



Institute for Law and Justice
Alexandria, Virginia

Training Evaluation Model: Evaluating and Improving Criminal Justice Training

Final Report

September 2007

Submitted to
National Institute of Justice

Prepared by
**Kelly Bradley
Edward Connors
Institute for Law and Justice**

Acknowledgements

The authors are deeply appreciative of the cooperation and support of all the individuals who made this project possible. We are indebted to Mark Gage, Marcia Williams, and Jim Foley of the National White Collar Crime Center; William Brunson of the National Judicial College and Nancy Yeend and John Paul Jones of the John Paul Jones Group; Liebe Geft, Sunny Lee, and Mark Katrikh of the Simon Wiesenthal Center Museum of Tolerance; and Steve Morrison, Michael Logsdon, Randall Milks, Regina Patrosi, and Michael Hummel of the National Corrections and Law Enforcement Training and Technology Center. These individuals showed the courage to open their programs to outside evaluators.

We are also truly grateful for the cooperation of the training participants who agreed to be interviewed multiple times and complete surveys for this evaluation. We are also genuinely indebted to Detective Dave D'Amico of the Monmouth County, New Jersey Prosecutor's Office, and Detective Alix Olson of the Madison, Wisconsin, Police Department, for allowing us to learn more about their efforts to combat crimes of hate and terror.

A study of this depth and length does not thrive without strong support from the Office of Justice Programs' staff who provided guidance and oversight to the project, including Dr. Edwin Zedlewski, Dr. Katherine Browning and Maggie Heisler from NIJ, and Todd Brighton and Elaine Snyder from BJA.

Additionally, we would like to thank Dr. James Wells, Dr. Kevin Minor, and Dr. Gary Corder of Eastern Kentucky University for conducting and writing the evaluation of the NCLETTTC training. They also contributed significantly to the development of the training evaluation model. Finally, we would like to thank the ILJ staff who worked so hard on this project, including Chera Baity, Tiffany Hall, and Joan Peterschmidt for assistance in data collection; Deborah Spence for assistance with the literature review; and Dr. Thomas McEwen, Barbara Webster, and Laurie Samuel for their contributions in writing this report.

This project was supported by grant number 2003-DD-BX-K101 awarded to the Institute for Law and Justice by the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. Points of view in this document are those of the authors and do not necessarily represent the official position or policies of the U.S. Department of Justice

Table of Contents

Chapter 1: Project Background and Report Overview

Project Summary.....	1
Need for a Criminal Justice Training Evaluation Model.....	2
Overview of the Report.....	3

Chapter 2: Planning for Evaluations

Types of Evaluations.....	6
Evaluation Planning Steps	9
Identify Program Goals, Objectives, and Evaluation Questions.....	10
Develop Conceptual Framework and Logic Model.....	10
Design Evaluation Methodology	11
Conduct the Evaluation.....	15
Analyze and Communicate Evaluation Results.....	18
Evaluating Criminal Justice Training Programs.....	20
Opportunities for Control and Comparison Groups.....	20
Challenges in Evaluating Criminal Justice Training	20

Chapter 3: Factors That Contribute to Successful Practitioner Training Outcomes

Training Objectives.....	22
Adult Learning Concepts	23
Instructional Methods	25
Practical Training Matters.....	28
Facilitation Skills	28
Communication Skills.....	29
Active Listening.....	31
Body Language.....	32
Sensitivity to Adult Students' Cultural Diversity.....	32

Chapter 4: Criminal Justice Training Evaluation Model

Kirkpatrick's Training Evaluation Model.....	34
Customizing and Expanding on Kirkpatrick's Evaluation Model for Criminal Justice Training	37
Conduct Needs Assessment	40
Design Training Plan	41
Develop and Test the Curriculum.....	43
Deliver the Curriculum	47
Evaluate the Training and Trainers and Revise	48

Chapter 5: Project Methodology

Key Decision Processes for Site Selection	65
Overview of Methods	68

Chapter 6: Cross-site Comparisons and Findings

Summary of the Training Evaluation Model's Applications	71
Needs Assessment.....	71
Training Plan.....	73
Develop and Test Curriculum.....	75
Pilot Test.....	77
Trainer Selection.....	78
Training Course Evaluation	79
Conclusions.....	85
Recommendations and Lessons Learned: Tips for Evaluating and Improving Criminal Justice Training.....	86
Costs of Training.....	96

Chapter 7: National White Collar Crime Center's Foundations of Intelligence

Analysis Training

The National White Collar Crime Center	99
History and Background	99
Center Services	100
Review of the Intelligence Literature	103
Intelligence-led Policing	104
National Intelligence Plan.....	105
Core Standards.....	106
Foundations of Intelligence Analysis Training.....	107
Program Overview.....	108
Evaluation Methodology.....	109
Evaluation Questions	110
Data Collection Methods and Framework	111
Study Strengths and Weaknesses.....	114
Evaluation Findings	114
Participant Reaction.....	115
Knowledge and Skills Gained.....	119
Behavior Changes	121
Discussion.....	127
Strengths of the Course.....	127
Recommendations for Change.....	128
APPENDIX 7-A: FIAT Development SME Participants.....	131
APPENDIX 7-B: NW3C FIAT Course Training Evaluation Materials.....	131
APPENDIX 7-C: Pre-Post FIAT Participant Self-assessment of Course Comfort Level.....	145
APPENDIX 7-D: Matched Pairs T-test Results of Pre/Post FIAT Course Comfort Level.....	146

Chapter 8: Simon Wiesenthal Center's National Institutes Against Hate Crimes and Terrorism Training

Simon Wiesenthal Center 148

Review of Hate Crimes and Terrorism Literature 148

 Hate Crime Defined 148

 Statistics 151

 Terrorism..... 152

 Training..... 153

Teaching Tools for Tolerance..... 154

 Tools for Tolerance National Institutes Against Hate Crimes and Terrorism 154

Evaluation Methodology..... 157

 Evaluation Questions 157

 Data Collection Methods and Tools 158

 Strengths and Weaknesses 167

Evaluation Findings 168

 Participant Reaction 169

 Learning/Knowledge Gained..... 178

 Attitude and Behavior Changes 182

 Organizational Impact..... 194

Discussion..... 195

APPENDIX 8-A: SWC Training Evaluation Materials..... 201

APPENDIX 8-B: Case Study of Monmouth County, New Jersey 216

APPENDIX 8-C: Case Study of Madison, Wisconsin 229

Chapter 9: NCLETTTC Advanced Leadership Techniques Training for First Responders, Corrections, and Security Officers

Introduction..... 240

 Context for Evaluation..... 240

 Overview of Literature Relevant to the Training..... 242

 Research Questions..... 245

Method..... 245

 Trainings and Participants..... 246

 Design, Instrumentation, and Data Collection..... 252

Results..... 258

 Level 1 Reaction Results 258

 Level 2 Knowledge Results 266

 Level 3 Behavior Change Results..... 268

 Level 4 Organizational Impact Results..... 279

Discussion..... 286

 Main Findings 286

 Strengths, Limitations, and Recommendations 291

APPENDIX 9-A: NCLETTTC Study Timeline, Milestone, and Workplan Chart.....	300
APPENDIX 9-B: NCLETTTC Training Evaluation Materials	305

Chapter 10: National Judicial College's Civil Mediation Training

National Judicial College.....	330
Literature Review.....	331
Civil Mediation Training	333
Program Overview	333
Evaluation Methodology.....	335
Evaluation Questions	335
Data Collection Methods and Framework	336
Evaluation Findings	340
Participant Reaction	340
Knowledge and Skills Gained.....	342
Behavior Change.....	345
Discussion.....	345
APPENDIX 10-A: NJC Training Evaluation Materials	348

Appendices

Appendix A: NW3C Fiat Course: Instructor Classroom Training Observation Assessment Instrument.....	358
Appendix B: Training Evaluation Model Project Evaluability Questions for Site Selection Screening.....	363
Appendix C: Synthesis Report on Evaluability Assessments of Training Programs.....	365
Appendix D: NW3C FIAT Training Evaluation Plan.....	371
Appendix E: Memorandum of Understanding Between Institute for Law and Justice and National White Collar Crime Center.....	379
References	384

Chapter 1

Project Background and Overview

The purpose of this project was to produce a training evaluation model that can guide evaluations of a wide range of criminal justice training programs. The study was conducted by the Institute for Law and Justice in partnership with Eastern Kentucky University. It was sponsored by the National Institute of Justice (NIJ) with funding from the Bureau of Justice Assistance (BJA). The project's overall goal was to help the Office of Justice Programs (OJP), U.S. Department of Justice, achieve more consistency and control over the hundreds of training programs for which it provides funding, and at the same time, increase the capacity of other criminal justice programs—federal, state, and local—to conduct their own training evaluations.

Project Summary

This study had two major objectives: (1) develop a flexible model for evaluating criminal justice training programs, and (2) test the model by applying it in the field to four training programs. The four programs that were evaluated to test the model had received BJA discretionary grant funding for training (commonly known as “earmarks”). They were selected in part because they permitted a test of the model in diverse environments: the programs were different in terms of learning objectives, intended audiences, instructional methods, subject matter, and other factors. The four participating training programs were

- *Foundations of Intelligence Analysis Training* (FIAT) offered by the National White Collar Crime Center. This was a basic analytical intelligence training curriculum for law enforcement and regulatory personnel.
- *Tools for Tolerance Institutes* offered by the Simon Wiesenthal Center. The purpose of this training was to give participants new perspectives on hate crime and terrorist acts, help them form multi-agency collaborations, and foster the development of strategic action plans.
- *Advanced Leadership for Law Enforcement and Corrections Professionals* offered by the National Corrections and Law Enforcement Training and Technology Center. This course was focused on teaching values-based leadership skills to agency leaders who are responsible for first responders and correctional and security officers.

- *Civil Mediation* training offered by the National Judicial College. This course familiarized participants with the civil mediation process and qualified them for certification in states that require it.

The research study teams began with a widely accepted training evaluation model (Kirkpatrick 1998) that was originally developed for the private business sector. The Kirkpatrick model was then refined to address training evaluation needs in criminal justice. The lessons learned by applying the model in the field were of great benefit in shaping the final model, helping to ensure that it would be applicable to all criminal justice training programs. The final model retains Kirkpatrick's evaluation framework but places greater emphasis on conducting a training needs assessment and on planning for a training evaluation when the training is geared for criminal justice audiences; and it makes changes to the Level 4 application.

Need for a Criminal Justice Training Evaluation Model

Compared to many other topics in the criminal justice evaluation research literature, little attention has been devoted to evaluating training programs. This holds true despite the facts that: (a) high caliber training programs are essential if criminal justice personnel are to perform their duties in a professional manner; and (b) each year, much time, effort, and money are devoted to staff training (Minor, Wells, Cobb, Lawrence, & Cox 2005). To the extent that criminal justice training programs have been evaluated, the focus has been on trainees' attitudes toward training and on the interrelated questions of what type of and how much training to offer (e.g., Brand & Peak 1995; Edwards 1993; Marion 1998; Ness 1991). Few training programs have been evaluated in terms of impact on the knowledge and behavior of program participants or impact on the organizations in which trainees work. There is an abundance of excellent literature on principles of program evaluation (e.g., Rossi, Lipsey, & Freeman 2004), but these principles have not been applied to criminal justice training on any significant basis.

Billions of taxpayer dollars have supported OJP-funded programs, a significant portion of which are training programs for criminal justice professionals. For example, a subset of federal spending on state and local criminal justice efforts in FY 2002 was \$94.5 million in BJA discretionary funds, which Congress had earmarked for 88 specific organizations. Of this amount, approximately \$25 million was earmarked for training grants (the four training projects evaluated in this study were among those FY 2002 grantees).

Evaluations are essential for determining whether OJP-funded training efforts are effective. NIJ has a responsibility to collaborate with other OJP agencies to support such evaluations. It sought an evaluation model that was flexible and practical, yet rigorous enough to guide evaluation planning where experimental or quasi-experimental designs were feasible.

Most training programs do assess participants' immediate reactions—they conduct what Kirkpatrick has called a “level one” evaluation—but far fewer programs or program sponsors are able to answer the more difficult evaluation questions: What specific knowledge, skills, or changes in attitude did participants gain as a result of the training? Were participants able to apply what they learned back on the job? Did their employers see positive changes in their organizations as a result of having invested in employee training? (Kirkpatrick 1998)

NIJ recognized that a more consistent approach to training evaluation was needed both to assist OJP agencies and Congress in making wise funding decisions (avoid funding ineffective training programs) and to assist grantees in conducting meaningful evaluations that could help them improve their training and document effectiveness. Although the Kirkpatrick model (explained in Chapter 4) offered an excellent framework for the planning of training evaluations of differing levels of complexity, it had not been fully explored in the criminal justice context. It was important to determine how such a model should best be modified or expanded to improve criminal justice training and training evaluations.

Overview of the Report

The audiences for this report include sponsors of criminal justice training programs; researchers and evaluators; and training program directors and trainers who may or may not have strong backgrounds in evaluation methodology. The chapter-by-chapter guide that follows is intended to help readers turn to portions of the report that may be of special interest to them.

Chapter 2: Planning for Evaluations

This chapter's purpose is to help “level the playing field” for readers who are not evaluation professionals by providing information about evaluation theory and design. It gives training developers and sponsors some of the tools they need to work effectively with evaluators. This is important because collaboration in the early stages of planning for a training program produces the strongest possible evaluation design and helps ensure that the design can actually be

executed. The chapter can also serve as a guideline for evaluators as they consider how to discuss evaluation planning with their clients.

Chapter 3: Factors That Contribute to Successful Practitioner Training Outcomes

This chapter reviews the importance of applying adult learning concepts to criminal justice curriculum development; discusses important considerations for matching course content with instructional methods and media; and explains how the learning environment contributes to successful outcomes. The chapter is intended to (1) help training developers increase the likelihood that they will achieve their training objectives, and (2) aid both training professionals and evaluators in interpreting evaluation results.

Chapter 4: Criminal Justice Training Evaluation Model

After reviewing the key features of the Kirkpatrick training evaluation model, this chapter explains how the model was enhanced in this project and then presents the complete, revised model. This is a step-by-step discussion of each element of the model. It covers simple assessments of participant satisfaction; evaluations of knowledge, skills, and attitudes learned; more complex and demanding evaluations of behavioral and organizational changes that may be attributable to the training experience; and the often overlooked task of evaluating instructors objectively.

Chapter 5: Project Methodology

This chapter first provides a detailed discussion of the criteria that guided the researchers in conducting ten evaluability assessments and in selecting the four training programs that participated in the evaluation. In addition, it provides an overview of the methodologies employed in each of the four evaluations. More detailed discussions of methodology and related issues are found in the individual evaluation reports (Chapters 7 through 10).

Chapter 6: Cross-site Comparisons and Findings

This chapter summarizes the key features of the criminal justice training evaluation model that were tested; discusses similarities and differences among the four training programs that participated in the project's test of the model; and presents our findings with respect to the applicability of the model and what we learned about outcome evaluations in terms of learning, behavior change, and where feasible, organizational impact. The chapter also includes policy recommendations for improving training and training evaluation and provides lessons learned for

OJP agencies involved in funding criminal justice training and organizations that develop and deliver criminal justice training. The chapter also includes a brief section on comparing costs of criminal justice training.

Chapters 7-10: Evaluations of the Four Training Programs

The next four chapters present the individual evaluations of the four training programs. Each evaluation is documented with background, methods, findings, and discussion. The four programs evaluated were:

- National White Collar Crime Center's *Foundations of Intelligence Analysis Training* (Chapter 7)
- Simon Wiesenthal Center's *National Institutes Against Hate Crimes and Terrorism* (Chapter 8)
- National Corrections and Law Enforcement Training and Technology Center's *Advanced Leadership Techniques for First Responders, Corrections, and Security Officers* (Chapter 9)
- National Judicial College's *Civil Mediation* (Chapter 10)

Chapter 2

Planning for Evaluations

Evaluating criminal justice training programs—like evaluating any program—involves systematically assessing whether a program operates the way it was intended and whether it has produced the intended outcomes. The best evaluations are planned concurrently with the program’s implementation; however, most evaluations are done after the program has been operating for awhile, at its conclusion, or at a future time after the program has ended. There are many different approaches to planning an evaluation of a training program, but the strongest evaluation is one that is planned during the curriculum development phase of the training, with the evaluation taking place concurrently with the training.

Stakeholders in the training evaluation process—for example, funding agencies, associations, training program directors, curriculum developers, trainers, and recipients of training—may not have the same understanding of evaluation methodology that professional evaluators do. This chapter provides these stakeholders with information they will need to communicate effectively with evaluators. The chapter first provides background information on various types of evaluations—their purposes and the questions one could expect to answer when choosing one type of evaluation over another. Next, it provides a detailed discussion of the steps taken in planning an evaluation. It concludes by pointing out some of the opportunities and challenges involved in conducting training evaluations. With this knowledge base in common, evaluators and program personnel can get more out of their joint planning efforts.

Types of Evaluations

Program evaluation is defined by the General Accounting Office (GAO)¹ as “...individual systematic studies conducted periodically or on an ad hoc basis to assess how well a program is working” (U.S. GAO 1998, p. 3). An evaluation should be purposive, analytic, and empirical (Maxfield 2001). That is, its purpose should be known, it should be based on logic, and the results should be based on experience and data.

¹ This office is now the Government Accountability Office.

Evaluation relies on social science research methods to examine whether a program is operating the way it was intended (known as process or implementation evaluation) and whether it has produced the intended program effects (referred to as outcome or impact evaluation). It provides an in-depth assessment of program need, performance, or benefit. Types of evaluations include:

- **Needs assessment**—answers questions about the conditions a program is intended to address and the need for the program
- **Assessment of program theory**—answers questions about program conceptualization and design
- **Assessment of program process**—answers questions about the program activities and operation, implementation, and service delivery (process evaluation or implementation evaluation)
- **Impact assessment**—answers questions about program outcomes and impacts (impact evaluation or outcome evaluation)
- **Efficiency assessment**—answers questions about program cost and cost-effectiveness (sometimes referred to as a Return On Investment (ROI) or cost benefit analysis)

The most common program evaluations examine both the process of a project (how it is being implemented) and the impact of a project (the consequences of the project for its participants). It is possible to conduct a process evaluation of a project (how it was implemented) without measuring the project's impact. However, it is not possible to conduct an impact evaluation of a program without first completing a process evaluation, because to assess the impact of a project, we need to first systematically assess *what is happening* inside the project. For example, if the evaluation finds differing outcomes across project participants, a process evaluation will help indicate whether all participants actually received equivalent services, were served by the same staff, and attended the program regularly.

A process or formative evaluation assesses the fidelity and effectiveness of a program's implementation by focusing on the activities and operations of the program (Rossi, Lipsey, & Freeman 2004). In essence, a process evaluation describes how a project was implemented, how it operates, and whether it is operating as stakeholders intended. Issues commonly investigated by a process evaluation include the following:

- What planning processes led to the application for program funding?
- Who was involved in the planning process? Were any key stakeholders omitted?

- What baseline information is available to document the need for services?
- How do the program activities fit into the larger local system for providing similar services?
- Has staff turnover occurred, and if so, how was the turnover addressed?
- What problems were encountered during program implementation and how were they resolved?

An impact or summative evaluation examines how well a project operates, what happens to participants as a result of the project, whether the project is effective in reaching stated goals, and whether there are any unintended/unwanted consequences. In essence, impact evaluations gauge the impact of a program through changes in participants that can be directly attributed to their program involvement. The evaluator needs a plan for collecting data that will be persuasive in demonstrating that the observed changes resulted from the program and not for other reasons. Issues commonly investigated by an impact evaluation include the following:

- Are program services being provided where none were previously provided?
- What impact did the program have on project customers?
- What impact did the program have on the community?
- What were the unintended consequences of program implementation?

Cost benefit or cost effectiveness analyses are a component of outcome evaluations but are not commonly performed. Cost benefit/effectiveness calculations are used to determine what the cost of a program is in comparison to monetary benefits or effectiveness in changes brought by meeting program goals. These analyses can be difficult to calculate because they require assumptions to be made about the dollar value of program-related activities. Typical questions include:

- Are resources used efficiently?
- Are the program's costs reasonable in relation to the benefits?
- Would alternative approaches yield the same benefits at less cost?
- What is the cost of a strategy in relation to its effectiveness?

Program evaluations are used for a variety of reasons including to assess if a new or innovative program (or project or training) shows promise, if funds are being used wisely, if a program should be expanded, or if evidence is needed to convince others about the merits and/or failings of the program. In practice, evaluation is most often called upon to help with decisions

about improving programs, projects, and components rather than decisions about whether to terminate a program or project. Decisionmakers may start out with global questions (“Is the program worth continuing?”) but they often receive qualified results (“These are good effects, but...”) that lead them to ways to modify present practice.

Evaluation Planning Steps

Planning a program evaluation depends on the specific questions that the evaluation poses.² Before deciding on a plan, an evaluator needs to know the following:

- What the program stakeholders or funding providers seek from the evaluation
- How the results will be used
- Timing, resources, and budget

Before an evaluation can be designed, it is important to decide what type of evaluation is best suited to your goals. That is, what is the purpose of the evaluation? Equally important is determining how the evaluation results will be used. The types of evaluation discussed earlier are shown in Exhibit 2-1. As the exhibit suggests, choosing the most fitting type of evaluation involves being clear on the evaluation’s purpose and the related questions that could reasonably be answered.

Exhibit 2-1: Evaluation Purpose, Questions, and Type³

Evaluation Purpose	Question to Be Asked	Type of Evaluation
Assessment of needs and determination of goals	To what extent are program needs and standards being met? What must be done to meet those needs?	Needs Assessment
Design of program alternatives	What services could be used to produce the desired changes?	Assessment of Program Theory
Review of program operation	Is the program operating as planned?	Process Evaluation
Assessment of program outcomes	Is the program having the desired effects?	Impact/Outcome Evaluation
Assessment of program efficiency	Are program effects attained at a reasonable cost?	Cost Benefit/Effectiveness Analysis

² An understandable evaluation guide and planning steps are presented on the BJA Center for Program Evaluation website at www.ojp.usdoj.gov/BJA/evaluation/

³ Source: Adapted from Rossi, Lipsey, & Freemman (2004, p. 40).

The basic steps in planning an evaluation are discussed below and include identifying program goals and objectives, deciding upon evaluation questions, developing a conceptual framework and logic model, designing a methodology, conducting the evaluation, and communicating the results.

Identify Program Goals, Objectives, and Evaluation Questions

An evaluation begins with the identification of program goals, objectives, and specific evaluation questions. Key stakeholders need to agree on the short and long-term goals of the program (e.g., “train police dispatchers to use new computer system”). While the overall *goals* may not be measurable in specific, quantitative terms, the most clear evaluation findings are based on using specific *objectives* and quantitative language. Objectives are focused, operationalized measures of the goals (e.g., “50 percent increase in the number of police dispatchers using the new computer system by the end of the year”).

