs

Mean implementation progress varied somewhat more for smaller subgroups of participants, such as individual trainings (Figure 5a). For example, mean progress implementing training-specific objectives varied from a low of 2.26 (N=12, SD=.81) or slight progress, for participants in 06-D1003 (Meeting the Needs of Juvenile Female Offenders) to a high of 3.50 (N=19, SD=.47) or moderate to substantial progress, for participants in 05-M103 (Correctional Leadership Development – Adult); this is illustrated in Figure 5a.

The full extent of outcome variation is most evident in the mean implementation ratings of individual participants, who ranged from no progress to exceptional progress (1.00 - 4.78) as illustrated in Figure 5b. For example, about 30% of participants reported slight or no progress implementing training-specific objectives, while about 15% reported substantial to exceptional progress.

### Potential Sources of Outcome Variation

As illustrated in Figure 5b, the 294 training participants who provided data varied considerably in their progress implementing training-specific objectives in their organizations. Multiple regression analysis (Figure 5c) revealed that:

- **Pre-Training Measures** (participant demographic or personal variables) explained or predicted about **4%** of the outcome variation.<sup>27</sup>
- **Training Measures** (training quality and relevance; trainer effectiveness) explained or predicted about **25%** of the outcome variation.<sup>28</sup>
- **Post-Training Measures** (resources and barriers<sup>4</sup> to implementing training objectives on the job, follow-up period, etc.) explained or predicted about **45%** of the variation<sup>29</sup> in progress implementing training-specific objectives in the organization.

(Note that explained variation from different models can not be summed.<sup>19</sup>)

In this case, post-training and organizational measures such as resources and barriers to implementation are clearly more predictive of progress implementing training-specific objectives than either pre-training and personal measures, or measures of training quality and relevance.

### Progress Implementing Participant-Identified Objectives

Overall participants reported a mean or average post-training progress of 3.05 (N=194, SD=.81), or moderate progress implementing participant-identified objectives in their organizations (Figure 6b). Larger subgroups of participants varied little in their implementation results. For example, participants' mean implementation progress was:

- 3.03 (N=103, SD=.88) in juvenile training programs
- 3.07 (N=91, SD=.73) in adult training programs
- 2.97 (N=67, SD=.97) in three month follow-up programs
- 3.09 (N=127, SD=.71) in six month follow-up programs

- 3.16 (N=95, SD=.71) in 2005 training programs
- 2.94 (N=99, SD=.88) in 2006 training programs

However, mean implementation progress varied somewhat more for smaller subgroups of participants, such as individual trainings (Figure 6a). For example, mean progress implementing participant-identified objectives varied from a low of 2.39 (N=6, SD=1.08) or slight progress, for participants in 06-D1003 (Meeting the Needs of Juvenile Female Offenders) to a high of 3.36 (N=15, SD=.90) or moderate progress, for participants in 05-D1002 (Meeting the Needs of Juvenile Female Offenders) as illustrated in Figure 6a.

Greater outcome variation is evident in the mean implementation results of individual participants (Figure 6b), who ranged from no progress to exceptional progress implementing participant-identified objectives in their organizations. For example, about 25% reported slight or no progress implementing participant-identified objectives, while about 30% reported substantial to exceptional implementation progress.

### Potential Sources of Outcome Variation

As illustrated in Figure 6b, the 194 training participants who provided data varied considerably in the amount of progress they reported on implementing participant-identified objectives. Multiple regression analyses (Figure 6c) revealed that:

- **Pre-Training Measures** (participant demographic or personal variables) explained or predicted **0%** of the outcome variation.<sup>30</sup>
- **Training Measures** (training quality and relevance; trainer effectiveness) explained or predicted about **7%** of the outcome variation.<sup>31</sup>
- **Post-Training Measures** (resources and barriers to implementing training objectives on the job, follow-up period, etc.) explained or predicted about **30%** of the variation<sup>32</sup> in implementation results for participant-identified objectives.