Formulating effective evaluation questions is critical to the success of an evaluation. The key question(s) to be answered by the evaluation may relate to program process, outcomes, the links between processes and outcomes, or explanations of why the program reached its observed level of effectiveness. The best questions are those that matter to key decisionmakers and stakeholders, while allowing for results that are useful, interpretable, and complete (Rossi, Lipsey, & Freeman 2004).

Develop Conceptual Framework and Logic Model

A conceptual framework (also known as a statement of theory, theory of program logic, or theory of program action) lays out the connections between the program strategy and tactics and the desired outcomes (Roehl 2002). A logic model is the graphical depiction of the conceptual framework. Developing a conceptual framework and logic model greatly simplifies designing the evaluation because it helps to identify which evaluation questions can and should be answered and which may not be feasible to address.

Care should be taken when identifying a program’s theory of program logic or action to avoid basing a program evaluation on faulty program logic flow (e.g., starting from weak or questionable premises, making too many leaps of faith in program expectations, or being too ambitious in what a program can accomplish using the means at hand). If the logic model is

faulty, valuable resources may be lost on ineffective program activities, and the program will be difficult (or impossible) to evaluate because staff will be collecting data that do not measure actual program relationships.

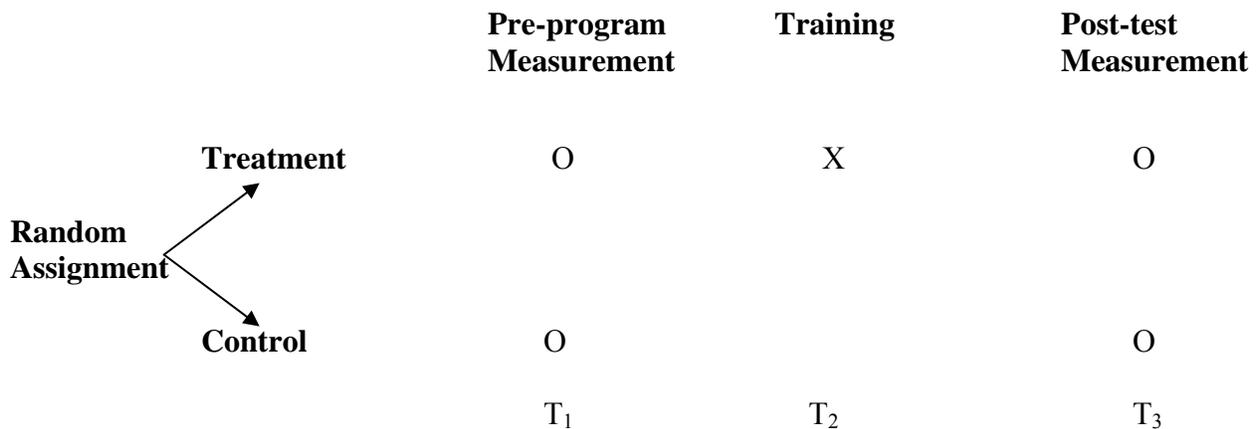
Design Evaluation Methodology

The most important component of an evaluation is the design. The type of evaluation design selected for a particular assessment depends upon the research question(s) being asked, cooperation of program sponsors and participants, resources available, and the time involved. In addition, the type of evaluation design selected is often influenced by the availability of data needed to design a scientifically valid study balanced against the needs of program personnel and recipients (Weiss 1998; Rossi, Lipsey, & Freeman 2004).

There are three types of designs: experimental, quasi-experimental, and non-experimental. In general, the more rigorous the evaluation design, the more confident we can be about the findings. When considering the evaluation design, however, one must take into account the competing pressures of having a sufficiently rigorous design while considering the practical aspects of time, money, cooperation, and the protection of human subjects (Rossi, Lipsey, & Freeman 2004).

The “gold standard” in research design is the experiment (see Exhibit 2-2). This is considered the strongest choice when assessing a causal relationship. In an evaluation using an experimental design, also called randomized control trial (RCT), the evaluator randomly assigns participants to either the treatment or control group. The treatment group receives the intervention; the control group does not. Outcomes are then observed with differences between groups being attributed to the intervention. According to Rossi, Lipsey, & Freeman (2004), the critical element in being able to attribute the group differences to the intervention is ensuring that the control and treatment groups are equivalent in composition and experiences. This equivalence comes from the participants being randomly assigned to those groups.

Exhibit 2-2: Experimental Design



O = Measurements
 X = Intervention
 T = Time period

With sufficient pre-planning, it is possible for training evaluations to use an experimental design. For instance, evaluators can take advantage of “wait-lists” for training and randomly assign half of the list to the training and half as a control group. The control group would not receive the training during the experimental period, but would receive it at the conclusion of the data collection phase of the evaluation. Consequently, the evaluators are able to conduct a rigorous evaluation, and all wait-listed participants are able to take the course as desired.

To conduct an RCT, individuals are not the only entities that can be randomly assigned. Workplaces, schools, or even entire communities can be randomly assigned. For example, in one evaluation of the Drug Abuse Resistance Education (D.A.R.E.) program, entire schools were paired by matching them on a number of factors. One school in each pair was randomly assigned to receive the program, while the other school served as the control (Rosenbaum & Hanson 1998).

However, it can often be difficult to randomly assign participants to a treatment or a control group. In these situations, the next best thing is to use a quasi-experimental design. In quasi-experiments, participants are *not* randomly assigned to a treatment or control group. Instead, the evaluator makes use of a real-life situation to form the groups, such as comparing

two police academy classes. In this example, the evaluator cannot randomly assign police recruits to the academy class but can assume that the classes are reasonably similar, so that a comparison is possible. As in the experimental design, one group receives the intervention and one group does not (Exhibit 2-3).

Exhibit 2-3: Quasi-experimental Design with Pre-post Non-equivalent Comparison Groups

		Pre-program Measurement	Training	Post-test Measurement
No Random Assignment	Treatment	O	X	O
	Comparison	O		O
		T ₁	T ₂	T ₃

Non-experimental designs include both reflexive designs and other types of data collection that typically rely upon qualitative data sources, such as case studies, interviews, and focus groups (see Exhibits 2-4 and 2-5). Reflexive designs involve comparing the targets with themselves (also know as pre-post or before and after designs). While these designs are the most frequently used in evaluation, they are the least rigorous.

Exhibit 2-4: Non-experimental One-group Pre-post Design

	Pre-program Measurement	Training	Post-test Measurement
Treatment	O	X	O
	T ₁	T ₂	T ₃

Exhibit 2-5: Non-experimental One-group Post Design

	Training	Post-test Measurement
Treatment	X	O
	T ₁	T ₂

Collecting data on project measures before and after the project intervention helps assess possible impact of the project. Data collection should begin before project interventions start (the baseline point) and continue throughout the project period and beyond. Measures can be plotted on a graph to show increase or decrease in variables or outcomes over time. Alternatively, pre-intervention and post-intervention periods can be compared (e.g., the number of arrests during a pre-intervention time period and for the same period after the project intervention). Some outcome measures (e.g., fear of crime assessed via neighborhood surveys) will not be available on a continual (e.g., monthly) basis. In this case, such information would be gathered before interventions begin and again after the interventions have been implemented.

Timing of the evaluation is also important. It is much easier to design a program assessment strategy while developing project activities than to implement one afterward. In this way, data collection forms and tasks can be built in from the start rather than collecting data retroactively. A summary of the benefits and trade-offs of each type of evaluation is shown in Exhibit 2-6.

Exhibit 2-6: Elements to Consider When Choosing an Evaluation Design

	Main Feature	Benefits/Trade-offs
Experimental Design	Random assignment of individuals to either treatment or control groups. Groups are usually matched on general demographic characteristics and compared to each other to determine program effects.	The strongest design choice when interested in establishing a cause-effect relationship. Experimental designs prioritize the impartiality, accuracy, objectivity, and validity of the information generated. They allow for causal and generalizable statements to be made about the population or impact on a population by a program.

Quasi-experimental Design	Features non-random assignment of individuals to treatment and comparison groups, as well as the use of controls to minimize threats to the validity of conclusions drawn. Often used in real-life situations when it is not possible to use random assignment.	Prioritizes the impartiality, accuracy, objectivity, and validity of the information generated. However, non-random assignment makes causal and generalizable statements harder to ascertain.
Non-experimental Design	No use of control or comparison groups. Typically relies upon qualitative data sources such as interviews, observation, and focus groups.	Are helpful in understanding participants' program experiences and in learning in detail about program implementation. No causal or generalizable conclusions can be drawn using non-experimental design.

Source: Little, P. (2002). Harvard Family Research Project: Selected Evaluation Terms. Pg. 5. Available: www.hfrp.org.

Conduct the Evaluation

Once the evaluation design has been determined, the next step is to implement the evaluation activities. To conduct the evaluation, it is necessary to develop sources of information for the evaluation measures and then collect the data. Evaluation sponsors and program managers play a crucial role at this stage, including cooperating with the evaluator; discussing nature, quality, and availability of data; and providing access to project materials (Rossi, Lipsey, & Freeman 2004).

The main sources of evaluation data include surveys, interviews, focus groups, program records, and case studies. Each is discussed below.

Surveys

Survey research involves a planned effort to collect needed data from a group of respondents. It is a common type of data collection method in social science research and is used extensively in evaluation. Surveys are most useful when there are no existing credible data sources, and thus the evaluator needs to generate original data. Surveys can be conducted in person, by telephone, mail, email, or fax. Participation in surveys should be voluntary and is often anonymous. If it is necessary to track respondents for data collection purposes, then the survey responses should be confidential.

Surveys allow the evaluator to ask a broad range of questions, including closed-ended and open-ended. Questions should also be kept simple and in multiple-choice format as much as possible, with limited numbers of open-ended questions. When using open-ended questions, it is

important to realize that content analysis of open-ended questions can become arduous and sometimes subjective. Surveys will generally produce both quantitative information (e.g., a score of from 1-5 on effectiveness of an instructor) and qualitative information (e.g., an answer to an open-ended question about the strengths and weaknesses of a curriculum).

There are several key guidelines when constructing surveys (e.g., see Maxfield & Babbie 1995). First, define the concepts that need measurement. That is, decide what information needs to be collected and what evaluation issues need to be addressed. Second, ensure that the questions (and answers if close-ended survey) are clear and unambiguous. Third, keep the questions short to avoid confusion and frustration by the respondents. Fourth, avoid negative questions, such as using the word “not” in a question. Respondents may unknowingly overlook negative words, thus calling the reliability of the survey into question. Finally, avoid using biased terminology.

There are a number of advantages to survey research. It can be more economical than some other types of data collection such as interviews. It is possible to reach a large population with relative ease, thus making collection of large samples feasible. Finally, since all respondents are asked the same set of questions, the evaluator increases the measurement value of the findings. The main disadvantage of survey research is sample bias—not everyone returns surveys and those who do return them may not be typical of the entire program population. To increase response rate, the evaluator should ensure that answer categories represent answers that respondents really want to give, keep the survey short, consider incentives for its return, and follow-up with non-responders. One way to increase response rate is to administer the questionnaire in a group setting (e.g., by collecting a number of participants in one room).

Interviews

Interviews can be face-to-face or by telephone and conducted in a formal, structured format; a semi-structured format; or open-ended. Structured interviews ask each respondent the same questions, in the same wording, and in the same order. Advantages to standardized interviews are that interviewer bias is minimized, data analysis is easier since all questions are identical, and it does not take as much time to conduct structured interviews as semi-structured or open-ended.

Semi-structured interviews are interviews that cover roughly the same set of specifically worded questions that are asked of each respondent, but with some flexibility to probe and ask follow-up questions. The advantage of a semi-structured interview is that it allows for a systematic framework of questions while building in the flexibility needed to explore individual differences.

Open-ended interviews begin with a list of topics to be covered but allow respondents to freely discuss the issues from their own, unique perspectives. In essence, an open-ended interview is a conversation between the interviewer and interviewee. Each open-ended interview will be completely unique. The strength of an open-ended interview is that questions are personalized and the evaluator can respond to both individual and situational differences. Weaknesses include the time involved to collect this type of data, sometimes requiring multiple conversations between interviewer and interviewee, and the potential of bias introduced by the interviewer.

Focus Groups

Focus groups are a technique for collecting data that was pioneered in market research. Focus groups involve participants who are willing to participate in a focused discussion to help evaluators understand the topic of interest (Krueger & Casey 2000). Recruitment of focus group participants is usually determined by a common characteristic that is important to the study. For instance, an evaluation of rape crisis counseling techniques could involve focus group participants who have received the training, or it could involve recipients of the counseling.

Focus groups have two advantages in evaluation. First, they obtain the views of six to twelve people in one place in the same amount of time it would have taken to interview one or two. A second, more important, advantage is that the views of each person are openly discussed in the group setting—allowing for argument, defense, justification, and learning over the course of the session. It is possible to gauge the strength of people’s commitment to their views, the resistance of their views to others’ arguments, and the changes that occur when different positions are aired. Questions posed to focus groups are generally kept to a manageable number (five to ten) and are phrased broadly (e.g., “How successful is this project in serving the needs of our clients?” or “Have we achieved full project implementation? If not, why not?”).

Program Records

Systematically gathering and reviewing program documents can be an important source of data in an evaluation because many programs collect data about and from the people who participate (Weiss 1998). These records can be either historic or current and include administrative, participant, program, or financial records.

A main advantage to this form of data collection is that it uses pre-existing program data in the evaluation. In addition, it can provide a useful framework for collecting additional program data. However, not all program records are up-to-date, accurate, or in a useful format for the evaluation. Sometimes evaluators change existing program forms or introduce new forms for staff to use during the data collection phase of the evaluation. It is important to understand that either changing or using existing program records in an evaluation can place a burden on the program staff who are responsible for generating the records the evaluators are seeking.

Case Studies

A case study is a systematic method designed to assemble information about an individual, group, setting, or event to give researchers insight into how something operates or functions (Berg 2001). The case study allows researchers to observe the “real-life” events of their subjects (Yin 2003). Case studies allow for an in-depth, rather than broad, examination of program results. Case studies often involve different data collection methods, including field observation, participant observation, and in-depth interviews; and review of archival records and other documentation, such as letters, memoranda, communiqués, agendas, announcements, meeting minutes, written reports, and objects (e.g., equipment).

The benefit of a case study is that it enhances understanding of a particular issue. For instance, a case study used to evaluate training programs can help answer questions of *how* training is being used and *why* it is or is not effective. Case studies are useful when it is necessary for the evaluation design to collect in-depth data but impossible for budgetary and practical reasons to collect such data from a large sample. The main disadvantage of a case study is that the findings are not statistically generalizable to other populations.

Analyze and Communicate Evaluation Results

Once evaluation data has been collected, it must be analyzed. According to Weiss (1998), the basic questions to be answered in evaluation analysis include:

- What happened in the program?
- How faithfully did the program adhere to its original plans?
- Did recipients improve?
- Did recipients of program services fare better than non-recipients?
- Was observed change due to the program?
- Did benefits outweigh costs?
- What characteristics of persons, services, and context were associated with success?
- What were the unanticipated effects?
- What limits are there to the applicability of the findings?
- What recommendations can be based on the findings?
- What new policies do the findings support?
- What are the implications for future programming?

Analytic methods should be determined during the design phase of the evaluation.

Quantitative data is analyzed through descriptive and/or inferential statistics. Use of descriptive statistics involves tabulating variables of interest, while inferential statistics allow relationships among variables of interest to be tested. The findings of inferential statistics can be generalized to a larger population. The strength of statistical methods is that the evaluator can draw conclusions in a precise, reliable, and valid way based on the data. One challenge with respect to use of statistical methods is having access to an evaluator with the expertise needed to properly conduct analyses. Another is that the manner in which data is coded and categorized can cause inaccurate conclusions to be reached.

A key, but often difficult, aspect of an evaluation is producing an evaluation report. The evaluation findings will have many different audiences and the results will be important to policy makers, funding agency staff, training providers, instructors, and recipients. Consequently, it is imperative that the evaluation report be thorough, accurate, and useful for the key stakeholders and decisionmakers. The report must be written and presented in understandable terms, not in convoluted or academic jargon. An evaluation report should include a short history of the program, evaluation questions, methodology and measures, copies of data collection instruments, detailing of design strengths and weaknesses, analyses, findings, and discussion. Some reports may include recommendations for change or a section on lessons learned.

Evaluating Criminal Justice Training Programs

Criminal justice training encompasses a wide range of programs—intelligence analysis, tribal justice, gun crime prosecution, victim assistance, leadership, and myriad others. Even so, the programs share some common goals, such as improving performance and professionalism; enabling criminal justice personnel to keep abreast of new laws, strategies, technologies, etc.; and, increasingly, using distance learning to overcome barriers associated with access to quality training.

Opportunities for Control and Comparison Groups

The criminal justice training process naturally lends itself to evaluation. First, training events typically have a clear purpose and a specific set of goals and objectives to achieve. Second, classes of trainees provide a natural and controlled sample. In addition, it may be possible to create a control group by using a waiting list (as discussed earlier) or to use comparison groups found from other training programs. Conducting a successful evaluation of criminal justice training requires the involvement of all key stakeholders, including training providers, instructors, training recipients, and personnel and managers from recipients' agencies. When evaluating criminal justice training programs, it is ideal to plan for the evaluation during the curriculum development phase of the training. That way, the necessary instruments can be designed as a part of the course.

Training evaluations provide key information on what happened in the course, what was learned, and how it was used. These evaluations require a significant time and cost commitment by the training providers, instructors, evaluators, and training recipients. Thus, it is important that the evaluation be designed carefully so that it actually assesses what the stakeholders and decisionmakers want to know.

Challenges in Evaluating Criminal Justice Training

Challenges in evaluating criminal justice training programs include methodological constraints, threats to validity, cost, and the interest and cooperation of key stakeholders. Methodological constraints in evaluating criminal justice training programs include difficulties in designing experimental studies, or even rigorous quasi-experimental studies. Often, evaluators have little control over who can and cannot participate in a training, thus making random assignment impossible. It can be equally difficult securing comparison groups.

Validity is an important consideration when designing an evaluation. In essence, the validity of a study relates to the accuracy of the results. Threats to validity include:

- **Maturation**—natural changes in the participants due to the passage of time that can effect the findings, such as gaining experience in a particular subject area apart from what was learned at a particular training.
- **History**—changes in the environment that can produce changes in the measures under study.
- **Selection**—how participants become part of a study. For instance, intelligence analysts who volunteer for a training may not be “typical” of all intelligence analysts.
- **Testing**—the process of testing and retesting the same participants. This process can influence the behavior of participants and cause them to answer differently on subsequent tests.
- **Mortality**—the dropout rate among participants. Those who drop out of an evaluation may be qualitatively different from those who remain.
- **External validity threats**—the degree to which the findings can be generalized to others.

The goal of an evaluation is to formulate credible conclusions. This process can be hampered by many factors, including the cost of evaluation and the interest and cooperation of key stakeholders. Evaluation is not cheap. It costs money and time to either hire external evaluators or use internal staff members. In addition, collecting data from training participants can be costly in terms of the expense to generate data and time involved on the part of evaluators and participants. If key stakeholders, including training participants, their respective agencies, and training providers and instructors, are not interested in participating in an evaluation, the results will not be useful.

Chapter 3

Factors That Contribute to Successful Practitioner Training Outcomes

Training Objectives

In the criminal justice field, we not only train adults, we train practitioners. This makes the objectives more specific but also more challenging. We are not dealing with young "sponges" who will soak up whatever materials we deliver. The participants in most criminal justice training programs are experienced professionals who work in the fields of policing, corrections, prosecution, courts, and others.

Psychology experts have spent years studying the motivations for adults to learn. Maslow's (1970) well-accepted work on human motivation and a hierarchy of needs postulates that adults have probably mastered their needs for hunger, thirst, and security; and many have attained love and self-esteem. One of the key remaining needs is for self-actualization. This final need can be seen in a person's desire to master a vocation, acquire peak experiences, enjoy a sense of accomplishment, and more. Thus, one of the primary ways to satisfy this need is to learn more and, by extension, participate in more training.

One of the most important aspects of developing a training curriculum is to be clear and accurate about the training objectives. What do we want to accomplish? Do we want to *inform* or *transform*? Bloom's taxonomy describes the lowest levels of learning as rote memory and comprehension—the students understand the information and can recall it later (Bloom, et al. 1956). This teaching involves simply transmitting information from teacher to student. The intermediate levels of learning, as noted by Bloom, include application and analysis, which involve students using logic to identify relationships and solve problems. The highest levels of learning are synthesis and evaluation. In this highest stage, *transformative learning* occurs and students learn to combine knowledge, facts, skills, and logic to make decisions and judgments about the merits of ideas, proposed solutions, or work approaches.⁴

⁴ See an excellent illustration of how Bloom's taxonomy is applied to adult learning in Spencer (1998) reference.

Not all training can involve transformative learning, but training that doesn't should not imply or advertise that it does. There is nothing wrong with training that merely informs students about a new or different subject. The training objectives must be clear and accurate so as not to mislead potential students or establish unreachable outcomes for the training organizations and instructors.

Along the same vein, being clear and accurate about training objectives enables training developers to select appropriate training materials and methods. Training materials and methods should match training objectives. If we are designing a course that is low on Bloom's scale, i.e., information dissemination, then we can use fairly simple materials and common methods such as one-way education (e.g., lecture, PowerPoint presentation, etc.). On the other hand, if the objective involves transformative learning, we need to build in materials and methods that enhance using logic to solve problems and make judgments, e.g., case studies, role play, etc. At this highest level of adult learning, some experts call for what they refer to as *interactive* teaching. Dr. Karen Spencer (1998) refers to *purposeful* teaching.

Training resources also often define the training objectives. It takes more time to deliver transformative learning than informative learning. The National White Collar Crime Center (NW3C) FIAT course attempted to train entry-level intelligence analysts in the fundamental concepts and practices of basic intelligence analysis over 40 hours. This was probably the bare minimum time for a course with these learning objectives. Even in 40 hours, most subject matter experts (SMEs) felt there was not enough time for student exercises. On the other hand, many courses that are designed to deliver *information only* can be presented in one day or less.

Finally, because evaluation hinges on training objectives, it is important to have evaluation in mind when the objectives are written. In addition, training objectives should be revised, if needed. Often, the objectives as originally written are never refined, even when the training changes significantly after some experience with a course.

Adult Learning Concepts

The field of adult learning is a multibillion dollar enterprise that encompasses a wide variety of interests, programs, organizations, and personnel (Merriam & Caffarella 1999). To understand learning in adulthood, it is important to know who the learners are, why they want to learn, and what methods of teaching will best facilitate their comprehension of new material.

When a criminal justice training falls short of achieving its objectives, the problem may be that the planners and developers had an incomplete understanding of adult learning principles, or that they made a poor match between the material to be covered and the instructional methods. Another possibility is that the material and instructional methods were in alignment, but not enough attention was paid to creating a comfortable learning environment. Practical matters such as room temperature, breaks, food and beverages, or seating arrangements are sometimes overlooked by evaluators, but they can have a significant effect on learning or even derail an otherwise excellent training program. This chapter addresses these issues to (1) help training developers increase their chances for successful outcomes, and (2) aid both training developers and evaluators in interpreting evaluation results.