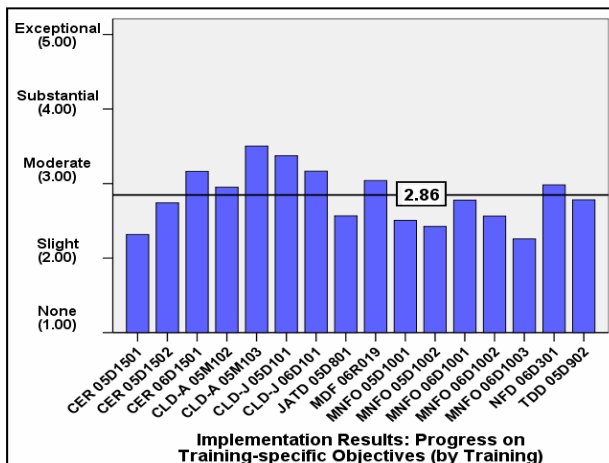


Figure 5a

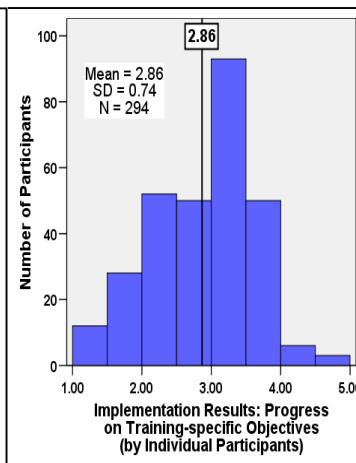


Figure 5b

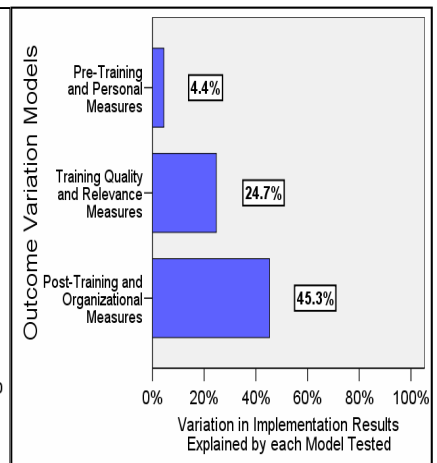


Figure 5c

Once again, post-training and organizational measures are clearly more predictive of outcomes, in this case progress implementing participant-identified objectives, than either pre-training or training measures.

**Interpretation and Practical Significance of Results**

Implementation results for both training-specific objectives (Figure 5a) and participant-identified objectives (Figure 6a) were **more modest** overall than the *training results* reported earlier (page 4; Figures 1a and 1b). Nonetheless, these findings are encouraging. In fact, a compelling argument can be made that

**“moderate progress”**  
implementing objectives  
in the organization

is of greater practical significance  
than **“substantial progress”** on objectives  
in the training environment.

Although logic and research suggest that progress on objectives during training is an important element in the transfer of training to the organization, that transfer is ultimately measured by implementation results.<sup>33</sup>

Findings from multiple regression analyses are generally positive and provide some insight into potential sources of the variation in participant implementation results. The **pre-training model** (personal and demographic measures) explained very little (**4.4%**) of the variation in implementation results for training-specific objectives and none of the variation in implementation results for participant-identified objectives. These are positive findings in that they reveal no evidence of bias (based on variables in the model, such as gender, age, rank, etc.) in the selection or treatment of training participants. There is also no evidence of implementation barriers based on those pre-training variables,

i.e., barriers that apply differentially based on gender, age, rank, etc.

Findings from the **training model** reveal that measures of training quality, relevance, trainer effectiveness, etc., are reasonably predictive of implementation results for training-specific objectives, explaining about **25%** of the outcome variation. While these results are not as positive as those from the pre-training model, they nonetheless are moderately encouraging and enlightening. For example, follow-up reports of implementation progress in the organization are several steps removed (in terms of time and the type of environment) from initial reports of training quality collected during training. For these initial reports of training quality to predict even 25% of the variation in implementation results, several months later in the very different environment of the organization, suggests that higher training quality may in fact improve training transfer to the organization.

With regard to participant-identified objectives, it is neither particularly surprising nor negative that **training measures** explained only **7%** of the variation in implementation results. Unlike training-specific objectives, which were standardized for each training and closely tied to training content, participant-identified objectives were diverse and not as closely tied to training content.

**Post-training and organizational measures**, not surprisingly, also explained more variation in implementation results for the standardized training-specific objectives (**45.3%**) than for the more diverse participant-identified objectives (**29.8%**). Of greater practical significance, however, is the fact that this model explained substantially more variation in implementation results, **for both types of objectives** than **either of the other models**. This echoes previous findings throughout the bulletin, in that it suggests the everyday realities of correctional organizations are important in the transfer of training to the workplace, and may thus be important to consider during training design and delivery.