The foundation for developing adult training courses is understanding adult learning principles. There are a number of key aspects of adult learning that, when understood and incorporated into a training program, contribute to more successful training. Examples of some of these principles from McCain (1999) include:

- **Learner directed**—learners are more willing to learn when they understand why they need to know the new information and skills being taught.
- **Active involvement**—active, not passive, involvement in the training is central to helping adult learners gain more from the training experience.
- **Immediate, relevant, and problem focused**—they like the learning to be problem-oriented and they want the training to have an immediate and relevant application to their profession.
- **Evaluable**—each concept and skill should be clearly defined. The corresponding knowledge, skill, or attitude change expected as a result of teaching the new concept or skill should be clearly indicated.
- **Based on experience**—learning is enhanced if it incorporates and builds upon one's facts and experiences. Training should be meaningful to the adult student's life situation. This encourages participation and helps give learners a deeper understanding of the materials being presented.
- **Multiple instructional methods**—people learn differently. Some learn verbally, some by example, and others require a hands-on experience; and still others through writing and reading. By providing a variety of instructional methods to present materials, an instructor can accommodate a wider audience in the learning process.

Equally important for criminal justice training providers to know, as Dr. Spencer (1998) points out, adult learners are also quick to be critical of unprepared teachers, poorly articulated

programs, and individuals or processes that interfere with their learning. They also expect to have their physical needs met (e.g., comfortable furniture, appropriate breaks, refreshments, etc.).

An important piece of the training evaluation model developed in this study relates to implementation of appropriate adult learning concepts. Critical questions include: How has what is known about adult learning styles informed the selection of training modality (e.g., classroom or workshop setting v. distance learning), teaching and learning methods (lecture, case studies, demonstrations, etc.), and instructional media (e.g., computer based instruction, video)? How has knowledge of the specific audience (obtained through job task analysis and/or a training needs assessment) informed the selection of training delivery methods and the curriculum development process? How do modality, methods, and media—and in what combination—result in learning for the target audiences?

Instructional Methods

Educators' emphasis today on matching teaching and learning styles stands in sharp contrast to the traditional scenario of passive listeners attending lectures given by teachers who rarely look up from reading their notes. However, some criminal justice trainers misunderstand or oversimplify adult learning theory. That is, they may select alternatives to lectures (exercises, games, discussions) for the sake of variety without considering their appropriateness for achieving the specific training objectives. Any of these methodologies (including lectures) may be effective and appropriate, but for producing specific types of results under particular circumstances.

Criminal justice training may build in a range of instructional methods, including lectures, structured activities completed individually or in groups (e.g., games or realistic management or problem solving exercises); small group discussions or conferences; case studies; role playing; and simulations. Each of these can be designed to promote learning and participant interaction, although each offers benefits and drawbacks for meeting specific training objectives. An excellent overview of training preparation, instructional techniques, and facilitation tips is provided by Dr. Karen Spencer in her materials on *purposeful* teaching (1998).

The lecture, for example, can be an appropriate and convenient way to present new ideas or facts in a relatively short time, is economical in terms of the material required to support it

(Fay 1988), allows the instructor to control the situation and ensure that specific content is covered (Watson 1979), requires little preparation on the part of learners (Watson 1979), can be easily updated, and is more effective than reading for some learners. Significant disadvantages are that lectures do not actively involve students in learning from doing or reflecting on the experience (Watson 1979); do not allow for differences in knowledge, experiences, and rates of comprehension (Watson 1979); and do not afford opportunities to learn skills such as writing, speaking, analyzing, and others (Fay 1988). Moreover, lectures are effective only for short periods of time, and they are not an effective means of altering behavior or changing attitudes (Watson 1979). As discussed in Chapter 4, this is a particularly important consideration for the training evaluation model refined in this study.

Structured activities can be dynamic and motivating, provide feedback and reinforcement, and allow trainees to learn more about themselves by having their actions analyzed by a group. However, these activities may only reinforce learning for those who enjoy them. Small group discussions can be effective in addressing objectives that do not require presentation of new information, such as increasing awareness of diverse viewpoints (Fay 1988), or correcting misinformation (Watson 1979). Lerda (1967) notes that structured activities such as pooling ideas, sharing facts, testing assumptions, and drawing conclusions "contribute to the improvement of job performance." (p. 155). On the other hand, group members must know the subject and be willing to discuss it; and discussions can easily be dominated by outspoken persons or those with the perceived highest status (Watson 1979).

Case studies are based on the premise that persons who solve problems in the training environment learn to solve problems after training (Fay 1988); one disadvantage is that case selection, preparation, and discussion take considerable time. Role playing offers potential for learning, from the experience itself and from observer reactions, about the effectiveness of various behaviors in particular situations. However, it can degenerate into hurt feelings or silliness and is difficult to implement if trainees are not comfortable with one another and in a comfortable setting (Watson 1979). All of these instructional methods—structured activities, conferences, case studies, role playing—have in common a requirement for trainers with strong skills as *facilitators*.

Many other learning activities can be built into training events, including topic tables (participants meet with each other and experts in a casual environment and discuss specific problems or needs); demonstrations; various types of computer-based instruction; and the matching of participants to mentors. In addition, opportunities for informal technical assistance can be created at key points in the training, permitting one-on-one information exchanges between participants and instructors or each other. This time can also allow sponsors and instructors to "take the pulse" of the participants—Is the training useful so far? Is it meeting their expectations for learning? What else do they need?

Some examples of training methods that enhance thinking skills or help bring about change in attitudes, values, or feelings are offered by Spencer (1998) and Capella (1994):

- **Case study.** A small group analyzes and solves an event, incident, or situation presented orally or in writing.
- **In-basket exercise.** In a form of simulation that focuses on the "paper symptoms" of a job, participants respond to material people might have in their in-baskets.
- **Critical incident.** Participants are asked to describe an important incident related to a specific aspect of their lives. This is then used as a basis for analysis.
- **Observation.** After an individual or group systematically observes and records an event using a specific focus (for example, leadership style, group interactions, instructor behavior), the data are analyzed and discussed (either one on one or in a group format).
- **Role playing.** The spontaneous dramatization of a situation or problem is followed by a group discussion.
- **Simulation.** This is a learning environment that simulates a real setting, with the focus on attitudes and feelings related to the situation presented.
- **Group discussion.** A group of five to twelve people have a relatively unstructured exchange of ideas focused on the attitudes and values they hold about a specific issue or problem.
- **Storytelling.** Participants "tell their stories" about an experience that all or most group members have in common.
- **Metaphor analysis.** Participants construct metaphors that describe, in a parallel yet more meaningful way, a phenomenon being discussed.
- **Game.** Participants take part in an activity characterized by structured competition to provide insight into their attitudes, values, and interests.

- **Exercise-structured experience.** People participate in planned exercises or experiences, usually using some instrument or guide, and then discuss their feelings and reactions.
- **Reflective practice.** Thoughtfully reflecting on one's actions, including the assumptions and feelings associated with those actions, can be done individually or as a part of a small group discussion.

Practical Training Matters

The comfort level of training participants can also be an important influence on training outcomes. An excellent curriculum delivered by top-notch instructors, but with too many participants who are uncomfortable or dissatisfied with details, will not be nearly as well received as a lesser quality training delivered to comfortable and satisfied participants. Creating the right training environment includes being concerned about the following areas (Parry 2000):

- Course pacing
- Cohesiveness of participants
- Comfort level of facilities—furniture, lighting, space
- Length and frequency of breaks
- Meals (provided or if eaten out)—given enough time to eat
- Quality and packaging of course materials—easy opportunity to see audiovisual materials
- Clearly and frequently communicated agenda
- Ability to participate
- Motivation to attend
- Homework assigned

Facilitation Skills

To reach a stage of transformative learning with adult students nearly always involves trainer facilitation skills. The teaching methods used at Bloom's highest learning levels, synthesis and evaluation, often rely on case studies, role play, or other problem-based techniques that require experienced facilitation skills.

This section discusses the importance of trainers using facilitation skills to enhance adult learning. NW3C is fortunate to have a training director with a Ph.D. in education who understands adult learning. She prepared an excellent module to train all the FIAT trainers in

facilitation skills. The other three courses evaluated used experienced trainers and expected that they already possessed facilitation skills. While this may have been true for some of the trainers in our sample, it is certainly not the case in general for most criminal justice training instructors.

The role of the facilitative trainer was also underscored in findings comparing the National Corrections and Law Enforcement Training and Technology Center (NCLETTTC) leadership training between the traditional classroom group and the online group. While admittedly basing these findings on a small sample, the online trainees showed some level of dissatisfaction with the training in terms of a lack of interactive participation opportunities; feelings not articulated by the trainees who received the same training from a live trainer (see NCLETTTC report in Chapter 9).

Effective communication in a learning environment involves three main areas: verbal content, vocal speech, and visual appearance. Content is obvious and is discussed in Chapter 4 of this report. The vocal area includes speaking clearly and articulately, pacing, volume, use of fillers (e.g., "ah"), use of jargon, etc. Visual appearance or "body language" refers to posture, gesturing with hands, making eye contact, facial expressions, and body movement.

An experienced facilitator will focus on all three areas to be effective. In contrast to most trainers, the experienced facilitator will devote as much time to the vocal and visual areas as the content areas.

The *process* of learning is critical for facilitators. Good facilitators focus on the training atmosphere and physical environment and how it can influence student behavior. They look to arrange things to provide for greater participant comfort. They think about such things as room setup, audiovisual impact, distances between tables/chairs, etc.

Communication Skills

One area that separates experienced facilitative trainers from average trainers is the effective use of communication skills. These skills draw in adult learners and get them to participate in the learning. Adult learning cannot occur without active student participation. Some examples of communication skills used by experienced facilitators include the following:⁵

⁵ Some of these examples are drawn from the following: Bens (2000), Hackett and Martin (1993), Schwarz (2002), Rees (2001).

- **Active listening:** Appear genuinely interested in other people's thoughts and feelings. Listen intently. Make eye contact.
- **Modeling:** Practice behavior that they want reflected back and try to make the training experience enjoyable. They try to be non-judgmental and watch their nonverbal messages.
- **Summarizing:** They use paraphrasing as a method of clarifying. (Check the perceptions of the group. For example: "Please correct me if I'm wrong, but I think Amy's comments summarized our last 10 minutes quite well by stating. . ."). Always summarize at the end of key parts of modules.
- **Focusing attention and pacing:** They keep the students on the topic and focused, using care to limit or reduce repetition.
- **Recognizing progress:** They acknowledge student progress, e.g., "Nice job! We just brainstormed 36 items in our 5-minute time period."
- **Waiting or Silence:** They practice the *Tao* principle that sometimes the hardest thing to do is nothing.
- **Scanning/Observing:** They attempt to nurture full participation from the group. They watch the groups' nonverbal cues in the form of body movement, facial expression, and gesture (which may indicate loss of attention, confusion, or discontent)—then they take a break, change the pace, change the topic, etc.
- **Inclusion:** They make sure everyone has an equal opportunity to participate and encourage those who have been silent to comment.

Facilitative teachers also use group processing techniques effectively in the training environment. Some of the more common techniques include (1) brainstorming and (2) force field analysis—Kurt Lewin's method for identifying "restraining" and "supporting" forces related to a given problem, policy, position, or issue (Lewin 1997). Some additional techniques noted by Spencer (1998) include the following:

- **Response cards:** Pass out index cards and request anonymous answers to your questions. Have the index cards passed around the group. Use response cards to save time or to provide anonymity for personally threatening self-disclosures. The need to state your answer concisely on a card is another advantage.
- **Polling:** Design a short survey that is filled out and tallied on the spot, or poll students verbally. Use polling to obtain data quickly and in a quantifiable form. If you use a written survey, try to feed back the results to students as quickly as possible. If you use a verbal survey, ask for a show of hands or invite students to hold up answer cards.
- **Subgroup discussion:** Break students into subgroups of three or more to share (and record) information. Use subgroup discussion when you have sufficient time

to process questions and issues. This is one of the key methods for obtaining everyone's participation.

- **Learning partners:** Have students work on tasks or discuss key questions with the student seated next to them. Use learning partners when you want to involve everybody but don't have enough time for small-group discussion. A pair is a good group configuration for developing a supportive relationship and/or for working on complex activities that would not lend themselves to large-group configurations.
- **Whips:** Go around the group and obtain short responses to key questions. Use whips when you want to obtain something quickly from each student. Sentence stems (e.g., "One change I would make in community policing is...") are useful in conducting whips. Invite students to "pass" whenever they wish. To avoid repetition, ask each student for a new contribution to the process.
- **Fishbowl:** Ask a portion of the class to form a discussion circle, and have the remaining students form a listening circle around. Bring new groups into the inner circle to continue the discussion. Use fishbowls to help bring focus to large-group discussions. Though time consuming, this is the best method for combining the virtues of large and small group discussion.

Active Listening

One of the key areas from the above communication skills list is active listening. These are techniques that can be taught to facilitative trainers. Some examples are illustrated below:

- Maintain good eye contact.
- Face the person or group head on.
- Listen for feeling as well as content; hear what is "between the lines."
- Don't confuse content and delivery. Assume the person has something to say even if she or he is having trouble saying it.
- Cultivate empathy. Try to put yourself in his or her place.
- Refrain from evaluating what is being said.
- Don't jump in the conversation too soon. Let people finish what they're saying.
- Pause a few seconds before giving feedback or answering a question. Take time to think about what was said.
- Show encouragement. Use simple gestures or phrases to show you are listening, e.g., say "uh-huh;" nod your head.
- Show support; say, "That's good; anyone else having anything to add?"
- Don't let people ramble. Try to help them come to their main point.
- Ask questions beginning with the words "what" and "how" (open-ended questions). Avoid questions that can be answered with a yes or no.

- Don't "jump ahead" to complete the person's sentence; you are making an assumption that may be incorrect.
- Be aware of your own emotional response to what you are hearing. It will affect how well you understand and can respond.

Body Language

Another area of critical importance for effective facilitation is visual appearance or body language. While we generally think of verbal skills as the most important facilitation skill, the role of nonverbal cues or body language is also critical to facilitative training. In training, these nonverbal messages are constantly flowing from student to facilitator and vice versa.

The experienced facilitator will be careful not to send out nonverbal cues or body language that can be interpreted as negative by the receiving audience. For example, standing up leaning against a wall with your arms crossed tends to suggest a closed mind or inattentiveness. This type of body language subtly inhibits the free flow of communication. Facilitators must also be keenly aware of the nonverbal cues given off by students whom they are training. Such cues can often be important indicators to test the pulse of the class.

Sensitivity to Adult Students' Cultural Diversity

The experienced facilitative trainer today has to be aware of the cultural diversity of the adult student participants. Some of the skills in this regard include the following:

- Cognitive and behavioral flexibility: able to adjust expectations and learning activities as the diverse needs, learning styles, and responses to training activities become manifest.
- Cultural self-awareness: understands the role of culture in the formation of one's own values, beliefs, patterns of behavior, and the like; awareness of one's own uniqueness as well as one's similarity to the prevailing cultural norms; aware in this cultural sense and able to teach this concept to others.
- Interpersonal sensitivity and relations: very adept at interpersonal relations and especially sensitive to the needs and concerns of learners; must be able to relate well to the wide variety of individuals who comprise the learner community.
- Tolerance of differences.
- Openness to new experiences and peoples: communicates that authentic openness to learners in patterns of thought, feeling, and action.

- Empathy: able to project oneself into the mind, feelings, and role of another; have the capacity to sense how the learner is doing and to respond appropriately; appreciate the learner's anxieties and difficulties as well as sense of accomplishment.
- Sense of humility: has real respect for the complexities, challenges, and uncertainties of cross-cultural learning; appreciates that training is not a perfect science, and that creativity in orientation design and technique is still possible and desirable; and has a deep respect for the intricate and varied nature of cultures.

Chapter 4

Criminal Justice Training Evaluation Model

Evaluating training requires a systematic and sound methodology to determine if the training is effective. This chapter addresses how to evaluate the effectiveness of training. It first explains the training evaluation model that was the starting point for this study. Next, it presents the expanded version of the model, tailored to the criminal justice environment.

This expanded model was developed by creating a preliminary version in consultation with a team of criminal justice training experts, conducting evaluations of four national criminal justice training programs using the model, and refining the model based on our experiences in the field to produce the final model presented here.

Kirkpatrick's Training Evaluation Model

Over 40 years ago, Donald Kirkpatrick (1998) developed a model for conducting evaluations of training programs. This model identifies four levels that successively build upon each previous level: (1) reaction, (2) learning, (3) behavior change, and (4) results (see Exhibit 4-1 below). Each level is described in more detail below.

Level 1—Reaction. This refers to determining how well participants *liked* a training program through the use of rating sheets distributed at the close of a training event. In essence, this is measuring “customer satisfaction” with the training. For the training to be effective, according to Kirkpatrick (1998), it is important that participants have a positive view of it. The majority of trainings conclude by distributing a rating form to participants, though the results are not always used to improve the training. The importance of measuring reaction includes the feedback it provides from the participants, the “message” it sends to participants that their opinion counts, and the quantitative information it affords training providers (Kirkpatrick 1998).

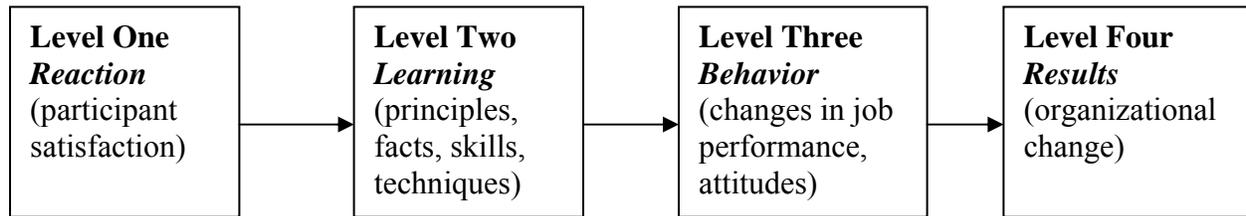
Level 2—Knowledge or Learning. This involves measuring participants' changes in attitude, improvements in knowledge, or increases in skills or techniques. Measuring learning is important because changes in behavior cannot occur if learning has not taken place. The key is to design evaluation instruments that can help ascertain what the participant learned as a result of the training. Measuring changes in attitude or increases in knowledge can be manageably tested

through paper-and-pencil tests (before and after the training), while skills can be measured through performance tests.

Level 3—Behavior Change. This level refers to the transfer of learning at the individual level—changes in an individual's job performance that can be attributed to the training. Essentially, a Level 3 evaluation wants to know what the training participants did (or did not do) once they returned to their jobs. Level 3 evaluations are more difficult and expensive to conduct than Level 1 and Level 2 evaluations because they consider factors that are not as straightforward to measure. For instance, measuring behavior change requires that the individual (1) learned the knowledge or attained the skill at training; (2) has the opportunity to apply this new learning or skill back on the job; (3) has incentives for applying the new learning or skill and continues to do so rather than slip back into old behaviors; and (4) can attribute the change to the training course (Kirkpatrick 1998).

Level 4—Results. At this level, the analysis shifts from changes observed in individuals to the impact on the trainee's organization. In particular, what are the measurable organizational results that can be attributed to the training? For example, if we trained an individual to operate a new radar device, are more cars now being stopped for speeding than previously? Conceptually, some have expanded Level 4 to include measuring return on investment (ROI) and cost benefits, while others consider ROI analysis separately from Level 4 (e.g., Phillips 1996).

Over the years, Kirkpatrick's model has remained the preeminent model for developing training evaluations; it is still the most widely referenced model in the training evaluation literature. It serves as a springboard for discussion in many publications by associations for human performance technology and training professionals in private industry (e.g., American Society of Training and Development (ASTD), International Society for Performance Improvement (ISPI)); in training evaluation references provided by the Office of Personnel Management (OPM 2002); and in literature encouraging educators and health care trainers to evaluate training impact.

Exhibit 4-1: Overview of Kirkpatrick's Model of Training Evaluation

There is a shifting of conceptual gears between the third and fourth levels in Kirkpatrick's framework, with the first three levels focused on the trainee and the fourth on the organization (Nickols 2000). Kirkpatrick did not provide a formal definition of results but discussed how the organizations that are paying for the training might benefit from it (e.g., reduced costs, turnover, or absenteeism; increases in quality or quantity of production; increased profits). As important as information on results would appear to be for organizations, evaluation is seldom conducted at this level—even in private industry—primarily because it is complex and expensive. It is not surprising, then, that few criminal justice agencies do so. Nevertheless, there are indications that such efforts can be worthwhile. For example, a training evaluation by Eastern Kentucky University found a high turnover rate (40 percent) among police academy graduates in Kentucky after one year of service as police officers. After the academy intervened by implementing a mentoring program for newly hired officers, the turnover rate was reduced to 20 percent (Minor *et al.*, 2002). It is quite possible, though, that some criminal justice agencies have limited expectations with respect to a training's impact on the organization, especially if the training lasts only a few days.

Over the years, Kirkpatrick has modified his guidelines somewhat, and many others have expanded on them (Nickols 2000; Phillips 1996; Brown 1997) or proposed alternatives (Holton 1996). However, the concept of the four levels has remained constant, and Kirkpatrick remains the most widely referenced model in the training evaluation literature. No doubt Kirkpatrick's model has endured in part because it is easily understood by both training providers and purchasers of training. It has been used and tested more than any other model, and it is easily accessible and malleable to a variety of different types of training program evaluations. The model needs expansion because as Kirkpatrick himself notes, "The model doesn't provide details

on how to implement all four levels. Its chief purpose is to clarify the meaning of evaluation and offer guidelines on how to get started and proceed." (Kirkpatrick 1998, p.3).

Customizing and Expanding on Kirkpatrick's Evaluation Model for Criminal Justice Training

Kirkpatrick's model provides a good starting point for evaluating training, especially training in the business community for which this model was created. For criminal justice, however, our experience suggests that planning for the evaluation of training outcomes using Kirkpatrick's or any other model is at the beginning stages, and that Level 4 evaluations of organizational impact and benefits compared to costs are almost never attempted. There were signs of change, however, shortly before this project began. For example, *Police Chief* magazine featured one article on the evaluation of training outcomes in law enforcement (Bumgarner 2001); and another article explained how the Los Angeles Police Department was applying Kirkpatrick's model (with assistance from a UCLA professor) as it undertook a massive restructuring of in-service training and development of a training evaluation design (Murphy & Gascón 2001).

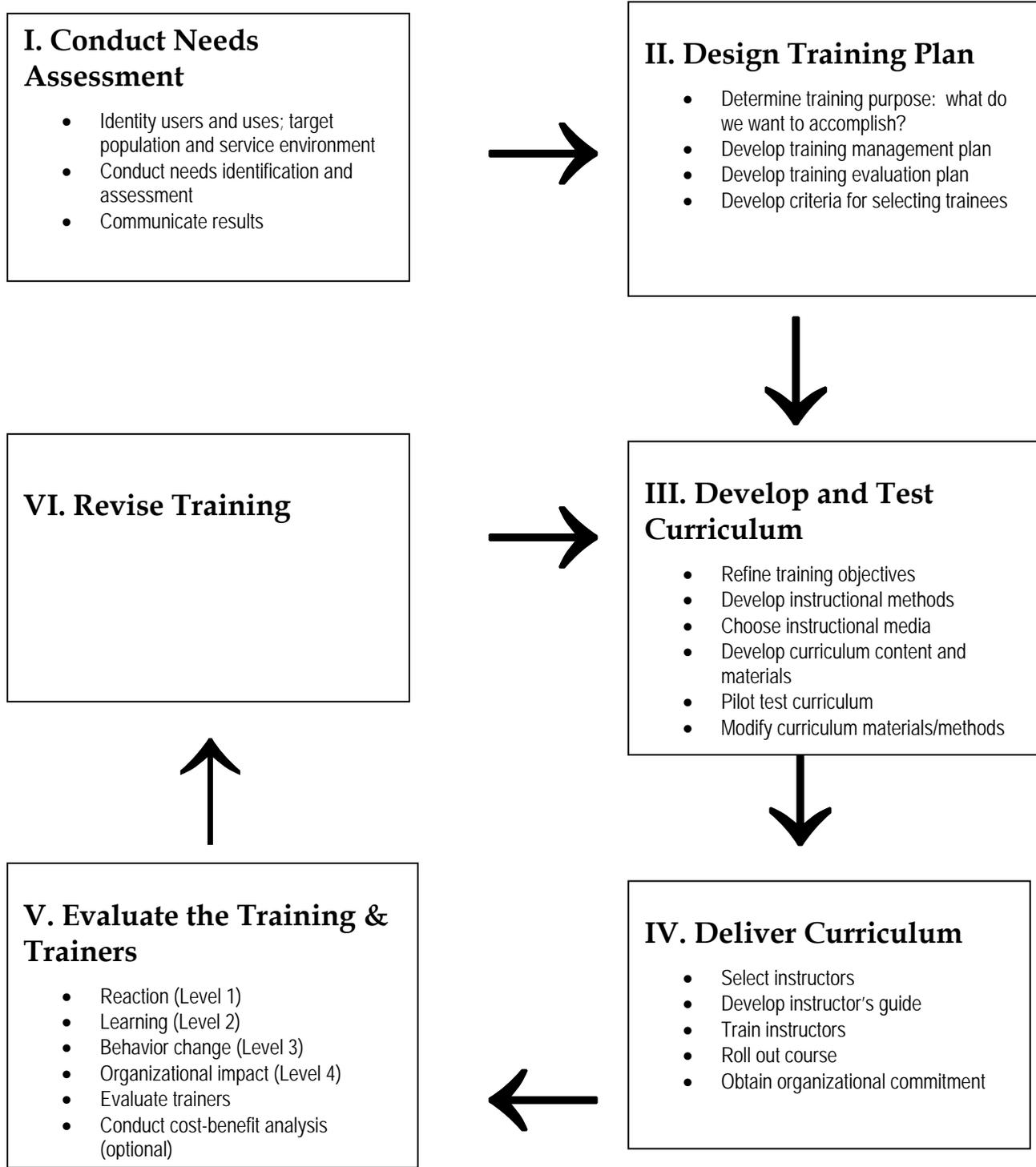
A training evaluation model for criminal justice training needs to be both comprehensive and flexible. It should be rigorous enough to help criminal justice organizations evaluate not only participant satisfaction but also learning, behavior outcomes such as improved job performance, organizational results, and ROI. At the same time, to be useful to the field, the model must be flexible. For example, an agency that simply does not have pre-training data or did not conduct a training needs assessment should still be able to implement portions of the evaluation model and get useful results. The evaluation process should be the same, whether an organization conducts the training evaluation itself or hires an outside consultant.

In this section of the chapter, we present our version of Kirkpatrick's model to evaluate criminal justice training programs. In Chapter 6, we present findings of how the model was applied to evaluate the four selected training programs. As shown in Exhibit 4-2, we have expanded upon Kirkpatrick's training model by including several stages of training development:

- Conduct needs assessment
- Design training plan
- Develop and test curriculum
- Deliver curriculum
- Evaluate the training and trainers
- Revise the training

It is important that evaluation planning and data collection be built into the earliest stages of training development. Note that there is some overlap in the training development stages. For example, development of a training evaluation plan is included in the second stage; however, the details of the evaluation design would also be dependent on formulation of measurable training objectives (noted under stage three).

Exhibit 4-2: ILJ's Expanded Training and Evaluation Model



Conduct Needs Assessment

A needs assessment for training is the first step in designing a training program. A documented needs assessment also aids the evaluator in assessing the justification for the training program. The needs assessment collects information from various sources on what training is needed and how it could best be delivered. A needs assessment can also be used to determine if an established training is meeting the needs of an organization or agency. A training needs assessment examines desired or actual behavior and involves such techniques as surveys, observations, interviews, and document reviews.

No federal training funds should be provided to any organization without a thorough and complete training needs analysis that clearly justifies the need for the training.

One example of needs assessment analysis involves a five-step process (Rossi, Lipsey, & Freeman 2004) that includes the following:

- Identify the users and uses of the analysis
- Identify the target population or service environment
- Conduct a needs identification
- Conduct a needs assessment
- Communicate the results

First, identify the users and uses of the analysis. This step clarifies why the analysis is being done and addresses its purpose. For example, is the analysis being conducted to determine whether training on a particular subject should be created, or is it to determine whether an existing training is still meeting participants' needs? The users of the needs assessment analysis include the key stakeholders: federal agency providing the training funds, organization developing and delivering the training, and agencies that will be paying to send students to the training.

Second, identify the target population or service environment. For new courses, identification of the target population consists of those who will need to take the training. For courses already developed and being reviewed, the target population consists of those who are taking the training and those who should be taking the training. It may be that some individuals need the training but are not receiving it and some are taking the training but do not need it. This analysis would identify those issues.

Third, conduct a needs identification. This involves identifying the underlying problems or issues and reasons training is needed. For instance, if a police department is switching to a new crime mapping system (geographic information system), the underlying problem and reason for training is to teach the department's crime analysts how to use the new system. The NW3C conducted a national review and found that none of the existing intelligence analyst training programs met the needs of beginning analysts. This led to their decision to develop the Fundamentals of Intelligence Analysis Training (FIAT).

Fourth, conduct a needs assessment. The needs assessment should produce solutions and recommendations for action. In the above example on the new crime mapping system, for instance, one solution for action might be to have the software manufacturer train the crime analysts. Another might be to have senior analysts learn the new system and then teach it to their subordinates. In the NW3C program example, the organization decided to develop the FIAT course through a consortium to meet the intelligence analyst training needs.

In developing training to deliver a skill or do a particular job, curriculum developers often need to collect background information in the form of a job task analysis that identifies the knowledge, skills, and abilities needed to perform the job. If a job task analysis is required, it would be documented in this stage but actually conducted in the curriculum development stage.

Finally, the fifth step is to communicate the results to the key stakeholders. This should be in the form of a report that conveys recommendations for action.

Design Training Plan

Creating a training plan is the second key step in designing a training program. The training plan can be thought of as a management plan for several distinct phases of training—goals, management/administration, and evaluation (Parry 1997). This plan should be available for the evaluator to review. In training that is just starting, an evaluation plan should be developed as part of this stage.

The starting point for developing any training is to identify, articulate, and document our goals—what we want to accomplish. How will the trainees think or function differently after the training? This part of the plan uses the results of the needs assessment stage.

Management and administration include developing the foundation or building blocks to develop the training, acquire the trainers, acquire the technology needed, advertise and invite participants, manage the logistics (training facility, food/beverages, housing, etc.), maintain training records, and more. This part of the plan includes establishing and projecting the costs of the training and resources required to deliver and manage the training. This report contains a more detailed discussion on training costs in Chapter 6.

The evaluation phase includes the planned evaluation of the training and the trainers. Planning for an evaluation of the training and the trainers during the curriculum planning stage provides for a more thorough and effective evaluation. During this stage, the funding agencies and training providers commit to funding the evaluation and collecting data to meet the evaluation requirements.

Based on this study, the team also observed two additional issues that are important to successful training but are often overlooked. They include:

- **Developing criteria for selecting trainees:** Some training is shot-gunned to the criminal justice field—whoever applies gets the training (because so often it is paid for by the federal government and delivered by contractors with a vested interest on reporting the "head count" or number of people trained). This leads to situations where personnel receive training that they never apply back in their department. For example, a patrol officer is trained in crime scene investigation but, once back on the job, the officer is never called to the scene of crimes to collect forensic evidence. Thus, the training participant criteria, in this example, should have specified that it was only offered to active crime scene investigators, so as not to waste training resources on a participant who never applied the training. The training criteria should be clearly stated in the training advertisements and the training organization should screen the applicants to ensure that they meet the criteria.
- **Obtaining organizational commitment:** This recommended key step, which is applied in Stage IV—Deliver Curriculum—stems from the above step. For each participant who applies to the training, the training organization should require the CEO, or a senior manager, of the participant's department to sign a statement

verifying that the participant will be in a position to apply the training back in the department. As noted earlier, much of the training in the criminal justice field is free to the participants (provided by federal funds). However, the participants' organizations do have to absorb costs associated with the trainee's salary and sometimes expenses related to travel, hotel, etc. This requirement would ensure a greater degree of vested interest on the part of the trainee's organization to apply the acquired training. At the application stage, trainee organizations should also be required to indicate their commitment to participate in follow up evaluations of the training.

In the earlier needs assessment, the training organization identified the need for training, existing gaps, target population, and recommended approach. In this stage, key questions that need to be answered in the creation of a training plan include (see Parry 1997, p. 13-14):

- What is the specific purpose of the training?
- What factors help or hinder the training purpose?
- What outcomes for the trainees are desired?
- What are the projected outcomes for the trainees' organizations?
- What is the projected life cycle of the training before it becomes dated or everyone has been trained?
- How long does the course need to be?
- What are the key topics and course content?
- Who has the expertise to develop the course?
- What are the costs for course development, instructors, and training?
- Can we conduct the evaluation in-house or contract out—what will it cost?

Develop and Test the Curriculum

Designing a training curriculum requires time, expertise, and commitment. It is important that the purpose of the training be clearly identified so that the training objectives can be refined, specific curriculum materials developed, and instructional methods and media designed. This stage involves the following steps:

- Refine learning objectives
- Develop curriculum content and materials

- Develop instructional methods
- Choose instructional media
- Pilot test curriculum
- Modify curriculum materials and methods

Develop the Curriculum

The curriculum development stage combines all the information developed in the previous two stages—needs assessment and training plan. We enter this stage with the following information available: the identification of need, purpose of the training and projected outcomes, target population, length of training, sources of funding, management plan, and evaluation plan.

Learning objectives. The first task in developing the training curriculum is to refine the learning objectives. The training organization's administrators and training developers have already agreed on the overall purpose and broad goals. For example, in NW3C's FIAT training, at this point they knew why they wanted to do the training—to train more inexperienced personnel to conduct standard intelligence analyses. The key task at this stage then becomes to define in more specific terms: what exactly should the participants be able to do? The learning objectives should include performance indicators (McCain 1999). The performance indicators identify the key measurable skills that participants should be able to do as a result of the training. This involves using action words such as: demonstrate, list, build, classify, conduct, or define. Dr. Spencer's monograph (1998) contains excellent examples of learning terminology aligned by Bloom's taxonomy.

The precise terms of the learning objectives are very important because they will be used to develop the details of the subsequent evaluation questions. For example, eventual test instruments (e.g., before and after knowledge or skills tests, follow up surveys, etc.) will be measuring knowledge gained in terms of the learning objectives.

Curriculum content and materials. The second task in this stage is to develop curriculum content and materials. This is where the actual content of the training is developed and documented. The course content needs to directly link to the learning objectives.

Course content is vital to the success of a training because “competent curriculum design is essential to training excellence” (Glenn, Panitch, Barnes-Proby, Williams, Christian, Lewis, et al. 2003, p.56). Key considerations for designing a curriculum are: purpose, content,

organization of content, and format. There are three main ways a curriculum can be organized (Glenn, et al. 2003):

- **Discipline-based**—focuses on rote learning by presenting information in discrete blocks, usually separated by subject matter.
- **Interdisciplinary**—focuses on transactional learning by encouraging participants to discover relationships and develop applications across common themes.
- **Transdisciplinary**—focuses on transformative learning by organizing the entire curriculum around a set of themes, skills, and problems.

Curriculum content can be divided into three areas: introduction, body, and summary (McCain 1999). The introduction should cover the importance of the course, the benefits participants can expect to receive, and what the training objectives are. The body of the course should begin with the underlying theory and prerequisite knowledge for the course, followed by the new skills, attitudes, or knowledge being taught. The summary phase of the curriculum reviews the training objectives and links what was learned or taught with these objectives.

With skill-based training, establishing learning objectives and developing the curriculum may require a job task analysis. That is, the training developers need to study and document the actual knowledge, skills, and abilities involved in what is planned to be taught (e.g., in the NW3C course, the job of entry level intelligence analyst). Documenting these performance activities, through interviews, surveys, and observations, leads to development of training to teach the knowledge, skills, and abilities required to perform the job.

This phase of curriculum development is often assisted by SMEs. After developing a detailed outline, the SMEs flesh out general topics and subtopics. This leads to identification of individual training modules. The development of specific training modules is often assigned to a specific SME with the requisite expertise. The SME will develop the contents of the module—often including teaching points, backup materials (e.g., journal articles, studies, and other exhibits), and more. The course project manager will then merge all modules—edit and reformat to make them all consistent—and produce the final curriculum.

Develop instructional methods. Another task in this stage is to choose the range and variety of instructional methods that would apply to the training. Instructional methods—the adult learning strategies used to teach criminal justice practitioner participants the learning objectives—can include lecture, role play, case studies, facilitated discussion, problem solving

through teams or individually, simulations, peer-to-peer teaching, and many other examples. There are many excellent texts and books that have been written on this subject, a few are cited in the reference section of this report, and Chapter 3 contains a fuller discussion of adult teaching methods and facilitator skills.

Choose instructional media. It is also important in this stage to select media that will be helpful in the learning process. Well-chosen media can hold trainees' attention and help communicate the learning objectives. Media include PowerPoint, overheads, white or chalk board, flip charts, video/DVD, and many others. Selection of media depends upon availability of equipment, skill and expertise of curriculum designers and instructors, setup of the training facility or classroom, time available, and other factors.

Pilot Test and Revise the Curriculum

Testing and revising the curriculum are two of the most important steps in training development. The pilot test of the training course can involve one or more "practice runs," depending on funding. Training pilot tests are either conducted in front of the SMEs and other invited experts and managers from the training organization or they are delivered live to a group of participants that emulate the target participants. Some organizations with tight budgets simply call the first roll out of the actual training the "pilot test." With adequate funding, the preferred method is to test the course in front of SMEs and others who serve as a "sounding board" or review panel. This is similar to Hollywood movie producers screening a movie in front of a live (but select) audience.

An important issue for consideration is who delivers the material in the pilot test. It may be the actual module developers, who are often the SMEs. At this stage, we may not have selected the eventual trainers. Under the best of circumstances, the SME module developers will end up being the trainers; however, that is not always the case. Some people are excellent at developing materials but don't have the personality or facilitation skills to be an effective trainer.

Pilot testing validates the training content; provides an opportunity to test Level 1 and Level 2 evaluation instruments; and allows the curriculum materials to be checked and double-checked for thoroughness and errors. It also helps get the timing down for how long it takes to deliver a module or conduct an exercise. It provides feedback on the clarity and usefulness of audio-visuals and other media. The results of the pilot test should be used to modify the

curriculum materials. The training organization is then ready to select trainers, train them, and roll out the course.

Deliver the Curriculum

After course materials have been tested and modified, the next stage of the training development process includes rolling the course out to the field. The number of times a course is delivered depends on the availability of attendees, trainers, and funding. This stage also includes selecting and training the instructors. It is important to document any changes in the curriculum or delivery, since this will affect the evaluation. This is also the stage where, as discussed earlier, we suggest that training organizations obtain organizational commitment in the training application and registration process.

Select Trainers

Trainers can come from many sources and may include practitioners working in the field who are knowledgeable about the training subject, professionals from the organization that developed the training, or expert training consultants hired specifically to conduct the training. The choice of trainer depends upon the needs of the training organization, the type of training being conducted, and funding resources. In addition to training budget and scheduling considerations, criteria for selecting trainers include:

- Depth of knowledge of specific subject matter
- Knowledge of the field (e.g., police, courts, corrections, etc.)
- Years of experience in the field
- Years of experience as an instructor
- Reputation and assessment of credibility
- Facilitator skills—verbal and nonverbal communication and listening skills
- Ability to teach/train

The hiring of paid consultant instructors to provide the training depends upon the training provider's budget; number of trainings to be provided; and, most importantly, finding qualified instructors. If the training provider finds a consultant who meets the necessary criteria for experience and has the budget to support this option, then this can be an excellent way to proceed. If the consultant is a training professional who is in the business of facilitating training, the training provider should get a quality product for the money spent. These individuals are often experts in applying adult learning principles to training.

If the training provider has scheduled numerous trainings across the country, however, hiring outside consultants may be too expensive. In this case, providers often train staff within their own agency or practitioners in the field to conduct the trainings. However, the old adage "you get what you pay for" certainly holds true in the training field. Regardless of whether the trainers are outside professionals or local practitioners, they should be paid to deliver the training. The NW3C tried to deliver the FIAT course, which lasted five days, with volunteer practitioners who were donating their time. They ran into a number of conflicts with scheduling, trainers not working out, and more. We think you simply don't get the best trainers if you don't pay them.

Training the Instructors

Even the most qualified trainers should be "trained" in the subject matter of the training course and provided refresher training in facilitation skills. If the SMEs who developed the training modules are the trainers, this step may not be necessary. This step often involves having new instructors observe other instructors deliver the training. A key tool in training new trainers is to develop an **instructor's guide**. This tool ranges in the degree of direction from providing minimal tips to providing elaborate step-by-step directions on what to say and do, when to do it, how to do it, etc.

The instructor's guide, sometimes referred to as a lesson plan, is designed to give directions in how to teach the training module. The guide often includes the following:

- Amount of time to allow for module, exercises, breaks, and other activities
- Background material to discuss with students
- Instructions on the main teaching points
- Types of media to use to illustrate points
- Types of instructional methods and facilitation tips (e.g., questions to be asked of class and suggested answers)
- Explanations and timing for exercises
- Suggestions and methods for debriefing exercises

Evaluate the Training and Trainers and Revise

Evaluating the training and the trainers is an important, often neglected, part of the training development effort. The evaluation serves to inform the key stakeholders—funding

agencies, training providers, instructors, and potential users—on the extent to which the training goals and objectives are being met.

Evaluating at the Four Levels

Ideally, a comprehensive (four-level) evaluation should be designed during the planning stages of the training (i.e., during Stage II—Design Training Plan). This can sometimes be challenging for several reasons: (1) the training could already be running before funding agencies request an evaluation; (2) funds may not have been dedicated to an evaluation; and (3) the Level 4 evaluation is often difficult—measuring results in terms of the impact of the training on trainees’ organizations is always a major challenge.

However with careful planning, cooperation, and adequate funds, most evaluations should be able to complete Level 1 through Level 3, and at least attempt Level 4. Each of the levels is discussed below. In Chapter 6, we present findings on how each level was applied to evaluate the four selected criminal justice training programs. See Exhibit 4-3 below for a summary illustration of evaluation levels, goals, questions, data, and sources.

Exhibit 4-3: Evaluation Questions, Data Collection Tools, and Sources

Evaluation Level	Goal	Evaluation Questions	Data Collection Tools	Data Sources
Reaction	Determine immediate reaction to the training provided.	What do participants think of the training—how satisfied were they? Was something missing or confusing? How engaged were the participants?	Satisfaction/ reaction survey with open-ended comments	Recipients of the training Instructors Observer SMEs
Knowledge	Assess if trainees can demonstrate that they acquired learning and can perform the intended skills.	What information and skills were gained?	Survey administered before and after training Performance tests Demonstrations	Recipients of the training

(Exhibit 4-3 continued, next page)

<p>Behavior Change</p>	<p>Ascertain if trainee's behavior changed as a result of the training.</p> <p>ay also want to assess attitude change</p>	<p>Have trainees transferred knowledge, learning, and skills to their jobs?</p> <p>Have trainees maintained attitude change over time?</p>	<p>Baseline data collection to establish “before” and “after” differences</p> <p>Surveys</p> <p>Interviews & focus groups</p> <p>Comparison groups</p> <p>Supervisor and employee questionnaires</p> <p>Observation</p> <p>Policy/procedure review</p> <p>Anecdotal data</p>	<p>Agency records & manuals</p> <p>Recipients of the training</p> <p>Trainee’s supervisors & colleagues</p>
<p>Results</p>	<p>Measure the effect the training had on the trainee's agency</p>	<p>Have organizational changes resulted from the employees’ training?</p> <p>Has the problem been solved?</p> <p>What is the cost-benefit or cost-effectiveness of the training?</p>	<p>Archival data</p> <p>Surveys</p> <p>Interviews</p> <p>Comparison groups</p>	<p>Agency records</p> <p>Client feedback</p> <p>Agency personnel feedback</p>

Level 1: Reaction

The first level, reaction, examines the training participants’ immediate feedback on the course, asking questions such as: Were course learning objectives met? How skilled and knowledgeable were the trainers? How relevant was the material to your job? Was the quality and presentation of materials sufficient? How effective were the instructional strategies? Answering this often involves administering a set of questions to the participants at the end of each training event.

Other alternatives at this stage include obtaining feedback from instructors or SMEs, who observed the training module, in terms of their perceptions of participants' reactions.

Some guidelines for evaluating participants' reactions to training from our experiences applying the evaluation model to the four training programs and Kirkpatrick (1998) include the following:

- Decide what you want to find out
- Design a form that will quantify participant feelings and reactions
- Encourage written comments and suggestions
- Obtain as close to 100 percent response as possible
- Develop standards to measure against
- Communicate results

Most criminal justice training providers evaluate trainee satisfaction with the training courses delivered but the value of the results is often debatable. Scarborough, et al. (1998) note that in policing, less attention is paid to trainees' attitudes toward the subject matter than to their satisfaction with other elements of the training. Kirkpatrick (1998) maintains that measuring reaction is important because funding agencies and training providers often make decisions about training based on participants' comments—because this is often the only evaluation data they receive. Kirkpatrick also notes that the results of trainee reactions can suggest motivation to learn, that is, if trainees do not like the program, they may not put much effort into learning. However, Kirkpatrick also points out that some learning may still take place despite low satisfaction ratings.

Another important opinion about training "customer satisfaction" ratings is that at this reaction level, "good evaluations can be bought" (McNaught 1991), for example, by providing impressive meals and facilities. In addition, reactions to a training can be influenced by the pacing of the course, selection of participants, length and number of breaks, personality of instructors, quality of course materials, motivation to attend, and course workload (Parry 2000).