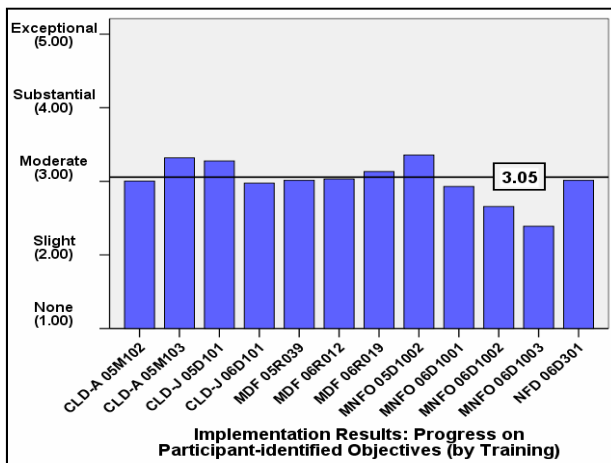


Figure 6a

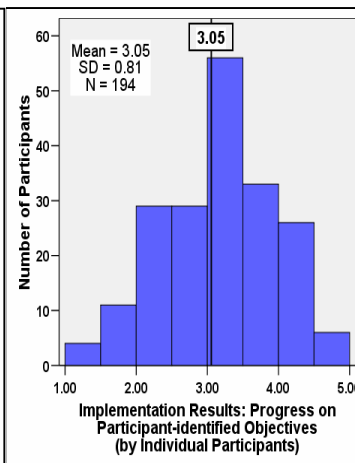


Figure 6b

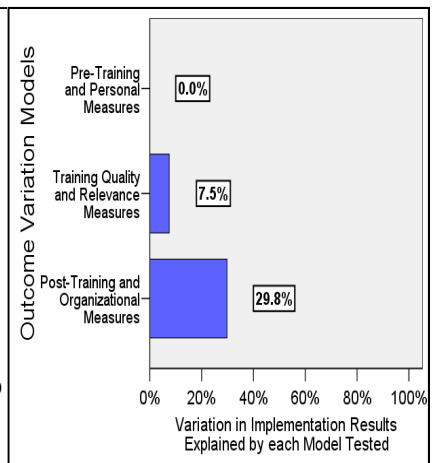


Figure 6c

# Summary and Recommendations

Data from over 4,000 completed surveys were analyzed for the bulletin (Table 1). Three outcomes were examined:

- **Training Results** (progress on training objectives during training)
- **Activity Levels** (training-related behaviors on the job, before and after training)
- **Implementation Results** (progress implementing training-specific objectives and participant-identified objectives in the organization)

Several hundred variables were considered as possible predictors of these outcomes; approximately 75 were selected, arranged into three models, and tested:

- **Pre-training measures** (participant demographics)
- **Training measures** (training quality, relevance, etc.)
- **Post-training measures** (organizational resources and barriers,<sup>4</sup> follow-up period, etc.)

Outcomes are summarized in Table 3, along with the predictive value of each model, for each outcome. **On average, outcomes were quite favorable:** participants reported moderate to substantial progress on training objectives during training; increased training-related behaviors on the job after training; and moderate progress implementing training objectives in their organizations. However, these are average measures; **individual participants varied considerably in their outcomes.** The extent to which each model accounted for variation in each outcome is summarized in Table 3.

All models were more predictive of activity levels than implementation results. This is to be expected given that implementation results are further removed from training (more distant outcomes). **Training measures** explained significantly more variation in all outcomes than did **pre-training measures**. These are **favorable and important findings**. For theoretical and practical reasons one would expect training measures, such as quality and relevance, to be predictive of participant outcomes. Current findings reveal that training measures were moderately to strongly predictive of training results, post-training activity levels, and some implementation results. Likewise, one would expect pre-training demographic variables such as age, race, and gender to **not** be very predictive of participant outcomes. Findings strongly support this expectation and reveal no evidence of bias in the selection or treatment of training participants.

**Post-training and organizational measures**, however, **proved to be the most predictive of all outcomes** where the model was tested. These findings highlight

*the limits of training and the importance of post-training environments when attempting to change behavior or implement training objectives in correctional organizations.*

As echoed in findings throughout the bulletin, efforts to improve trainings that are already moderately to highly rated may meet with diminishing returns more quickly than expected. The findings presented in this bulletin suggest that better returns on investment for NIC and its constituent organizations may be achieved by focusing resources on the organizations to which participants return after training.