Decide What You Want To Find Out. Training programs need to get the participants' reactions to a variety of areas—but shouldn't overwhelm them with too many items. Most training organizations divide what they need to know into several areas: subject matter, instructors, and accommodations. The subject matter obviously covers feedback related to the course itself—was it relevant to trainee's job; presented in an interesting way; good balance of lecture, exercises, discussion, etc.; good use audiovisual aids and handouts; and more. This part of the assessment might also ask for reaction to the schedules, timing, pacing, breaks, etc.

Instructors can be rated on a number of different areas: knowledge of subject matter, pacing, use of interactive methods, facilitative skills, helpful manner, communication skills, and more. It is very important that all instructors know the characteristics that will be used to rate

their performance. There is a more detailed discussion of evaluating instructors later in this chapter.

Finally, training providers often need to ask about the participants' satisfaction with facilities (comfort, convenience), service, meals, etc.

Design Form to Quantify Feelings and Reactions. The ideal training reaction form provides for a lot of information to be collected in the shortest time possible—the training is over and participants want to leave. Thus, most opt for simple variations of Likert scales that measure both positive and negative reactions. Some examples include:

- Excellent Very good Good Fair Poor
- Strongly agree Agree Disagree Strongly disagree
- Completely satisfied (5) (4) (3) (2) Completely dissatisfied (1)
- High (5) (4) (3) (2) Low (1)
- Very helpful Helpful Not helpful Useless
- Too much OK Not enough

Some social scientists don't like forcing all participants to make choices when they might feel uncomfortable for some reason, so the option is to add the following choices to the above:

- Unable to judge
- Not applicable
- Does not apply

Encourage Written Comments and Suggestions. The satisfaction ratings provide part of the participants' reactions. Also needed are participants' reasons for the reactions and suggestions to improve the program. Thus, opportunities for open-ended comments are critical to make sense of the reaction ratings. Participants should be especially encouraged to write comments and be given adequate time.

Obtain Response Rates. The bane of all training providers is trying to get a 100 percent response rate of participant evaluations and for them to be *candid*. There are many techniques to maximize return rates such as asking for evaluations before the training is over (e.g., "We will take our final break as soon as everyone completes the evaluation form."), collecting forms at the door as students leave, constantly stressing the importance of the forms, etc.

Develop Standards. Standards are developed through a series of trainings with similar rating forms. Simply put, you tally the responses to each question using the weighted scales and divide by the number of participants who answered the questions. For example, using the following five-point scale:

Excellent (5) Very good (4) Good (3) Fair (2) Poor (1)

With the following ratings:

How do you rate the meals? The answers were as follows from 20 students

10 _____ Excellent

6 _____ Very good

3 _____ Good

1 _____ Fair

0 _____ Poor

The calculations are: $10 \times 5 = 50$; $6 \times 4 = 24$; $3 \times 3 = 9$; $1 \times 2 = 2$; $0 \times 1 = 0$ for total of a total of 85, which divided by the 20 students comes to an average rating of 4.25. With a large enough sample, this now becomes our standard to measure satisfaction with future meals.

Communicate Results. The main stakeholders for this information are the instructors, training provider organization, and funding agency. These are the people who can make decisions to change instructors, modify modules and materials, and more.

Level 2: Knowledge or Learning

The second level of evaluation measures whether knowledge acquisition or learning took place by the participants. The obvious theory is that learning must occur before behavior can change. Measuring learning can involve establishing what knowledge was learned, what skills were developed or improved, and/or what attitudes were changed. Determining this often involves a pre and post-test survey, performance tests, or demonstrations of the skill learned (e.g., first aid techniques). The ability to identify the overall purpose and specific learning objectives of the training are critical for an assessment of learning (and for evaluation at all subsequent levels).

Some guidelines to evaluate learning include:

- Evaluate knowledge, skills, and/or attitudes both before and after the program
- Use a paper-and-pencil test to measure knowledge and attitudes
- Use a performance test to measure skills
- Use the results to take appropriate action

Measuring knowledge and skills is more straightforward than measuring attitude change. With knowledge, a pre/post-test on the curriculum content will provide a good measure of knowledge acquisition. To evaluate skill acquisition, the trainee is asked to perform the skills taught. Measuring skills can be either through the demonstration of the actual skills taught or by simulation (e.g. paper and pencil testing, role play). Attitude change, however, can be quite challenging to measure, although a variety of techniques can be used, including organizational climate surveys, employee attitude surveys, projective or sentence completion tests, or simulation games (Parry 2000).

In criminal justice training, the most common Level 2 evaluation involves measuring the new knowledge (cognitive learning) gained by trainees using before and after questions asking about policies, procedures, facts, and other information. The tests are mostly true/false and multiple choice for ease of scoring. The pre-test provides the basis for what the trainees know and don't know about the training subject. The post-test provides results on the extent to which trainees learned from the training. A second post-test is sometimes administered four to six months after the training to measure the extent to which the learning has been retained.

The challenge in constructing before/after knowledge tests is to measure learning at the highest levels of Bloom's taxonomy (Bloom, et al. 1956). It is relatively easy to "pitch" the cognitive functioning of test question items at Bloom's levels 1-4 [(1) knowledge or memorization, (2) comprehension or understanding, (3) application of what has been learned in concrete situations, (4) analysis of material for relationships]. It is more difficult to develop questions that measure Bloom's more advanced cognitive levels 5 and 6 [(5) synthesis or integration of material (6) evaluation and being able to make substantiated judgments about the material].

In developing the tests, educators typically collapse the categories for ease of question construction (see Minor et al. 2005; Usova 1997) as follows:

- Knowledge/Comprehension: Information that trainees need to remember and understand
- Application/Analysis: Material that trainees should be able to break down and make relevant to actual situations and concrete relationships
- Synthesis/Evaluation: Diverging pieces of information that trainees need to be able to integrate for purposes of making judgments supported by facts

Two popular methods of estimating the reliability of knowledge tests include parallel forms reliability and internal reliability. Parallel forms reliability can be estimated when multiple versions of a test are administered to the same trainees one time per version and scores from the various forms are correlated. Internal reliability can be calculated anytime a single version of a test has been given one time. Various methods exist for estimating internal reliability, but one of the best and most common is Cronbach's alpha,⁶ which is based on correlations among items and conveys the extent to which high (or low) scores on a given item are associated with high (or low) scores on other items as well.

However, not all training can be evaluated at Level 2 with before/after knowledge tests. In the planning stages, the training providers and funding agencies need to decide if the course will receive a Level 2 evaluation and then gear the test to the learning objectives. The learning objectives must be clear that the "trainees will learn to ..." (specific actions). In the four training programs evaluated as part of ILJ's project, several used before/after knowledge tests; none used performance tests (see Chapter 6 for more discussion on this).

Level 3: Behavior Change

What happens after the training when trainees return to their jobs? How much transfer of learning knowledge and skills or attitude change occurs back in the police department, corrections facility, or court? A Level 3 evaluation examines how the work behavior of participants may have changed because of the training.

Key evaluation questions for a Level 3 evaluation are: What is the extent to which participants have transferred knowledge and skills to their jobs? What behavioral changes resulted from the training? Answering these questions requires collecting baseline data to establish participants' knowledge, skills, or attitudes before the training, and then documenting

⁶ See Allen and Yen (2002).

changes that occur after the training back on the job. Available data collection tools include surveys, interviews, focus and comparison groups, supervisor and employee questionnaires, observation, policy/procedure review, and anecdotal data.

Based on our experiences in this project and Kirkpatrick's studies, some guidelines used to evaluate behavior change include:

- Use a control group, if practical
- Allow time for behavior change to take place
- Evaluate both before and after the program
- Survey and/or interview the following: trainees, their immediate supervisor (or their subordinates), and others who often observe their work behavior
- Obtain a significant sample size

Level 3 in our revised evaluation model assumes that individual behavior change (e.g., in job performance or attitude) will occur that can reasonably be attributed to the criminal justice training. As the International Society for Performance Improvement (ISPI) (2000) points out, "just because a student can demonstrate increased knowledge or skill is no guarantee that behavior on the job will actually change" (p. 4). Evaluation at the third level can be relatively simple for skill-based training (e.g., where the trainee learns CPR), but for many training objectives, evaluation of behavioral outcomes is not a straightforward process. In the training design phase, specific learning objectives must describe what the trainee will be able to do after completion of training. These objectives must be clear, honest, and realistic.

In addition, a task for evaluators (at both the third and fourth levels) is to validate whether training was in fact the correct performance intervention to begin with (ISPI 2000; Bumgarner 2001). For example, one pitfall is to expect training to accomplish too much with respect to changing attitudes (Buerger 1998). A classic example is cultural diversity training, which is difficult to design and is often arranged as a quick fix for an entrenched problem that might be more effectively addressed through other interventions.

Assuming a job performance issue is one that can appropriately be addressed by the type of training provided, and that the trainee did learn, any number of factors may still inhibit the transfer of learning to performance on the job. These include motives and expectations; lack of incentives or consequences; lack of environmental support, resources, tools, data, or information;

and individual capacity (ISPI 2000; Bumgarner 2001). Of course, the possibility that changes in job behavior result from influences other than the training must also be addressed by developing the most rigorous methodology feasible.

Control group. In this method, the evaluator finds a "control" group to compare to the trainees (experimental group), for example, other police officers on the same patrol squad who did not receive the training (e.g., training in problem solving). The evaluator then compares problem solving activities of the trainees with the control group to try and find evidence that the trainees are performing better because of the training. However, it is difficult to control for all the extraneous variables that could occur to affect the results—the groups may not be equal in all factors, shift work may make a difference (different opportunities for problem solving), and many other factors. As described in Chapter 2, implementing these types of evaluation designs is challenging and costly.

Interviews/surveys. It is relatively cheaper and easier to conduct follow-up interviews or surveys with the trainees six months or so after the training to ask about the extent to which they are applying the training. However, for obvious reasons, the trainees may not admit that their behavior hasn't changed. You may still be able to obtain some useful information from trainees if the questions are objective—asking for factual demonstrations—and not subjective opinions.

For more objective feedback, the evaluator needs to contact the trainee's supervisor or peers who are knowledgeable about the trainee's work behavior. Kirkpatrick (1998) cautions to look for persons who are best qualified, reliable, and available. The last point is very important. If written surveys are used, then availability is not an issue. However, if the evaluation must be conducted in a lengthy phone interview, then availability becomes critical. In both methods, mail survey or phone interview, incentives should be developed to obtain the highest response rates possible. This links back to the need to obtain organizational commitment from the trainee's agency in the application process.

Level 4: Organizational Impact

A Level 4 evaluation examines the impact of the training on the organization. A Level 4 evaluation is results-oriented, asking questions such as:

- How much did work quality improve at the organization level?
- How much did work productivity increase at the organization level?
- What tangible benefits have been received?
- What is the return on investment?

The key evaluation question for a Level 4 evaluation is what effect has the training had on the organization? In particular, a Level 4 evaluation assesses if the problem has been solved and/or the need has been met at the organization level. This is the most complex, expensive, and time-consuming evaluation design because it is challenging to isolate the results of training. Data collection techniques include archival records, surveys, and interviews.

Our expanded model departs from Kirkpatrick at this level. Kirkpatrick called this level "results." We call it "organizational impact." Kirkpatrick's model was applied in the private sector where companies often sent entire work units (e.g., sales force) to training and expected "bottom line" results such as increased sales, increased profits, improved quality, increased production, and more. Criminal justice organizations cannot measure profit and loss or sales. Thus, we must gear Level 4 to the type of organizational impact that resonates with the criminal justice field. At this level, we are measuring the impact at the larger than individual level—the impact in the field. This can involve a variety of outcomes such as improved citizen satisfaction, implementation of new department-wide policy changes, and more. As with Level 3, the changes are easier to measure when the training involves a specific skill or knowledge. It is more difficult to measure at Levels 3 and 4 training topics such as leadership, communication, motivation, and managing change.

A separate aspect of Level 4 evaluations considers cost benefits/effectiveness and the training's return on investment. Cost benefit analysis measures both cost and training outcomes in monetary terms. Cost benefit analysis allows for the discovery of whether training expenditures are less than, similar to, or greater than training benefits (Yates 1999). Cost effectiveness analysis differs from cost benefit in that it does not rely upon monetary measures for determining outcome (Welsh, Farrington, & Sherman 2001). Cost effectiveness analysis only measures the monetary value of training costs, while measuring effectiveness in terms of whether training objectives were met (Yates 1999). Typically, the key question for criminal justice agencies is: If we invest X amount of money in a training program, what benefits should we expect from that investment? Criminal justice agencies may be interested in the training

costs at their internal organizational level—how much do we invest in recruit training and what do we get out of it. However, in terms of sending personnel to the types of training we evaluated in this project, the costs to the criminal justice agency are minimal because of the federal funds used to develop and deliver the training.

On the other hand, the federal agencies providing the training funds should want to know cost benefit comparisons of the training they pay for. How much does effective training cost? Cost considerations for the development and management of training are discussed more in Chapter 6. The study team did not conduct cost analyses of the four training projects evaluated because we felt that we would be pushing the bounds of cooperation since each of the four organizations were voluntarily allowing us to evaluate their training.

Some guidelines to evaluate organization impact include:

- Use a control group, if practical
- Allow time for results to be achieved
- Measure both before and after the program

Level 4 evaluations are particularly difficult for criminal justice training. First, the goal of the training may have been to increase trainees' knowledge, but this will not necessarily change the fundamental way the trainees' agencies conduct business. For example, training for investigators may have the objective of teaching about human trafficking. This new information about human trafficking, however, will not necessarily translate into a measurable change that will affect the investigators' respective organizations—investigating more human trafficking cases. This does not mean that the training was insignificant; it just means that the training is not measurable at a Level 4.

Second, training may have objectives that yield measurable results but still not on an organizational level because many criminal justice agencies send one or two personnel to a training, not entire units or sections or departments. Many examples for Kirkpatrick's model involve training conducted in the private business sector, such as when a company sends its entire sales staff to marketing training. It would be expected that organizational changes occur as a result—an increase in measurable sales. With criminal justice training, on the other hand, most departments cannot afford to train their entire staff all at once. Thus, Level 4 evaluations

need to accommodate the unique needs of criminal justice agencies when planning for the measurement of results.

Control groups. The benefits and issues in using control groups are the same under this level as discussed in the previous Level 3. The evaluation design is time-consuming and costly to set up and implement and it is difficult to control for all the various variables that may be as much, if not more, influential of organizational outcomes than the training.

Comparing the Four Levels of Evaluation

The four evaluation levels are successively more difficult to implement. A Level 1 evaluation is the most common and often involves simple participant reaction surveys. A Level 2 evaluation adds a pre/post-test component to measure knowledge, skill, or attitude change. Evaluating training programs at Levels 3 and 4 is more challenging, requiring comparison groups or before and after interviews or surveys with the participants, their supervisors, and colleagues. A Level 4 evaluation can be particularly challenging because it can be difficult to link organizational change measures directly to the training. Parry (1997) estimates that 85 percent of all business training programs are evaluated at Level 1, but fewer than 10 percent are measured at a Level 4. The number of criminal justice training programs conducting a Level 4 evaluation are likely even lower.

Indeed, there are situations in which it is simply not possible to conduct a Level 3 or 4 evaluation (Parry 1997). For instance, some behaviors cannot be measured objectively; the behaviors may never occur (e.g., a manager is required by state POST standards to receive CPR training but never has occasion to use it on the job); they may cost too much to measure; or it may not be possible to attribute the behavior change to training.

There are distinct advantages and disadvantages in each of the four levels. The chief advantage of a Level 1 evaluation is that the instruments are easy to create, administer, and interpret because the results are easily tallied, scored, and summarized. As discussed earlier, most criminal justice training programs build a Level 1 evaluation into their course. The main disadvantage to a Level 1 evaluation is that it could be “bought” with nice classroom facilities, extra breaks, entertaining instructors, and quality snacks and treats. These "customer satisfaction" ratings do tell what the participants thought of the training, instructors, facilities, etc, but are not informative about what the participants learned, or whether the training

objectives were met. A Level 1 evaluation provides some important feedback about the training environment and the instructors, but it should not be considered a stand alone evaluation. For useful evaluation of training, a Level 2 evaluation is minimally necessary, with a Level 3 (and ultimately a Level 4) the goal.

The chief advantage of a Level 2 evaluation is the relative ease in developing the instruments. A Level 2 evaluation instrument measures what the participants learned. Thus, an appropriate instrument will be a before/after test of the training's learning objectives. For instance, if the purpose is to teach a new skill, then a skills test would be appropriate. If the purpose is to convey knowledge or change attitudes, then the before/after test should reflect this new learning or change in attitude. It is important to note, however, that while it can be a straightforward process to develop an instrument to measure learning, it can be challenging to develop one that truly measures what the planners wanted participants to learn from the course. For example, when writing a multiple choice knowledge test, it is often easier to write a fact-based question than an abstract one. So it is much more difficult to measure synthesis or judgment than memory and recall.

A Level 2 evaluation is useful when it is important to measure what the participants learned at the training and it is not be feasible to complete a Level 3 evaluation, which measures how participants' behavior changed as a result of the training. For example, the skill or knowledge taught was not expected to produce a specific, concrete, measurable change in behavior; there are not enough resources (financial and time) to conduct an adequate Level 3 evaluation; potential evaluation participants and/or their agencies are not cooperative with the necessary follow up required for a Level 3 evaluation; or changes in behavior cannot be attributed solely to the training.

A Level 3 evaluation is more complex than a Level 2 but can be very informative in terms of providing feedback to key stakeholders about how the training participants actually used what they learned at the training. In essence, a Level 3 evaluation uses performance measures as an indicator of training effectiveness. This is most easily accomplished when there are tangible objectives, such as increasing the number of police problem solving activities, or increasing the number of intelligence analyses performed per investigator. A Level 3 evaluation relies upon quantifiable, observable outputs, and employees who exert some control over production of these

outputs and how they spend their time. Without these factors, a Level 3 evaluation will be much more difficult to conduct.

A Level 4 evaluation is the most difficult, yet extremely informative type of evaluation for training programs. It provides good measures of training effectiveness, yet it is very challenging to design and execute. For a training to be measurable at Level 4, it needs to meet the requirements of a Level 3 evaluation, plus examine how the trainees' organizations did or did not benefit from participation in the training. Being able to determine if the "impact" being measured can be attributed directly to the training is very difficult. Not only must the training have provided tangible and measurable skills, knowledge or change in attitude among participants, but the evaluation design must allow sufficient time to pass for these results to be achieved; and it needs to eliminate other influencing factors that may achieve the same results that the training sought to accomplish.

While a Level 4 evaluation is the "gold-standard" for evaluating training programs, each of the other three levels answers specific, important questions about the training. With all the money that the U. S. Department of Justice (USDOJ) puts into criminal justice training, training providers should always produce a Level 1 evaluation, which is common, and also produce a Level 2 evaluation when the training is clearly skill or knowledge-based. In addition, training providers should be required to at least attempt a Level 3 evaluation with a significant sample of the trainees. A Level 4 evaluation should be optional but encouraged.

Evaluating the Training Instructors

Earlier in this chapter, we discussed criteria to consider in selecting instructors for criminal justice training and noted potential advantages and disadvantages of using consultants who specialize in training, training provider staff, or practitioners working in the field. This section addresses the need to incorporate an assessment of the instructor's performance into any training evaluation. An instructor's delivery of the curriculum material, group management and facilitation skills, and certain individual characteristics (credentials, personality, mannerisms, etc.) affect participants' learning, but these factors can be difficult to measure objectively.

There are two main ways to evaluate instructors: obtain feedback from students and conduct classroom observations. As part of this project, our team developed and briefly tested an instrument to evaluate instructors and also observed how the four training projects handled it.

Three of the four relied mainly on student feedback. The NW3C did prepare a standard form to have SMEs and other experts evaluate the instructors (see Appendix A for both forms). In addition, during the pilot test of the FIAT course, NW3C brought in a number of prospective instructors and had them audition a module, videotaped the presentations, evaluated the instructors using the standard form, and critiqued the instructors (in private).

Some of the standard criteria used to evaluate instructor performance are shown in the list in Exhibit 4-4 below.

Exhibit 4-4: Examples of Selected Criteria to Evaluate Training Instructor Performance

- Introduced learning objectives at beginning of module
- Responsive to trainees—asks about their needs
- Comfortable in front of class
- Provides constructive feedback to trainees
- Available to trainees outside of class
- Makes training materials relevant to trainees
- Blends theory and practice
- Uses effective questioning and feedback skills
- Encourages multiple trainees to participate
- Uses clear and concise speech
- Acknowledges multiple points of view on controversial subjects
- Effectively uses media to make points
- Demonstrates an awareness of and manages time well
- Demonstrates comprehensive knowledge of subject
- Explains key principles and skills
- Demonstrates experience-based knowledge of subject
- Relates module subject to other course topics

Revising the Training

The final step in the model is revision of the training. The evaluation will have informed the evaluators and key stakeholders about what the training participants thought of the training, what they learned, how their performance and/or knowledge changed, and possibly how the organization benefited from the training. This information can then be used to modify the training as needed to make it more efficient, closer to meeting targeted goals, or better suited to its audience. Revision of the training can involve something as simple as providing more breaks or a bigger classroom, to providing regionally-based training locations, to completely changing the instructors or overhauling the entire curriculum. It is important that once the training is revised, the evaluation continues so that the training can be constantly updated and perfected.

Chapter 5

Project Methodology

In 2003, the National Institute of Justice (NIJ) funded ILJ's competitively-selected proposal to (1) develop an evaluation model to evaluate criminal justice training programs; and (2) apply the model to evaluate selected criminal justice training programs. In FY 2002, Congress had earmarked \$94.5 million in Bureau of Justice Assistance (BJA) discretionary funds for 88 specific programs, about 25 percent of which were training programs. Our directive was to apply the training model to four of those programs.

This chapter explains how the four programs were selected as test sites for the model and provides additional information about the methods used by the study team.

Key Decision Processes for Site Selection

At the beginning of the grant period, NIJ provided the study team with a list of 15 FY 2002 earmarked criminal justice training program grants. From this list, four programs were to be selected for evaluation. The 15 programs were:

- Center for Task Force Training
- Law Enforcement Innovations Center
- Littleton Area Learning Center
- National Association for Court Management
- National Center for Rural Law Enforcement
- National Crime Prevention Council
- National Judicial College
- National Corrections and Law Enforcement Training and Technology Center
- National Training & Information Center
- National White Collar Crime Center
- Oregon Federal Law Enforcement Training Center
- Regional Training Academy, Springfield, MO
- Roger Williams University
- Simon Wiesenthal Center
- University of Mississippi School of Law

The decision to focus the training evaluation model on only a subset of these 15 training projects was a limiting factor in demonstrating the robustness of the evaluation model. For example, we proposed testing the model on a police recruit training class but NIJ and BJA felt

that by legislative policy the training evaluation funding was tied to the above Congressionally-earmarked projects.

To begin the process of paring down this list, ILJ convened an expert working group of six academic and practitioner professionals from the fields of criminal justice and education.

The working group developed the following criteria for selecting the projects to be evaluated:

- Diversity of subject matter and audience: law enforcement, courts, corrections, etc.
- Diversity of type of audience: line level, management/leadership, administrative, technical/information technology, etc.
- Diversity of training objectives: skill-based, knowledge-based, behavior-oriented, etc.
- Diversity of training modalities: classroom, self-paced on PC/CD, distance learning, etc.
- Level of cooperation and commitment from grantee: willing to engage in Level 3 or Level 4 evaluation; also, commitment to participate from trainees' agencies
- Degree of front-end planning: course is based on needs assessment, selection criteria for trainees
- Focus of analysis: individual vs. organization, based on number of trainees per unit at an agency. That is, if one out of 100 patrol officers is being trained, this would not be expected to show much impact back in the agency unless training is in a specialty and the agency is small.
- Adequate training dosage
- Involvement and commitment of trainees' agencies in trainees' acquisition of the training—realistic expectations
- Likelihood that trainees have time and opportunity after training to practice and apply new learning
- Attrition factors: trainees dropping out of training before finishing
- Cost, if any, to trainee's agency
- Local evaluation ongoing or planned

Using these criteria, we were able to eliminate five of the training programs that did not meet enough of the criteria (e.g., training was one day, training not yet developed, and more): the National Association for Court Management, National Center for Rural Law Enforcement, National Crime Prevention Council, Oregon Federal Law Enforcement Training Center, and the Regional Training Academy in Springfield, Missouri.

To further refine the list, evaluability assessments guided by the selection criteria and consistent protocol were conducted on the remaining 10 training programs by study team and working group experts. These evaluability assessments used guidelines developed by NIJ under the Analytical Support Program (see presentation by Banks, D. (2005) at NIJ National Research and Evaluation Conference; see also, Justice Research and Statistics Association 2003).⁷ On each of the ten grant-funded training programs, a senior research staff person (working group members were all experienced social science Ph.D.s) reviewed the grant application and other materials (e.g., progress reports) provided by BJA staff, conducted extensive telephone interviews with the program director and staff, and talked to the BJA grant monitors about the training programs.

All ten evaluability reports were analyzed and compared in a detailed spreadsheet. ILJ prepared a synthesis report, which was reviewed by the expert working group members and NIJ staff (see Appendix C for synthesis report). While all criteria were important in selecting training programs to be evaluated, particular weight was given to diversity of subject matter and audience, level of cooperation, and level of training dosage. The result was selection of the following four training programs for evaluation:

- National White Collar Crime Center's *Foundations of Intelligence Analysis Training* (FIAT)
- Simon Wiesenthal Center's *National Institutes Against Hate Crimes and Terrorism* training
- National Judicial College's *Civil Mediation* training
- National Corrections and Law Enforcement Training and Technology Center's *Advanced Leadership Techniques for First Responders, Corrections, and Security Officers*

Each of these training programs had considerable strengths and some weaknesses. The final four chosen reflected diversity in training target audiences (law enforcement, corrections, judges, mixed criminal justice teams) and the first three listed above were focused on teaching a measurable skill, or knowledge, or attitude change. The National Corrections and Law Enforcement Training and Technology Center (NCLETTC) course was geared toward

⁷ See Appendix B for the evaluability guide.

leadership, a more challenging training topic to evaluate. Finally, NCLETTTC was the only training program with an online component, and this played a role in its selection.

Overview of Methods

Over a period of several months, study team members communicated extensively with the selected organizations, visited their headquarters, discussed the evaluation efforts in detail, and engaged in a variety of negotiations to obtain their cooperation.

The next task was to create individual evaluation plans for each of the four programs. These evaluation plans presented the methodology and data collection plans for each of the four training programs (see Appendix D for an example). After each of the training organizations reviewed the plans, some modifications were made. We then negotiated and executed a memorandum of agreement with each organization (see Appendix E for an example). It is clear that this project could not have been possible without the extensive cooperation of the CEOs and staff of the four training organizations.

Over the next year or more, study team members implemented the evaluation plans with the four training programs. With varying degrees of success, described more fully in the next chapter, we applied the four levels of evaluation to assess the training programs. The principal methods for collecting data included the following at each training evaluation level:

- **Level 1** (participant reaction) data were written surveys by trainees and evaluator observation. At each site, the evaluation team observed at least one entire training class and used the existing participant reaction surveys that each training organization had developed for the training. At one program, evaluators also conducted reaction interviews with a sample of training participants.
- **Level 2** (knowledge gained) data were obtained by pre/post training knowledge tests. These tests were created either by the training developers or the evaluation team. With the exception of the Simon Wiesenthal Center training, which used pre/post interviews instead of written tests, knowledge tests were used throughout.
- **Level 3** (behavior change) data were obtained by surveys or phone interviews. When possible, surveys and interviews were also conducted with the training participants' supervisors.
- **Level 4** (organization impact) data, where possible, were obtained through interviews, case studies (for Simon Wiesenthal Center), and surveys (National Corrections and Law Enforcement Training and Technology Center).

In sum, we were able to show the strengths and weaknesses of attempting to collect all levels of evaluation data using the modified Kirkpatrick's model.

Chapter 6

Cross-site Comparisons and Findings

The training evaluation model developed and tested in this project is designed to aid evaluators and key stakeholders of criminal justice training programs in conducting evaluations of such programs. The model aims to be flexible enough that an evaluation can be designed during the initial planning phase of training or implemented after a training has already been developed and even carried out. An effective evaluation provides useful feedback to improve the training. For an evaluation to be useful, the evaluation plan must be carefully thought out and conducted in a manner that increases its practical value and decreases uninformative data gathering techniques.

This chapter includes a discussion of our experiences in implementing the training evaluation model steps and provides cross-site comparisons of the four training evaluations.

This project tested the training evaluation model described in Chapter 4 by evaluating four trainings that had received large earmarks from BJA:

- **National White Collar Crime Center (NW3C) *Foundations of Intelligence Analysis Training (FIAT)***. This training was developed by the NW3C in conjunction with the International Association of Law Enforcement Intelligence Analysts (IALEIA), Law Enforcement Intelligence Unit (LEIU), and the Regional Information Sharing Systems (RISS) to provide a standardized, basic analytical intelligence training curriculum for entry-level law enforcement with the goal of being the standard-bearer course in intelligence analysis training.
- **Simon Wiesenthal Center (SWC) *National Institutes Against Hate Crimes and Terrorism*** training was designed to help participants develop new perspectives on hate crime and terrorist acts, aid them in forming multi-agency collaborations, and foster the development of strategic action plans for combating hate crime and terrorism.
- **National Corrections and Law Enforcement Training and Technology Center (NCLETTTC) *Advanced Leadership Techniques for First Responders, Corrections, and Security Officers*** training. This training was developed for criminal justice personnel in leadership positions who are responsible for the professional development, policy implementation, tactical mission execution, and formal evaluation of staff and personnel within the first responder, corrections, and security officer arena. The goals of the course were to enhance and build on a supervisor's leadership abilities and provide practical application of values-based leadership within a variety of organizations.

- **National Judicial College (NJC) *Civil Mediation*** training was developed to familiarize participants (mostly judges) with the civil mediation process and qualify them for certification in states that require it. At the completion of the course, the participants should be able to start and conduct a mediation session, outline common standards of conduct and ethical considerations, handle special problems and avoid classic errors, and summarize the interpersonal dynamics of mediation.

The four training programs that participated in testing the training evaluation model were selected because (a) they each had the goal of teaching a specific skill or type of knowledge, and (b) they represented different aspects of the criminal justice system. This diversity allowed us to determine how the model held up across a range of criminal justice professions.

The executive directors of the organizations providing the training, the developers, instructors, and other personnel were willing and eager to participate in the evaluations of their training. The organizations had already planned some degree of evaluation. They were proud of their training products and happy to receive feedback to make them even better.

Summary of the Training Evaluation Model's Applications

The purpose of this project was to produce a training evaluation model and test it through four diverse evaluations of criminal justice training programs. The result is a flexible model, based on Kirkpatrick's model (Kirkpatrick 1998), that can serve those in the curriculum planning and development stages as well as those who have already created a course and now are looking for evaluation help. The model, as illustrated in Exhibit 4-2 in Chapter 4, covers the following phases of training development:

- Conduct needs assessment
- Design training plan
- Develop and test curriculum
- Deliver curriculum
- Evaluate the training and trainers
- Revise the training

Needs Assessment

A needs assessment analysis is an important first step that provides feedback on what training is needed, how the training should be delivered, or whether an existing training is

meeting the needs of an agency or organization. As described in Chapter 4, needs assessment analysis involves a multi-step process to identify the users of the assessment, identify the target population for the training, determine why the training is needed, conduct the assessment and analysis, and communicate the results.

In this stage, the training should be completely justified before funding is invested in the curriculum development. Otherwise, training that is duplicative of other training or doesn't meet an established need might be developed and waste precious resources.

The four projects varied in applying the needs assessment step before developing their training plans. NW3C did a good job of conducting a needs assessment as a *sine qua non* to developing the training. If they could not demonstrate that the FIAT course was filling a void or gap in current training of entry-level law enforcement intelligence analysts, they were not going to commit to the course development. They brought together key stakeholders and experts (the two main law enforcement intelligence analysts associations,⁸ representatives of the Regional Information Sharing Systems—funded by USDOJ since the early 1980s to sharing intelligence among state and local law enforcement, and other experts in intelligence training). They reviewed existing intelligence training (e.g., Anacapa Sciences—leader in intelligence training for over 30 years, Drug Enforcement Administration, Federal Bureau of Investigation, Federal Law Enforcement Training Center, and others). Thus, NW3C determined, through extensive research, that their proposed training course was unique and not duplicative of existing training. They were also encouraged by a report from the Global Intelligence Working Group (GIWG, 2003) that documented the void in intelligence training for state and local law enforcement. They found that the target audience for the training, newly hired intelligence analysts in state and local law enforcement, did not have a standardized, entry-level, basic course that could provide the fundamentals of intelligence analyses.

The other three organizations that developed courses did not put as much effort and resources into formally documenting the initial needs assessment phase, although they each conducted some degree of research and analysis. The SWC, in developing the hate crimes course, determined that, while others were presenting hate crimes training to some degree, their

⁸ International Association of Law Enforcement Intelligence Analysts and Law Enforcement Intelligence Unit

approach of bringing in multi-disciplinary teams was unique. NCLETTTC determined that there was no other online course for police and corrections leadership. NCJ identified a gap in training judges regarding civil mediation skills.

Thus, a rigorous needs assessment stage should be a requirement for any OJP-funded training. Too often, organizations have received earmarked funds to develop some type of criminal justice training that was duplicative of training already being offered somewhere else. This type of "reinventing the wheel" has become all too common in criminal justice training.

Training Plan

Once the need for training has been justified, the second phase of training development involves creating a training plan. The purpose of the training plan is to think through all the management and administration components of the training. The training plan items and questions that were presented in Chapter 4 should be answered in this stage.

Often, little attention is given to conducting the training needs assessment or developing a training plan. Instead, designing and delivering the curriculum are the objectives or “starting place” for many organizations in creating a new course. We believe that this does not provide a strong foundation or backbone for the course.

In this stage, the training organization needs to clearly identify, articulate, and document the goals—what do we want to accomplish through training. How will the trainees think or function differently after the training? This part of the plan uses the results of the needs assessment stage.

Management and administration include developing the foundation or building blocks to develop the training, acquire the trainers, acquire the technology needed, advertise and invite participants, manage the logistics (training facility, food/beverages, housing, etc.), maintain training records, and more. The best training course can founder without a strong and experienced organizational foundation to support the administration, management, and logistics.

In our study, all four of the organizations had extensive experience in managing and delivering training. Each had demonstrated an experienced approach and foundation to support previous training. However, the projects varied in the degree to which they actually documented

a "training plan" in writing. For the most part, the organizations just folded the new training into their existing training management practices. The requirement for a documented training plan should be clearly spelled out by the funding agencies.

The trainings evaluated were carefully planned. The training planners and developers designed the training to fulfill a need, laid out clear training objectives and goals, and worked to ensure that the training fulfilled the need. They also operationalized specific goals that the training participants were to take away from the training. One of the keys to evaluating training at the four levels is having measurable objectives that can be isolated. For instance, the NW3C was teaching participants the skills needed to be an intelligence analyst; the SWC was helping participants develop both strategies and networks for fighting crimes of hate within their communities; and the NJC was teaching judges and other judicial personnel how to conduct mediations of civil cases.

The more specific the skills or learning are in the development of the training, the easier they are to evaluate. The more difficult training evaluations tend to be about training involved in learning to be better *leaders* or *managers*—broad areas of learning.

The agencies also varied to the extent and degree that they developed detailed cost plans for the training courses.

Criteria for Selecting Trainees and Obtaining Organizational Commitment

As we discussed earlier, training organizations need to apply some degree of selection criteria to prospective trainees so that federal funding agencies have some certainty that the resources put into the training will result in trainees using and applying what they have learned back home in their agencies.

The NW3C training had fairly rigid trainee selection criteria—entry level intelligence analyst from a member law enforcement organization. In the other three examples, the criteria were broader. In our research, we found that most of the OJP-funded training did not apply rigid selection criteria to potential training participants. In fact, many organizations that we talked to opened the training broadly to "criminal justice personnel."

Of our four training examples, only the NJC charged tuition for training. The other three, as with most OJP-funded training, did not charge trainees for the training.

The SWC training required teams of criminal justice professionals from each attending jurisdiction. This requirement resulted in time-consuming recruitment efforts from the SWC staff to try and secure all needed members of a team. For example, each team was expected to have a judge attend the training, but many teams found it hard to find judges who were both available to attend and interested in the subject matter. NW3C advertised its new FIAT course to member agencies only. The other trainings evaluated, NJC and NCLETTTC, advertised their respective Civil Mediation and Advanced Leadership trainings as they did all of their trainings, through their course catalogs and websites. Participants signed up without either of these organizations relying upon heavy recruitment efforts or incentives.

However, none of the four agencies studied required the trainee in the application process to obtain any degree of commitment from the trainee's organization that the trainee would be allowed and encouraged to apply the training back in the organization. NW3C attempted to put a statement related to this requirement on the training application website but then decided against it. Again, we feel that the trainees' organizations must affirmatively display a degree of vested interest in the training so that the potential for the trainee to apply the learning later is optimized.

Evaluation Plan

In this second stage of the training development and evaluation model, the training organization should develop an evaluation plan. The application of that plan is discussed in stage V (see Exhibit 4-2 in Chapter 4). The reason that the organization should develop the evaluation plan in this second stage is to clearly emphasize the need to develop and implement the evaluation concurrently with the development of the curriculum and delivery of the training. Too often, training evaluation, with the exception of Level 1 (participant reaction) is an afterthought—conducted several years after the training has been underway.

More on training evaluation and how the four training organizations applied it is discussed later in this chapter.

Develop and Test Curriculum

This stage involves refining learning objectives, developing course content and materials, selecting instructional methods and media, pilot testing the course, and revising it based on feedback. The four training courses studied varied in how they handled each of these areas.

Develop Learning Objectives and Course Content

In each of the four courses evaluated, the training organizations did a good job of articulating the learning objectives and developing course content and materials. They each established specific learning objectives as discussed in Chapter 4. They each also used a variety of SMEs to some degree. For example, the NW3C used at least six to eight SMEs who worked nearly 12 months to develop course content and materials. During the year, they met two or three times as a group to review the work to date and provide feedback to each other. On the other hand, NJC turned to a pair of experienced contractors who very quickly customized a course for judges that they had already developed and delivered in the private sector for mediators in civil cases.

The key point is that for training course development, SMEs vary (in-house staff, practitioner, outside contractor, etc.) but are absolutely essential.

Each of the trainings studied were lengthy enough to provide an adequate dosage level to test the training evaluation model. Two of the trainings were 40 hours, one was approximately 30 hours, and one was 16 hours. The length of these trainings represented a well thought out attempt to balance the need to meet the training goals with the needs of participants. For instance, the shortest training evaluated was the NCLETTTC course, which specifically targeted policymakers such as police chiefs, lieutenants, and other supervisors. Training longer than two days would have been difficult for these managers to attend. On the other hand, the 40-hour NW3C FIAT course trained entry-level analysts who could take more time off from their jobs.

We did not conduct a formal content rating and comparison of the actual training texts or notebooks in each course, although they were reviewed by training experts on the evaluation team according to the following criteria:

- Organization—how the material was arranged
- Readability—how understandable were the concepts (clear and concise writing)
- Content—current and informative subject matter
- Instructional value—usefulness and relevancy of the material in transferring knowledge

Each of the four courses studied did a good job on the training notebooks. In each case, the students' feedback on the materials was generally positive.

One key to useful course materials, especially in skill-building training, is that the student can use the course materials back home on the job to look things up; the course materials should contain useful examples and techniques, not merely principles.

Another measure of "success" to a certain degree is feedback that the trainees distributed copies of the training materials to other employees back at work. This sharing of materials helps to disseminate the usefulness of the training, although this would be difficult to measure—it seems to be an "intuitive" theory. This does raise the issue of using copyrighted materials in federal government training. Training organizations that use federal funds to deliver training should be wary of using copyrighted materials that cannot be reproduced by training participants when they get back home. If the training is free to criminal justice practitioners, then they should not have to pay a fee to a contractor to reproduce the training materials and distribute them to their peers back at work. This inability to reproduce training materials caused some dissatisfaction with one of the training programs studied.

Instructional Methods and Media

The instructional methods and media were as varied as the trainings themselves. The SWC training relied heavily on guest speakers, interactive media, and tours of the Tolerance Center and Holocaust Museum. The NCLETTTC classroom course relied mostly on lecture and exercises; they also had an online course. The NW3C reinforced lectures with extensive classroom exercises. The NJC training relied heavily on lectures followed by simulated role play.

All four courses attempted to use adult learning principles to some degree. NJC's role play, which probably involved nearly half of the 40-hour course, showed the most adult learning concepts. The role play exercises used real case examples and involved all class members. In the FIAT training program, NW3C's training director delivered adult learning and facilitation training to most of the instructors. The other two training programs included adult learning methods and facilitative instruction to a lesser degree.

Pilot Test

Of the four training programs evaluated, the NW3C conducted the most elaborate pilot test of the FIAT course. NW3C's FIAT pilot test, which was done before the SMEs (members of key intelligence analyst associations) and key members of the NW3C staff, lasted over three

days, and included videotaping of presentations and extensive critique of instructors and course materials using standard checklists. Based on this pilot test, the course was substantively revised. One of the key areas learned in the pilot test was the length of a module—how much material for a given module could be delivered in what amount of time.

On the other hand, in developing the Civil Mediation course, the NJC relied on experienced trainers who had already developed and delivered a similar course to just modify the course for the new audience of judges. In this case, the "pilot test" was the first delivery of the actual course.

What we found overall in the review of pilot tests was that the expected time for exercises is often underestimated. They always take longer in reality than the developers think. Additionally, the time spent on debriefing a completed exercise in front of the whole class to enhance the learning experience for all trainees is often underestimated. Unfortunately, when time is running short, this is the area that many trainers opt to cut from the module—when in reality, this may be the most important part of the module.

Trainer Selection

The four training programs differed in the process of selecting the type of instructors used in the various trainings evaluated. The SWC training used instructors who were professional consultants to provide a version of a training the Center conducts all over the world. The NW3C training relied on a mix of training instructors from within NW3C and intelligence analyst practitioners. These intelligence analyst practitioners were not necessarily professional instructors but had years of field experience. The NJC and the NCLETTTC relied on professional consultants who worked in the field they were teaching about. The NJC instructors were active civil mediators as well as trainers with many years experience. The NCLETTTC used an instructor who had both experience teaching at the university level and first-hand experience in positions of authority in the U. S. Army and a local sheriff's office. In summary, all the instructors observed had the requisite education, subject matter knowledge, and experience as trainers. They differed as to their level of facilitation skills.

In addition, our observations confirmed that all instructors made themselves available to trainees during breaks and after classes. However, the training organizations could have done a better job providing for follow-up contact after the course was completed—when the trainees

returned to their jobs. SWC attempted to build in this contact via its website but had difficulty implementing it during our study observation period. Other courses, such as NCJ and NW3C, allowed for informal follow-up—the trainees could email the instructor—but did not arrange for any type of organized and formal follow-up.

NW3C and SWC developed and implemented a formal training program for instructors along with instructor guides. The other two organizations relied on the fact that the trainers were experienced instructors. In the case of NJC, the instructors were delivering a course they had created for the private sector. NW3C made the most concerted efforts to evaluate instructors using the form in Appendix A.

Training Course Evaluation

One of the most critical aspects of this model is evaluation. The evaluation phase includes the planned evaluation of the training and the trainers. Planning for an evaluation of the training and the trainers during the curriculum planning stage provides a more thorough evaluation. Many training developers and curriculum designers overlook the importance of incorporating and budgeting evaluation into training program development. Only through carefully planned and implemented evaluations can meaningful training feedback be obtained and effectiveness determined. Without this, key stakeholders will never really know the true impact or value of the training.

In our criminal justice training program evaluation examples, each of the organizations planned for a Level 1 evaluation—participant reaction through feedback surveys at the end of the course. NW3C also conducted instructor evaluations using observers. SWC had been doing some degree of Level 3 evaluation—held several focus groups ("plenary sessions") with team leaders six months or more after the training to discuss what the teams had accomplished.

The organizations were not planning Level 2 (participant learning) or Level 4 (organizational impact) evaluations. Other than SWC, the organizations were not planning Level 3 evaluations.

The findings about the levels of evaluation that were planned with these four criminal justice training programs are not uncommon. Unless training organizations require that participants develop an action plan to implement the training before the training ends, and the

participants' agency demonstrates commitment to use the training, it becomes much more difficult to expect any Level 3 or Level 4 change or to monitor participants' translation of training content into action.

In reviewing criminal justice training around the country as part of this project, we could find very few training organizations that implemented anything more than a Level 1 evaluation of the training.

Level 1 Applications

Each of the training programs implemented a variation of a standard participant training reaction survey form. SWC used a separate feedback form for each day of training. The others administered the forms at the end of training. Most were able to obtain adequate samples. The forms used were all professionally acceptable examples.⁹ Examples are in the appendices of each individual evaluation report.

In the SWC training, the evaluation team also tested using personal interviews by phone to obtain participants' reactions to training one to six months after the training. This method worked and produced rich and detailed feedback. However, it was time-consuming and expensive. Even with many attempts to each person, we were only able to complete interviews with 65 of 105 participants (62 percent response rate). The difference compared to obtaining participant feedback immediately after training was that because of the time lag—one to six months after training, the evaluator was able to obtain some feedback on learning (Level 2) and implementation on the job (Level 3).

We also tested an instructor's rating instrument in the NCLETTTC training program evaluation.

Each of the training programs used the feedback and results from the Level 1 surveys to varying degrees. NW3C actually made changes to the training schedules and curriculum after the first two training sessions due to participant feedback.