**Based on these findings, consideration should be given to the following recommendations:**

1. As part of needs and viability assessment processes, establish minimum standards for an acceptable balance of resources and barriers to implementing training objectives in an organization.
2. Develop and implement a procedure to evaluate whether a training applicant's organization meets the minimum standards.
3. Amend training target audiences to include these minimum standards and either: a) select training participants consistent with the amended target audience, or b) provide technical assistance or other aid to improve the balance of resources and barriers when accepting training participants from organizations with an unfavorable balance.
4. Expand the CEO endorsement to include statements that the CEO has reviewed and supports the training objectives, and will assist the training participant in implementing those objectives and resulting action plans. (The granting, or rejecting, of this expanded endorsement can be considered part of #2 above.)
5. Training evaluation follow-up periods should be six months instead of (or in addition to) three months, and should be standardized as much as is feasible.

Outcomes		Predictors <sup>1</sup>		
		Pre-Training and Demographic Measures	Training Quality and Relevance Measures	Post-Training and Organizational Measures
Training Results	3.76 <sup>a</sup> (SD=.68, N=427)	4.0%	30.7%	
Activity Levels (Before Training)	52.8% <sup>b</sup> (SD=33%, N=365)	15.3%		
Activity Levels (After Training)	70.4% (SD=29%, N=255)	22.2%	57.6%	70.4%
Implementation Results (Training-Specific)	2.86 <sup>c</sup> (SD=.74, N=294)	4.4%	24.7%	45.3%
Implementation Results (Participant-Identified)	3.05 <sup>d</sup> (SD=.81, N=194)	0.0%	7.5%	29.8%

<sup>1</sup> Percentage of variation in each outcome that was explained or predicted by each model (set of predictors).

<sup>a</sup> 3.76 represents more than moderate progress; almost substantial progress.

<sup>b</sup> The pre-training activity level is intended more as a baseline measure; not an outcome per se.

<sup>c</sup> 2.86 represents approximately moderate progress.

<sup>d</sup> 3.05 represents moderate progress.



## Future Directions

As mentioned in previous bulletins, training evaluation results can be beneficial in improving both future trainings and training evaluations. Training participant progress on training objectives, post-training behavior changes, and implementation results in the workplace, as described in this bulletin, provide insight to guide design of more efficient and streamlined trainings and training evaluations. These improvements, together with the future directions of this evaluation project and the bulletin series, are discussed below.

**The central finding of this bulletin pertains to the importance of post-training factors in the successful implementation of training objectives in correctional organizations.** Findings suggest that NIC may want to increase the emphasis placed on selection criteria for their trainings, particularly with regard to resources and barriers<sup>4</sup> to implementation in the applicant's organization. This is not to suggest that NIC should exclude applicants from organizations displaying a balance of resources and barriers that seems unfavorable to workplace implementation; in fact, these may sometimes be the very organizations that could benefit the most from training. Nor is it to suggest that the selection process should be "stacked" in favor of organizations with an abundance of resources and a paucity of barriers; these will sometimes be the organizations that least need NIC training services. What is being suggested is that an unfavorable ratio of resources to barriers should be a screening factor taken into account during the process of assessing need for training and, correspondingly, the viability of transferring training substance to the workplace.

Similarly, this factor may impact decisions about the design and delivery of trainings. For example, personnel from organizations with an unfavorable balance of resources and barriers may need certain specialized types of training (e.g., technical assistance) in addition to (or other than) the trainings typically offered. This approach could help teach such organizations to cultivate and develop resources as a prelude to implementation and then use these resources to help minimize the impact of barriers. In this manner, training recipients and their respective organizations might be better prepared to contend with the complexities and difficulties of post-training environments, thereby circumventing some of the frustration and demoralization that can accompany a failure to implement training objectives as intended. This would help insure that budgeted training dollars are utilized at maximum efficiency.

A related consideration in selecting participants for a given training is the mixture of those from organizations with favorable balances and those from organizations with unfavorable balances – with the goal of achieving a sort of balancing of balances. The point is that a properly planned and structured mix could afford participants from organizations with unfavorable balances the opportunity to learn

from and collaborate with their peers who work in organizations having more favorable ratios. At base, all such approaches speak to the quality of the fit between the training being delivered and the organization(s) in which it will be implemented.