If possible, it is often useful to analyze the Level 1 evaluation data (as well as Levels 2 and 3) in terms of training participants' demographic and background characteristics. For example, can we determine any differences in reactions/satisfaction, test results, or job

⁹ The reaction form used with the NCLETTTC training was developed by the study evaluation team.

performance by trainees in terms of their sex, age, race/ethnicity, work experience, education, or other characteristics. Then, test to determine whether differences, if any, are statistically significant (apply possible range of statistical social sciences tests). Illustrations of this type of analyses are found in the evaluation reports on the NW3C training (Chapter 7) and the NCLETTTC training (Chapter 9).

Level 2 Applications

Several different evaluation instruments were implemented at the Level 2 evaluation stage. See Exhibit 6-1 below for illustrations.

Exhibit 6-1: Evaluation Instruments Used at Level 2 (Learning) for the Four Training Programs

<u>Training Program</u>	<u>Evaluation Instruments</u>
NW3C	Pre/post knowledge test; completed at the training
NJC	Pre/post knowledge test; completed at the training
NCLETTTC	Pre/post knowledge tests; (1) completed at the training; (2) completed six months after the training
SWC	Pre-training interviews (team leaders) Follow up surveys by mail (team leaders) Reaction interviews with participants (1-6 months after training)

Knowledge tests were implemented before and immediately after the training in three of the four training programs. In the SWC training we tested a pre-interview, post-survey technique. In addition, a follow-up post training test was administered to a sample of attendees at the NCLETTTC training. The tests all produced useful results.

It was not too difficult, time consuming, or expensive to develop pre/post tests that revealed cognitive measures at the lower to mid levels of the Bloom learning taxonomy (knowledge, comprehension, application, analysis). Research and training program staff

accomplished this with multiple choice and true/false questions that mirrored curriculum content. However, it was more challenging to try and measure the higher levels of cognitive functioning—synthesis and evaluation. This effort deserves more research attention in the future.

Application of the follow-up knowledge test also produced useful results—in fact, in the NCLETTTC training program, showing some dissipation of learning six months or more after the training; although the test was not accurate in pinpointing the reasons why this occurred. The testing process also suffered a low response rate.

The pre-interview, post-survey technique also worked but had some shortcomings. First, while the results did show some changes in knowledge and attitudes, the changes were not readily quantifiable because answers tended to be varied. Second, the test was only done with the SWC training team leaders—not all team members—because of the costs involved and difficulty of contacting all team members. In retrospect, we should have conducted the post-test by phone interview not mail survey. While more time consuming and expensive, the results might have contained more detail.

Level 3 Applications

Two main types of evaluation instruments were implemented at the Level 3 evaluation stage. See Exhibit 6-2 for illustrations.

Follow-up surveys were mailed to training participants in all four training programs. In two of the programs, NW3C and NCLETTTC, follow-up surveys were also mailed to the training participants' immediate supervisors. In the NW3C evaluation, we also tested a self-assessment rating tool that measured the participants' comfort levels with applying the main intelligence analysis techniques in the workplace.

In the SWC training, we also analyzed the strategic plans that each training team developed during the training. While the plans did show that the teams were able to use synthesis and evaluation skills to make judgments and decisions about future courses of action to deal with hate crimes (a Level 2 skill learned), the real value of the exercise was for evaluators to assess the progress that teams made in implanting the plans six months or more after the training. These were very specific tasks that could be followed up.

Exhibit 6-2: Evaluation Instruments Used at Level 3 (Behavior Change) for the Four Training Programs

<u>Training Program</u>	<u>Evaluation Instruments</u>
NW3C	Pre/post self-assessment performance tests (post-test by mail) Follow-up self-report survey by trainees (mail) Follow-up survey of trainees' supervisors (mail)
NJC	Follow-up self-report survey by trainees (mail)
NCLETTTC	Follow-up self-report survey by trainees (mail) Follow-up survey of trainees' supervisors (mail)
SWC	Follow-up self-report survey by team leaders (mail) Analyses of strategic plans

In all cases, the test instruments worked to varying degrees and produced a variety of information. Each individual evaluation report contains more detailed findings of what these instruments produced.

We did find some common themes as follows:

- It was challenging to obtain anything close to 100 percent sample with trainees or supervisors. This raises several issues: (1) would better incentives have helped? (2) would obtaining organizational commitment up front have helped? (3) would conducting the follow up by phone interview have helped?
- The instruments needed to be more objective and contain more quantifiable information. We needed to obtain more "facts" than "opinions." The best instrument would be able to measure the number of times the trainee now

performed the skills he or she was trained to perform (compared to the number of times listed in pre-training reports).

Level 4 Applications

Our most challenging goal was to conduct Level 4 evaluations using quasi-experimental designs. However, this proved more difficult than we thought. Factors that affected our design choice included cooperation and assistance from the training participants' agencies and potential comparison group participants.

Key to a quasi-experimental design is obtaining a comparison group. At the outset of this evaluation, we had hoped to obtain comparison groups from a variety of sources—for example, colleagues of those attending training, matched by level of experience and education, and from comparable training classes offered by other providers. The main factor that affected our evaluation designs and prevented us from conducting a quasi-experimental design was the lack of adequate comparison groups.

With two of the trainings, NJC and SWC, it simply was not possible to obtain a comparable comparison group. The NJC *Civil Mediation* training had a small class of 25 judges offered twice a year. It was not reliable to try and match the judge who came to training with another single judge back in the same court. In addition, most of the judge-trainees did not immediately begin to use the civil mediation skills learned in the training when they returned to work. The SWC training involved a jurisdiction sending a multidisciplinary team—judge, prosecutor, police officer, probation officer, teacher, etc. It would have been difficult to match each team with another comparable team back home.

With the NW3C's FIAT training, we proposed two comparison groups: an alternative training class, and colleagues of the FIAT trainees. Neither of these groups worked out. After a national search, we were able to find only one training class for intelligence analysts similar to NW3C's FIAT course—in fact, the organization with the class partnered with NW3C to help develop the FIAT course. However, we simply could never get adequate cooperation from the executive director and staff of the alternative training program to negotiate an arrangement to use them as a comparison group. Also, because of the tight training schedule for the FIAT course, we could not arrange for a suitable comparison group from the trainees' agencies.

We did attempt two types of Level 4 evaluation with the SWC and NCLETTC programs. In the SWC program, evaluation team members conducted on-site case studies at two of the teams that had received training—teams from New Jersey and Wisconsin. While clearly a small sample, the results did show that both sites implemented changes based on the training. The Wisconsin team formed a county-wide hate crimes task force and the New Jersey team changed police academy training and formed a juvenile offenders program—both organizational impacts.

In evaluating the NCLETTC program, study staff tested organizational assessment instruments before and after the training. While the results were mixed, the utility of the measurement instruments on an organization-wide basis showed promise.

More detailed discussions of the methods, challenges, and findings for evaluating each program are provided in the individual evaluation reports in Chapters 7-10.

Conclusions

One of the main lessons learned from the test of the training evaluation model is that evaluating training—although a logical and rather straightforward process on paper—can be challenging in practice, especially if there is (1) no planning for evaluation during the training planning phase and (2) the aim is to conduct a meaningful evaluation at Levels 3 or 4. The higher level training evaluations can be especially complicated if training recipients are from agencies that are scattered across the country. This gives the evaluation team little control over agency participation in the evaluation follow-up.

While planning an evaluation is not a simple task, it is possible and should be a mandatory aspect of all OJP-funded criminal justice training. Currently, most criminal justice training programs do conduct a Level 1 assessment of participants' reactions to the training. In most criminal justice training, training providers should expand to a Level 2 evaluation to measure the extent to which changes occurred in learning and knowledge. No training is planned without a purpose. All training programs have some goals and objectives that they want to meet to benefit the recipients. These can take the form of information or skills to be learned, or attitudes to be changed. Training providers should plan to capture data that indicate whether the training participants achieved the training objectives. This can be done through pre-post tests, conducted immediately prior to and at the conclusion of training.

In addition, criminal justice training providers should commit to evaluate behavior change (Level 3), at least with a sample of the participants. A Level 3 evaluation is very informative and not nearly as complicated or costly as a Level 4 evaluation. The results provide critical information to the training providers, funding agencies, and attendees. If a Level 3 evaluation shows that training is achieving behavioral changes in trainees back at their jobs, then funding agencies, such as BJA and other OJP agencies, would be justified in providing continuing funds for these training programs. Federal agencies that fund criminal justice training should require, and provide funds for, Level 3 evaluations as a routine aspect of their financial support for training programs.

In the end, all stakeholders, including the training provider, funding agencies, attendees and their agencies want to know whether the training was effective. Being effective means: Did the participants learn the material identified in the training objectives? Did the learning have a positive impact on their work and/or their agencies? A successful training program will have a high quality curriculum in place and capable instructors. A quality curriculum requires advance preparation in determining the need for the training, soliciting advice from SMEs, balancing the need for lengthy, detailed training against the needs of participants and their respective agencies, and recognizing that seemingly minor details such as length of breaks, comfort of trainees, and location of training are important. Even the best curricula will be ineffective if the instructors are not highly qualified in the subject area or do not have the ability to facilitate and communicate. Thus, it is equally important to select and train qualified instructors and evaluate their performance.

Recommendations and Lessons Learned: Tips for Evaluating and Improving Criminal Justice Training

The study team learned from the experiences of using the expanded Kirkpatrick training evaluation model to evaluate four diverse national criminal justice training programs. Our experiences afford us the knowledge to make a number of recommendations for NIJ, BJA, and other OJP agencies that fund criminal justice training. We also posit some *lessons learned* for federal agencies and criminal justice organizations involved in the business of developing and managing training.

With some trial and error, hard work from all study team members, great support from the four training organizations that were evaluated, and support from NIJ and BJA staff, we demonstrated that the model works. We also have provided insights into what parts of the model work well—and should always be done; and what parts are the most challenging—and may be attempted with enough planning, resources, and support.

It is our view that using the evaluation model developed and tested in this project will allow OJP agencies to obtain more consistency and control over large and expensive criminal justice training programs. OJP should take the lead in promoting *best practices* in the criminal justice training field. Developing a rigorous training evaluation effort will encourage training organizations to implement these best practices because their training programs' effectiveness and costs/benefits will be documented and available for comparison.

Recommendations

1. OJP should develop training standards for organizations that receive federal funding for criminal justice training. These standards should focus on the process of developing, delivering, and evaluating training, not the substance.

Every state in the U.S. has a criminal justice standards and training commission. These commissions establish standards and approve training courses delivered to certified police and corrections officers in their states. They set requirements for certification and training. They have extensive experience in developing learning objectives, lessons plans, and tests. OJP should identify a group of key experts from state training commissions, training administrators from state and local law enforcement, and academic training professionals and bring them together to develop standards for OJP-funded criminal justice training. Some suggested standards are contained in some of the below recommendations. This effort would help produce more consistency in the quality of criminal justice training around the nation.

2. OJP-funded criminal justice training should require a comprehensive needs assessment and training plan before training funds are committed.

No OJP federal training funds should be committed to any organization without a thorough and complete training needs analysis that clearly justifies the need for the training. Too often, organizations have received earmarked funds to develop some type of criminal justice training that was duplicative of training already being offered somewhere else. This type of

"reinventing the wheel" has become all too common in criminal justice training. In addition, the training plan will require the training organizations to document comprehensive plans to manage and support the training effort. The training plan should also detail the selection criteria that the training organizations will use to recruit training participants. As well, the training organizations should describe how they plan to obtain commitment from training participant's agencies that the agencies will ensure that the trainees will use the training acquired when back on the job. The details of needs assessments and training plans are contained in Chapter 4 of this report.

3. All large-scale OJP-funded criminal justice training should include evaluation plans that incorporate evaluation Levels 1-3.

An evaluation plan should be included with any OJP-funded training that exceeds \$500,000. The evaluation should be funded at a minimum of 10 percent of the total training funds. The federal agencies need to review each evaluation plan and decide if it can be implemented by the training organization's staff or if outside experts (e.g., academics, consultants) should manage it. In the evaluation plan, the training organization needs to demonstrate that it can obtain the commitment of the training participants' agencies to cooperate with the evaluation requirements.

A Level 1 evaluation is required to show participants' immediate reaction and satisfaction with the training. A Level 2 evaluation is required—where applicable based on training that attempts to transfer knowledge, skills, or attitude changes—to show that (1) suitable trainees are taking the course and (2) they are learning or changing attitudes. The goal of the Level 2 test should be to strive to reach the highest levels on Bloom's taxonomy of learning.

A Level 3 evaluation should also be a requirement with at least a significant sample of participants. At a minimum, information should be obtained from participants' supervisors six months after the training to assess the extent to which the trainees are using the acquired training on the job. While not a requirement because of the complexities and challenges involved, any training organization that proposes a Level 4 evaluation should be given priority in the funding decisions.

4. All large-scale OJP-funded criminal justice training should include a pilot test of the course before it is implemented in the field.

The pilot test should be described in the training plan. A detailed pilot test may be the most important step in the training development process—it needs to be done well. What we found overall in the review of pilot tests was that the expected time for exercises is often underestimated. They always take longer in reality than the developers think. Additionally, the time spent on debriefing a completed exercise in front of the whole class to enhance the learning experience for all trainees is often underestimated. Unfortunately, when time is running short, this is the area that many trainers opt to cut from the module—when in reality, this may be the most important part of the module.

5. All large-scale OJP-funded criminal justice training should be required to offer a system of follow-up technical assistance to support the ongoing learning process.

National-level OJP-funded criminal justice training should offer some type of organized follow-up for student questions via email or website. Instructors should be encouraged and paid to answer follow-up questions as part of their training responsibilities. In one of the projects evaluated, many of the attendees commented that this would be an extremely helpful addition to the training. Follow-up technical assistance would be especially valuable when a training course involves developing and implementing policies or strategies.

6. OJP should fund applications of the training evaluation model developed in this project with law enforcement and corrections recruit training academies as continued research on the utility of the training evaluation model.

The application of the training evaluation model developed in this project was limited to four BJA-funded earmarked training programs. The model should be tested on police and corrections entry-level training academies. Police and corrections agencies spend hundreds of millions of dollars a year training new recruits. These agencies probably train over 50,000-60,000 new officers every year. This doesn't even include new law enforcement and corrections officers trained at the federal level (easily another 15,000-20,000 per year).

Training of new recruits at the academy level could truly benefit from more comprehensive evaluations. For example: To what extent are field training officer evaluations and other department evaluations informing the academy's curriculum development process? What new topics should be offered or given more in-depth treatment? What topics should be eliminated or modified based on Level 3 evaluation results (i.e., the knowledge gained is not

being applied in the field to the extent anticipated)? What can academies do about the perpetual issue of new officers “unlearning” certain techniques or attitudes because of other influences on the job? The potential improvements to academy training would include improved consistency, improved quality, better focus of resources, and potentially cost savings (by eliminating and changing what isn't working).

It would also prove useful to examine the training by what works with certain individuals by comparing training outcomes according to age of recruit, sex of recruit, minority status, and education levels. The findings might help academies tailor the training in certain subjects to the individual needs of the learners.

In the police and corrections training academies context, we recommend that the research also involve the following: (a) more research on developing pre-post knowledge tests (Level 2) that measure the highest cognitive levels on Bloom's taxonomy; and (b) more research on applying comparison group methodologies at Levels 3 and 4.

7. OJP should fund applications of the training evaluation model developed in this project with online and computer-based training courses in the criminal justice field.

In this project, we attempted to apply the training evaluation model to an online version of the training delivered by the NCLETTTC. The test demonstrated that the processes of the model work. However, more testing is needed with online and computer-based training courses. Training providers—criminal justice agencies and training organizations—are rapidly expanding the delivery of training by computer and the Internet. Clearly the new delivery modes are more efficient and flexible than traditional classroom training—many more students can be reached often at the student's own time and pace. However, very few evaluations have been conducted of computer-based courses designed specifically for police and other criminal justice personnel. The limited evaluation findings in the current study did show some differences compared to classroom training.

Before offers of online and computer-based criminal justice training advance from a trickle to a proliferation, OJP should take a leadership role in evaluating these courses, identifying strengths and weaknesses, and developing best practices.

8. *OJP should fund research into the motivations for why criminal justice personnel continue to learn through training and academics.*

In a recent *Time* magazine interview, Bill Gates when asked about the future of improvements in learning through technology, responded by saying, "Learning is mostly about creating a context for motivation. It's about why should you learn things." Clearly, motivation is one of the keys to learning.

OJP should also take a leadership role and fund research into the motivational factors that encourage or discourage criminal justice practitioners from continuing to advance their learning in both the academic field and the training field. Once we have identified these motivational factors, we can encourage criminal justice agencies, training organizations, and academic institutions to build on them to redesign learning for criminal justice practitioners.

Lessons Learned

These lessons learned are directed at OJP agencies that fund criminal justice training and also all the training organizations that develop and deliver criminal justice training. In our research around the country, we have found many capable organizations that deliver quality criminal justice training, like the four programs that were evaluated as part of this project. Below are some lessons from these projects and others that can help improve criminal justice training.

1. Begin the evaluation at the earliest developmental stages of the training.

By committing to an evaluation early on, the training organization can design all the materials, methods, and other aspects of training with evaluation data collection in mind. All the evaluation instruments (Level 1—reaction; Level 2—knowledge tests; and Level 3—follow-up to assess behavior change) can be designed to clearly fit the training. Instructors can be retained who are supportive, not fearful, of the evaluation. Everyone involved can commit to the cycle of development → evaluation → revision depicted in Chapter 4, Exhibit 4-2.

2. Develop clear, measurable, and meaningful learning objectives.

The learning objectives are the foundation of the training and the focus of the evaluation. The course content is designed to meet the learning objectives. Take time and care in crafting

the learning objectives. Each word in the statement of objectives is important. The statements must be clear and meaningful. Instructors should teach to the objectives and check and recheck to ensure that they are teaching to the objectives. Evaluators will evaluate against the objectives—to what extent were they achieved?

3. Screen prospective training participants and obtain organizational commitment.

Training organizations should develop criteria for prospective applicants to meet to attend the training. Applicants should be screened to determine that they meet the criteria. This can be done via an application process over the organization's website. This is necessary to ensure that the organization is training practitioners who will apply the learning on the job. As well, training organizations should obtain some degree of commitment from the applicant's criminal justice agency that they will encourage and support the application of the learning when the trainee returns to his or her job.

4. Ensure that the training is developed using adult learning principles.

Criminal justice training organizations have the best intentions of using adult learning principles but often get caught up in a situation of limited time and resources and resort to standard classroom training. The emphasis on adult learning principles should drive the development of the curriculum and teaching methods. We have provided a number of examples and references in this report. Adult learning principles should become standard practice in criminal justice training. The main emphasis of the pilot test should be to reinforce the use of adult learning principles and practices.

5. Evaluate instructors and encourage them to develop facilitation skills.

There is a tendency for training organizations to use experienced practitioners who have delivered training in their agencies, or other agencies, to deliver criminal justice training. Many of these instructors, while very knowledgeable and experienced, sometimes don't have well-developed facilitation skills. Even the best instructors can benefit from additional education in developing and honing facilitation skills. In major training programs, the training organizations should develop and deliver a module on facilitation skills to all prospective instructors.

Instructors should always be evaluated in the student reaction survey and the instructors should be aware of the rating criteria being used. However, in many cases, where resources are available, instructors should also be evaluated by trained SMEs using standard rating forms (see examples in the Appendix A of this report).

6. Take the training pilot test seriously and devote ample time to conduct it carefully.

Take more time to hold the pilot test than you think you need—you will always need more time. The pilot test should not be the first time the training course is offered. The pilot test should be a *trial run* conducted in front of SMEs and others who were involved in developing the training modules. To use a football analogy, the pilot test should be an internal scrimmage, not an exhibition game against an opposing team. Actually spend ample time role playing the exercises—and debriefing of the exercises. These are often the most critical, yet under-practiced, parts of the training.

7. Evaluation staff should be observers and attend each delivery of the training that is being evaluated.

This project's evaluation team erred in not attending each training class being evaluated. In order to conserve evaluation funds, we sometimes relied on the instructors or administrative staff of the training organizations to distribute the evaluation instruments. In retrospect, this was a shortcoming of our methodology. Some mistakes were made and some data were not collected accurately. This was a lesson for us all.

8. Level 1 evaluation (reaction) tips:

- Keep the student reaction ("satisfaction") form simple—stress the need for comments to explain their rating choices
- Obtain participants' reactions at the end of each day of training
- Attempt to get 100 percent sample—use encouragement and incentives
- Obtain reactions from instructors after every module they teach

9. Level 2 evaluation (knowledge) tips:

- Training objectives will govern the relationship between test items and the levels on Bloom's taxonomy; however, wherever possible, attempt to construct tests that reach the highest cognitive levels on the taxonomy—may need to use open-ended questions.

- Consider administering the post-training knowledge test again six months subsequent to the training with a sample of participants to measure learning retention or dissipation.
- Obtain participants' commitments (and their agencies' commitments) during the training application stage to complete the follow-up surveys or cooperate with the interviews.

10. Level 3 evaluation (behavior change) tips:

- In conducting follow-up assessments of the application of learning, ensure that the trainees have had ample time, opportunity, and resources to use the learning on the job.

In the NW3C evaluation, most trainees were using the learned techniques as soon as they returned to their jobs. In the NJC evaluation, most of the judges were going to wait until they retired before they actually became court mediators, although many did apply some of the techniques when they returned to their jobs.

- The most efficient process is to send self-report follow-up surveys to trainees six months or more after the training.

This survey should ask for responses in terms of specific and quantifiable frequencies where the learning is applied on the job (e.g., how many times a day (or shift, etc.) do you perform XXX). The responses can be compared to the responses on a pre-training survey.

The timing for the follow-up varies from training to training. In the NW3C training, six months was ample time for the intelligence analysts to have applied their new skills. However, in the SWC team training, some of the teams had not yet even met again to work on implementing their strategies six months after the training.

- A more objective process is to additionally send the same survey to the trainee's supervisor.

The trainees themselves may have reasons to be less than objective about demonstrating that they are applying what they have learned. Hopefully, their supervisors will be more objective. However, as mentioned above, the supervisors must have some incentive and commitment to take the time to respond to the surveys. Obviously, if resources permit, phone interviews versus mail surveys provide more detailed feedback on the application of the training.

- Finding a comparison group to evaluate trainees against would be an effective implementation of Level 3.

Establishing a quasi-experimental design with an experimental and control group is challenging yet not impossible. The lesson we learned is that the evaluation design clearly needs to be established at the beginning of training in collaboration among evaluators, training organization, and the trainees' agencies. It would work best where an agency sends one of several work units to training. For example, a police agency sends one of several patrol squads to learn how to implement a specific skill (e.g., use GIS mapping data for problem solving). We can then compare the problem solving statistics of the squad of trained patrol officers with squads that were not trained.

11. Level 4 evaluation (organizational impact) tips:

- While implementation of a Level 4 evaluation is challenging and rare, the feedback can be very powerful to substantiate the value of the training.