Although the pilot phase of the Training Evaluation Project (2005-2006) has concluded, the project continues to evolve and mature. An example of this, discussed in previous bulletins, is an Evaluation Matrix that has been developed to assist researchers and training evaluators in the selection of appropriate evaluation types and level of rigor for a given training situation. Ongoing development of the Evaluation Matrix is being guided by evaluation findings. For example, based on findings presented in the current bulletin, the *Needs Assessment* section of the matrix has been expanded to *Needs and Viability Assessment* to reflect the importance of considering the post-training and organizational issues (such as the balance of resources and barriers) from the outset. Likewise, based on feedback from earlier findings, the development of a simplified version of the Evaluation Matrix is underway. The simplified version is designed to provide instructional designers, trainers, and other practitioners with a menu of options from which to choose when designing an in-house training evaluation.

In addition, based on earlier and current findings, trainings are now being evaluated with standardized, condensed, and validated instruments, administered with less complex, more concise protocols. Furthermore, these new training evaluations are addressing needs that were identified in the pilot phase. Some of these needs include: collecting trainer demographics, allowing trainers to evaluate the training participants, collecting more detailed data on organizational resources and barriers, providing relevant NIC personnel with a preliminary evaluation report of their training within two weeks of training delivery, implementing training evaluations institute wide, and assisting NIC personnel in taking over many evaluation functions.

Future bulletins in the series will continue to present findings from the trainings that were evaluated in the pilot phase of the evaluation project. The next bulletin will focus on evaluating a particular type of training program that was offered multiple times.

For more information please contact:

Dr. James B. Wells  
President and Chief Research Consultant  
Commonwealth Research Consulting  
4160 Kentucky River Parkway  
Lexington, KY 40515

jbwells@cwrc.us  
(859) 806-5748



# Notes

<sup>1</sup> The project began with cooperative agreements 05A28GJF9 (March 2005) and 06PEI01GJM1 (July 2006) between NIC and Eastern Kentucky University's Center for Criminal Justice Education and Research. This was considered the pilot phase of the project. During this period it was referred to as the NIC Training *Academy* Evaluation Project (a reflection of its scope). Via cooperative agreement 07PEI12GJQ7 (August 2007) with Commonwealth Research Consulting, Inc., NIC expanded the scope of the project to potentially include any NIC training.

<sup>2</sup> The outcome data were drawn from portions of 427 overall training evaluations and 378 current applications follow-up surveys, while data tested as predictors were drawn from all five of the surveys in Table 1.

<sup>3</sup> The time frames were six months for evaluations of Correctional Leadership Development and Management Development for the Future, and three months for all other training programs.

<sup>4</sup> Resources and barriers were derived from content analysis of responses to open-ended questions on the Overall Training Evaluation and the Current Application Follow-up surveys that probed for these data. The most commonly identified resources were: personnel/staff; staff development/training; funding/infrastructure; support of management/administration; and support of a key person. The most commonly identified barriers were: staff/organizational resistance; insufficient funding/infrastructure; excessive workload or insufficient time; and agency structure/policy. See Bulletin 1: *Participant Demographics, Overall Evaluation of Training, and Applicability Ratings* (February 2007) for additional details on identified resources and barriers to the transfer of training substance to the workplace.

<sup>5</sup> The location of the bulletins is subject to change. If a search of the NIC website does not locate the bulletins, please contact Dr. James Wells (see page 13). Bulletin 1: *Participant Demographics, Overall Evaluation of Training, and Applicability Ratings* (February 2007). Bulletin 2: *Participant Evaluation of Trainers* (July 2007).

<sup>6</sup> All 453 training participants were informed that participation in the evaluation was voluntary, and that standard precautions would be taken to preserve confidentiality. Signed consent forms were collected from all participants. See Bulletin 1: *Participant Demographics, Overall Evaluation of Training, and Applicability Ratings* (February 2007) for a demographic profile of the training participants.

<sup>7</sup> Participants completed trainer evaluation forms for each trainer. The number of trainers per training varied from three to nine. See Bulletin 2: *Participant Evaluation of Trainers* (July 2007) for additional details.

<sup>8</sup> The current applications survey collected information on the current activities of participants and their colleagues, e.g., the ways in which they were (or were not) applying training content to their jobs. The title and content of this form varied somewhat from training to training.

<sup>9</sup> Dillman, D. A. (2000). *Mail and telephone surveys: The total design method* (2nd ed.). New York: Wiley. (The Total Design Method is an established protocol for maximizing survey response rates.)

<sup>10</sup> Prior to the establishment of NIC's Division of Research and Evaluation, training evaluations were not standardized, and data were collected, analyzed, and reported by individual training. Approximately 36 individual mini-reports based on individual training evaluations were produced during the pilot phase. The current research design employs a set of standardized evaluation tools designed to be used across trainings.

<sup>11</sup> The subject of the survey questions was adjusted to reflect the composition of each class, e.g., individuals (you), teams (you or your team), executives/directors (you or your designee), etc.

<sup>12</sup> Although there was not enough data to include race in the model, bivariate analyses revealed no significant correlations between race and any of the outcome variables, i.e., training results, activity levels, or implementation results.

<sup>13</sup> As used in this bulletin, percentage of explained outcome variation refers to adjusted R square.

<sup>14</sup> Adjusted R square, F scores, degrees of freedom, and significance levels are provided in endnotes as appropriate. The full output of multiple regression modeling, including the preparatory and follow-up procedures typically used in support of it, are voluminous and beyond the scope of this publication.

<sup>15</sup> The training-specific objectives are too numerous to list here; there were approximately 10-20 per training. If necessary, please contact Dr. James Wells for additional information.

<sup>16</sup> Adjusted R square = .040;  $F_{2,379} = 8.89$ ,  $p < .000$

<sup>17</sup> Adjusted R square = .307;  $F_{3,402} = 60.90$ ,  $p < .000$

<sup>18</sup> It is possible for different models to explain or predict some portion of the same variation. For example, if **pre-training measures** explain 15% of the variation in training results and **training measures** explain 40% of the variation, it is not accurate to say that pre-training measures and training measures together explain 55% of the variation.

<sup>19</sup> For over 20 years researchers have estimated that organizational resistance and other factors typically block up to 90% of training substance from transferring to the workplace. For example, see Ilian, H. (2004). Levels of levels: Making Kirkpatrick fit the facts and the facts fit Kirkpatrick. In B. Johnson, V. Flores, & M. Henderson (Eds.), *Proceedings of the 6th Annual Human Services Training Evaluation Symposium* 89-104. Berkeley, CA: California Social Work Education Center.

<sup>20</sup> The activities are training specific and too numerous to list here; there were approximately 15-30 per training. If necessary, please contact Dr. James Wells for additional information.

<sup>21</sup> Survey items regarding pre/post activity levels were not part of the evaluation protocol for five trainings. For two others, these items were erroneously omitted from the pilot survey.

<sup>22</sup> For example, the mean pre-training activity level of the 31 participants who provided initial data is 34.6% and the mean post-training activity level of the 11 who provided follow-up data is 39.1%. Thus the change in mean scores for MNFO 06-D1002 is a 4.5% increase. However, the mean change score for the 11 participant's who provided both pre and post data is a 3.0% decrease (Figure 4a). (On an individual level, 7 of these 11 participants reported post-training decreases in activity levels.) The contradictory results stem from differences in the groups of participants who did and did not provide follow-up data. For example, the mean pre-training activity level of the 11 who later provided follow-up data was 42.4%, while the mean pre-training activity level of the 20 who later did not provide follow-up data was 30.3%.

<sup>23</sup> Adjusted R square = .153;  $F_{5,314} = 12.51$ ,  $p < .000$

<sup>24</sup> Adjusted R square = .222;  $F_{5,209} = 13.25$ ,  $p < .000$

<sup>25</sup> Adjusted R square = .576;  $F_{5,149} = 42.96$ ,  $p < .000$

<sup>26</sup> Adjusted R square = .704;  $F_{4,125} = 77.60$ ,  $p < .000$

<sup>27</sup> Adjusted R square = .044;  $F_{1,238} = 11.92$ ,  $p < .001$

<sup>28</sup> Adjusted R square = .247;  $F_{4,206} = 18.18$ ,  $p < .000$

<sup>29</sup> Adjusted R square = .453;  $F_{3,126} = 36.66$ ,  $p < .000$

<sup>30</sup> Adjusted R square = -.058;  $F_{15,146} = .414$ ,  $p < .973$ . (Note that this model is not significant.)

<sup>31</sup> Adjusted R square = .075;  $F_{2,150} = 7.164$ ,  $p < .001$

<sup>32</sup> Adjusted R square = .298;  $F_{3,93} = 14.61$ ,  $p < .000$

<sup>33</sup> The extent (if any) to which training results, activity levels, or implementation results effect the organization will be the subject of a future bulletin on organizational change.