The standard methodology to implement the Level 4 evaluation is to use a quasi-experimental design with a control and experimental group as discussed just above and in the report in Chapter 4. The evaluators are looking for evidence of the training's impact at the organizational level. While better defined Level 4 measures exist in the private sector, criminal justice organizations do have some useful measures such as work output, work quality, customer satisfaction, policy changes, legislative changes, and more.

However, there are alternatives to using comparisons groups at Level 4. In this project, we attempted two variations of Level 4 evaluation methodologies—the case study approach and using organizational climate surveys. Both examples showed strengths and weaknesses.

One of the main issues with attempting a Level 4 evaluation is that enough personnel need to receive the training to have an organizational impact. Too often because of resource constraints and scheduling, criminal justice agencies are only able to send one or two personnel to outside training versus an entire unit, squad, or division. Thus, with so few members of a unit being trained, we can't expect to see a change impact on the entire unit's work.

This is why the application of the Level 4 evaluation stage may work best in evaluating criminal justice entry-level training—if the agencies would cooperate in establishing control groups.

Costs of Training

The main challenge to conducting cost benefit or ROI analyses of training is to identify actual costs and tangible benefits. This is often more difficult than it seems.

Throughout the report, we have discussed the challenges to identifying the training outcomes—Levels 3 and 4 evaluations. In addition, it is often challenging to isolate and calculate training program costs. For example, some training organizations don't keep cost figures in a detailed and itemized way. For example, administrative staff may work on several projects in a week and not be able to extract that they spent X hours on the specific training program being evaluated. If outside SMEs are not being paid by the training organization but volunteering their time, is this still a type of cost that should be captured?

There are a variety of questions regarding cost benefit analyses that should be sorted out and become a part of any effort by OJP to develop criminal justice training standards for training organizations using federal funds to deliver training.

While it may be difficult to conduct comprehensive cost benefit studies, it should be more straightforward for OJP to simply compare the costs of training by organizations using federal funds. The cost benefit analysis includes identifying the costs to the trainee's agency, a difficult calculation to make and obtain. On the other hand, cost comparisons of training costs require cost information only from the training organizations that receive the federal funds.

Training program costs for training organizations can be subdivided into the following cost areas:¹⁰

- **Training development costs:** This includes mainly staff time devoted to developmental efforts such as the project manager, staff professionals (curriculum content), staff administrative personnel (e.g., word processing), staff technical personnel (e.g., media design, video production, etc.), and outside SMEs (curriculum content). Costs in this category would also include travel and lodging costs if SMEs or the organization's staff had to travel from another location to meet and hotel costs if the development meetings took place in a hotel.
- **Materials and supplies:** Costs in this category include copying or printing student notebooks, surveys/tests, and other training supplies such as student nametags/tents, pads of writing paper, 3x5 cards, and other training materials.

¹⁰ See Head (1994) for additional examples and explanations.

- **Facilities support:** This includes all costs associated with holding the training in a facility (e.g., hotel) such as room rent, food and beverages, audiovisual equipment rent, and more.
- **Administrative/management support:** Costs in this category include management, staff, and instructor labor time involved in delivering the training. This typically includes task areas such as advertising and application efforts (e.g., website development costs), registering students, managing students at training, logistics efforts during training (dealing with hotel staff on facilities, food and beverages, audiovisual equipment, etc.), and training time. Costs in this category (although they could also be in "materials and supplies") also include shipping materials to the training facilities. Additional costs in this category often include travel and lodging for staff and instructors. If the training organization is paying the travel and lodging of the training participants, the costs are accounted for in this category.¹¹

Staff costs are calculated by multiplying the labor time (in hours) allocated to a training task or area by the hourly salary rate of the individual. Costs for outside SME instructors are simply the arranged fees for services. Costs for facilities, materials, supplies and other direct costs are the actual costs.

Thus, the accumulation of all the above costs adds up to the total costs for developing and delivering the training program. Obviously, these costs continue to increase with each additional training class.

To make comparisons among training programs, the total program costs need to be divided by an acceptable denominator such as number of students trained (cost per student) or number of students trained in a training day (cost per student per training day). Since the development costs have already been invested, as the training is delivered to more students, the cost per student decreases.

Any tuition charged for attending the training needs to be added to offset the training costs. In the four training programs evaluated, only NJC charged tuition (and also provided scholarship assistance in certain cases).

The above cost comparisons are helpful to some degree but would be much more useful if you also had evaluation data on training effectiveness. These are all decisions that can be

¹¹ In the four training programs evaluated, only SWC paid travel and lodging for attendees.

made in implementing recommendation number one above—to develop standards for training organizations using federal funds to train criminal justice practitioners.

Chapter 7

National White Collar Crime Center's Foundations of Intelligence Analysis Training

The National White Collar Crime Center

The National White Collar Crime Center (NW3C) is a Congressionally-funded, non-profit corporation with a mission to provide resources and support to state and local law enforcement agencies involved in the prevention, investigation, and prosecution of economic and high-tech crimes. Further, as a result of the nationwide emphasis on combating terrorism, the NW3C supports and partners with other organizations to address homeland security initiatives centering on economic and high-tech crimes. As a private, non-profit organization, the NW3C has no investigative authority; instead, it assists law enforcement agencies in understanding and combating economic and high-tech crimes.

History and Background

The predecessor to the NW3C, the Leviticus Project, was originally established in 1978 as an anti-crime program dealing with fraud in the coal mining field. The Leviticus Project, federally-funded since 1980, evolved from exclusive focus on coal mining fraud to covering all oil, natural gas, and precious metal fraud, to its current broader scope focusing on all types of economic and high-tech crimes. The Leviticus Project changed names to NW3C in 1992 to better reflect its expanded mission. Since 1992, NW3C has continued to expand in providing training, research support, and technical assistance on economic and high-tech state and local crimes. After the attacks on the United States of September 11, 2001, its mission was further expanded to include domestic homeland security initiatives.

The NW3C is headquartered in Richmond, Virginia, and operates offices in Fairmont and Morgantown, West Virginia. NW3C activities are directed by an elected board composed of a chairperson and seven members from different designated regions across the country. Over 1,700 agencies are members of NW3C, including local and state law enforcement agencies, state regulatory agencies, and local and state prosecution offices. While there is no fee to become a member, membership is required to attend training sessions and receive investigative support.

Organizations that endorse the activities of the NW3C include the International Association of Chiefs of Police (IACP), Fraternal Order of Police (FOP), National Organization of Black Law Enforcement Executives (NOBLE), and National Sheriffs' Association (NSA).

Center Services

The NW3C provides a wide range of services for its members including training, research, intelligence and analysis support, fraud complaint management, and annual economic crime summits.

Training

The NW3C provides classroom and computer-based trainings for law enforcement personnel at various locations across the country on the investigation of economic and high-tech crimes. Specifically, NW3C offers the following courses:

- FIAT (Foundations of Intelligence Analysis Training): Provides training for law enforcement analysts on the tools necessary to successfully use intelligence analysis
- FIPS (Financial Investigations Practical Skills): Provides training for economic crime investigations
- FREA (Financial Records Examination Analysis): Provides training on using computers to examine and analyze financial records and present evidence
- BOTS (Basic On-line Technical Skills): Provides basic instruction for law enforcement new to on-line investigations
- BDRA (Basic Data Recovery and Analysis): Provides basic instruction to assist in the investigation and prosecution of cyber crime cases
- ADRA (Advanced Data Recovery and Analysis): Provides advanced training on computer data recovery and analysis. Six separate ADRA classes are available including:
 - Windows 95 through ME
 - Windows NT through XP
 - Online platforms, including Internet Explorer, Netscape Navigator, America Online, and IM software
 - Hard drive and computer media forensics through *ILook Investigator*
 - Email forensics
 - Automated forensics tools

The NW3C has trained more than 60,000 law enforcement personnel nationwide. Training classes are held in classrooms at NW3C or hosted by member agencies and held in their facilities. In addition, NW3C provides training via DVD, CD, video, and the Internet. For members of NW3C, training is provided free, although students must pay for their own transportation, lodging, and meals.

Intelligence and Analysis Support

The NW3C provides two types of investigative support: short-term support on requests for information on specific individuals and businesses and long-term analytical support on major financial crime investigations. NW3C fulfills the short-term requests by accessing databases such as Lexis-Nexis and ChoicePoint. For long-term analytical support, NW3C has trained analysts with expertise in developing and analyzing databases containing financial information (such as bank and credit card statements).

Short-term investigative support is provided from a staff of six trained personnel in the Investigative Support section of NW3C. The staff handles over 700 case requests each year—one request may have many individuals and businesses for which information is sought by an investigator. In total, the NW3C personnel for short-term investigative support do more than 15,000 queries each year. To obtain services, a member agency completes a form indicating the information it needs for an investigation. Generally, a request consists of a list of individuals or businesses for which the investigator is seeking current address and other information available from public databases. Once the request form arrives, NW3C staff access databases supported by Lexis-Nexis, ChoicePoint, and others to collect information on the specified individuals or companies. For example, a list of 15 individuals or businesses usually will require about two hours of effort on the part of staff. The NW3C staff edits and formats the information they obtain for easy readability, after which the results are emailed or faxed back to the requesting member. These services are available to all NW3C member agencies at no cost, and turnaround time of results back to the agency is almost always within one business day.

Eight analysts within the Investigative Support section provide analytical support on major cases. These analysts are assigned about 20 new cases each year. They are usually asked to analyze large amounts of financial data that make up the primary evidence of a financial crime case. Analysis most often requires significant data entry by the analyst, leading to the

development of spreadsheets, exhibits, and link analysis charts to illustrate the connections and flow of money or other data in a case. An analyst may spend several months on one case. The board of directors must approve each long-term analytical support effort. The requesting agency submits paperwork to the NW3C outlining the facts of the case and the support that is needed. This level of analytical support is not available within the requesting agency, and the investigation and prosecution could not move forward without it.

Fraud Complaint Management

In partnership with the FBI, NW3C established the Internet Crime Complaint Center (IC3) in May 2000 to address fraud committed over the Internet. The mission of IC3 is to serve as a mechanism for receiving, developing, and referring criminal complaints regarding cyber crime. The IC3 website (www.ic3.gov) provides a convenient and easy-to-use reporting mechanism for victims to report Internet fraud. In addition, the IC3 provides a central repository for the sharing of fraud data by all law enforcement and regulatory authorities. Complaints received by the IC3 include Internet auction fraud, non-delivery of merchandise and payments, credit/debit card fraud, investment fraud, business fraud, confidence fraud, and identity theft. IC3 analysts review and evaluate each complaint in order to refer it to the appropriate law enforcement agency. According to the *IC3 2003 Internet Fraud Report* (NW3C & FBI 2004) IC3 received 124,506 complaints in 2003, which reflects a 60 percent increase from 2002. Of these complaints, with a total dollar loss of \$125.6 million, 95,064 were referred to law enforcement.

Research

NW3C created a research section in September 1995 with the mission of identifying the impact of economic and high-tech crime as a means of increasing both public awareness of the problem and law enforcement's ability to deal with it. The research section conducts original research on economic and high-tech crime, maintains a library of information on these issues, and evaluates different NW3C training programs and conferences. The research section of the NW3C uses a range of study methods including surveys, field research, secondary data analysis, policy evaluation, and program evaluation. Recent research initiatives undertaken by the center have produced a variety of papers and research reports, including:

- Check fraud
- Cyberstalking
- Credit card fraud
- Disaster fraud
- Embezzlement and employee theft
- Health care fraud
- Identity theft
- Insurance fraud
- Internet gambling
- Telemarketing fraud

Additional tasks performed by the NW3C research section include providing analytical support to the IC3 and the White Collar Crime Research Consortium. Finally, through the extensive library holdings of NW3C, the research team is able to provide legal guidance related to economic crimes, including case law, novel and complex legal issues in white collar and high-tech crime, and legislative developments.

Economic Crime Summits and Outreach Seminars

The NW3C has held regional economic crime summits for the past 10 years. These summits highlight current trends and initiatives in economic crime for law enforcement and fraud prevention specialists from the public and private sectors to share knowledge, skills, and experiences. Representative attendees include members of academic organizations, government agencies, private corporations, victim interest groups, and crime prevention specialists. In addition, NW3C holds free one-day seminars several times a year throughout the country. The seminars focus on electronic law enforcement and computer crime trends, although the specific topics vary according to need. These seminars are for prosecutors, crime prevention, and enforcement professionals.

Review of the Intelligence Literature

In the wake of September 11, 2001, intelligence analysis has become increasingly more important to federal, state, and local law enforcement agencies. Intelligence analysis can take many forms and can be defined in a multitude of ways. Intelligence analysis has long been used by the military to track movement of wartime enemies and by federal law enforcement agencies to investigate criminal enterprises, including organized crime and illegal drug distribution networks. After the September 11th terrorist attacks, the importance of intelligence, including

conducting, disseminating, and coordinating with other agencies, was viewed as a principal mechanism for preventing “future tragedies” (U.S. Department of Justice 2005).

The Global Intelligence Working Group (GIWG), formed after September 11, 2001, defines intelligence as “...the combination of credible information with quality analysis information that has been evaluated and used to draw conclusions (U.S. Department of Justice, 2004, p. 3). In a slightly different vein, Smith (1997) states that intelligence involves “...the collection and analysis of information to produce an intelligence end product designed to inform police decision making at both the tactical and strategic levels” (p. 1). Others differentiate short-term, tactical analysis from long-term, strategic analysis (Peterson 1997).

Intelligence-led Policing

Intelligence-led policing is a model of policing that applies the collection and analysis of information to facilitate crime reduction and prevention by informing police decision making at both the tactical and strategic levels. In essence, intelligence serves to guide operations rather than the reverse (Smith 1997). The concept of intelligence-led policing originated in Great Britain in the 1990s and quickly spread to Australia and Canada.

More recently, intelligence-led policing concepts have been developing in the United States. Since the terrorist attacks on September 11, 2001, there have been calls for improved crime-related intelligence with a greater emphasis placed on coordinated intelligence planning and sharing between law enforcement agencies. It is thought that improving agencies’ capacity to share intelligence could significantly improve their ability to protect public safety. In fact, President Bush “...pledged to make information sharing an important tool in the nation’s war on terror” (U.S. Department of Justice 2005, p. 1).

The U. S. Department of Homeland Security (DHS) was created in November 2002, with the chief goal of implementing a comprehensive national security strategy. Prior to creation of DHS, there were more than 100 government agencies and programs responsible for various aspects of national security, including but not limited to intelligence, border patrol, communications, emergency preparedness and response, and immigration. Paramount to its mission, the DHS unifies border and transportation security policy, coordinates disaster response, creates a central point for analysis and dissemination of intelligence, and fosters research and

development efforts (DHS 2005). Key to accomplishing DHS goals to provide for a secure homeland is a coordinated, comprehensive national domestic intelligence plan.

National Intelligence Plan

To meet the need for improved and coordinated intelligence, the International Association of Chiefs of Police (IACP) held a *Criminal Intelligence Sharing Summit* in early 2002.¹² The summit brought together law enforcement executives and intelligence experts for the purpose of producing an intelligence sharing plan that would coordinate criminal intelligence data from across the United States at all levels of government. One product of this summit was the National Criminal Intelligence Sharing Plan (“Plan”), endorsed by then U.S. Attorney General John Ashcroft. The Plan addresses two goals: (1) create a coordinating council comprised of local, state, tribal, and federal law enforcement executives; and (2) address the legal impediments to transferring criminal intelligence between law enforcement agencies. The purpose of the Plan was explained as follows by Attorney General Ashcroft:

This Plan represents law enforcement’s commitment to take it upon itself to ensure that the dots are connected, be it in crime or terrorism. The Plan is the outcome of an unprecedented effort by law enforcement agencies, with the strong support of the Department of Justice, to strengthen the nation’s security through better intelligence analysis and sharing (U.S. Department of Justice 2005).

Barriers to a national intelligence plan include the absence of a coordinated process for generating intelligence; a tradition of resistance to cooperation between local, state, tribal, and federal law enforcement organization; deficits in intelligence analysis; and differing technologies used in different agencies. To break down these barriers and achieve summit goals, the participants recommended:

- Promoting intelligence-led policing
- Protecting civil rights
- Building trust among law enforcement agencies
- Remediating analytic deficits
- Remediating information deficits
- Addressing training issues
- Addressing technology issues

¹² For more information, see *Recommendations from the IACP Intelligence Summit, Criminal Intelligence Sharing: A National Plan for Intelligence-Led Policing at the Local, State, and Federal Levels*, IACP, 2002. Available at <http://www.theiacp.org/documents/pdfs/Publications/intelsharingreport.pdf>.

A second product of the summit was formation of the GIWG, created to provide specific recommendations for developing the National Criminal Intelligence Sharing Plan. The GIWG operates under the auspices of the U.S. Department of Justice's Office of Justice Programs. Included among its recommendations, and key to the FIAT course, is its plan for a national model of intelligence training. In developing this model, the GIWG had three main goals: (1) identify specific training topics for each level of personnel (police executives, managers, general law enforcement, intelligence officers, and intelligence analysts) involved in intelligence gathering; (2) make specific recommendations on both the objectives and delivery of training; and (3) develop the model curricula through collaboration with relevant agencies and groups (GIWG, 2003). In its efforts to meet these goals, the GIWG reviewed local, state, national, and international training curricula, only to conclude that there was a lack of national-level training standards and that no single national agency was coordinating intelligence training. In summary, the GIWG developed core standards to serve as a blueprint for developing the key knowledge necessary to achieve intelligence-led policing.

Core Standards

Increasingly, intelligence analysis is being seen throughout the law enforcement community as an important mechanism for building a foundation for criminal investigations and prosecutions. According to BJA (2005), "Analysis is an integral part of every major investigation an agency opens. Often, small pieces of information that may appear insignificant can be a major part of a larger picture." Law enforcement is supported by intelligence analysts who not only help solve crimes and increase the ability to prosecute cases, but also identify crime trends and develop threat, vulnerability, and risk assessments. The increased recognition of the importance of competent intelligence analysis has resulted in calls for adequate training of analysts. According to research conducted by the GIWG and reported by the Criminal Intelligence Training Coordination Strategy Working Group (2004), "...the law enforcement community cited the lack of sufficient training for personnel as a significant impediment to enhancing their intelligence function" (p. 1).

To support law enforcement, analysts must provide accurate, timely, and relevant information. This can entail fostering relationships with other law enforcement personnel and national and local analytical staff. More importantly, it requires core minimum

training standards. The GIWG, through the Plan, established core minimum training standards in six areas: law enforcement officers, law enforcement executives, intelligence commanders and supervisors, intelligence officers and collectors, and intelligence analysts. These training standards address the intelligence needs, role and mission, core training objectives, and recommended training length and delivery for all levels of law enforcement personnel involved in intelligence gathering, sharing, management, and operations. The GIWG recommends:

- Two-hour training for *law enforcement officers* emphasizing that they are the largest and most viable resource for collection of intelligence information.
- Four-hour training for *law enforcement executives* focusing on the management and regulation of intelligence gathering and dissemination.
- Twenty-four hour training for *intelligence commanders and supervisors* highlighting the daily intelligence functions within the agency.
- Forty-hour training for *intelligence officers and collectors* stressing their role in collecting, evaluating, and compiling intelligence information.
- Minimum of 40 hours of training for *intelligence analysts* emphasizing their job of providing actionable intelligence through critical thinking, logic skills, and research and analysis of raw data.
- Forty-hours plus of training for a *train-the-trainer* program teaching people how to deliver the different intelligence courses, from the two-hour training for law enforcement to the forty-hour training for intelligence analysts.

The FIAT curriculum provides both the recommended 40 hours of training for intelligence analysts and a separate instructor development component that serves as a train-the-trainer program.

Foundations of Intelligence Analysis Training

The *Foundations of Intelligence Analysis Training* (FIAT) course was developed by the NW3C in conjunction with the International Association of Law Enforcement Intelligence Analysts (IALEIA), Law Enforcement Intelligence Unit (LEIU), and the Regional Information Sharing Systems (RISS). This consortium formed a partnership to develop the FIAT course after recognizing a shortage of intelligence analysis training and a lack of standardized courses in this field. The purpose of FIAT is to provide a standardized, basic analytical intelligence training curriculum for law enforcement and regulatory personnel with the goal of being the standard-bearer course in intelligence analysis training.

The NW3C, IALEIA, LEIU, and RISS consortium developed the FIAT course to fill the void of affordable, quality analytic training programs. Before FIAT, intelligence training was either internal training provided for specific agencies or commercial training provided by individuals not currently working as intelligence analysts. The consortium, using the standards set forth by the GIWG, developed a national training curriculum that would be instrumental in standardizing training on intelligence analysis. Indeed, the FIAT course meets the GIWG intelligence training standards for intelligence analysts and is a 40-hour classroom course taught by practitioners with intelligence analysis experience.

Program Overview

The FIAT course targets law enforcement and regulatory personnel who have not received formal, basic intelligence analysis training. To develop the training, the NW3C assembled leaders in the intelligence analysis community to serve as subject matter experts (SMEs) for a two-day meeting in January 2003 (see Appendix 7-A for a listing of attendees). The purposes of the SME meeting were to (1) review and amend a proposed topic outline compiled by NW3C staff and (2) create training objectives for each training module. During the meeting, the associate director of NW3C proposed that they form a partnership with IALEIA, LEIU, and RISS to develop the course and to create a pool of potential training instructors.¹³

Through the cooperative effort of the SMEs and FIAT partnership organizations, the course curriculum was developed by July 2003. The five-day, 40-hour curriculum is divided into three sections: Introduction to Intelligence Analysis, Intelligence Analysis as a Thought Process, and Analysis Methods and Skills. The curriculum covers the following modules:

- Introduction to Intelligence Analysis
 - History of Intelligence Analysis
 - Purpose of Intelligence Analysis
 - Intelligence Models and Core Competencies
 - Intelligence Cycle
 - Legal Issues and Ethics
 - Resources

¹³ NW3C, IALEIA, and LEIU have a history of cooperative initiatives. For example they teamed up to produce the CD, *Turnkey Intelligence: Unlocking Your Agency's Intelligence Capabilities*, which was distributed at the 2002 International Association of Chiefs of Police conference.

- Intelligence Analysis as a Thought Process
 - Critical Thinking
 - Creative Thinking: Brainstorming and Mind Mapping
 - Fundamentals of Logic
 - Inference Development: Competing Hypotheses, Assigning Probability
 - Recommendations & Development
- Analysis Methods & Skills
 - Crime Pattern Analysis
 - Association Analysis
 - Flow Analysis
 - Communication Analysis
 - Financial Analysis
 - Strategic Analysis
 - Indicator Development
 - Products of Intelligence: Reports and Presentations

In September 2003, key SMEs and other interested individuals were invited to the NW3C offices in Fairmont, West Virginia, for a three-day instructor development training. The purpose of this training was twofold. First, the FIAT curriculum and exercises were pilot-tested for thoroughness, consistency, errors, and logic. Second, the NW3C partnership wanted to ensure that every individual who taught the FIAT curriculum had adequate training on both the curriculum materials and teaching strategies. Instructor development training participants were given the opportunity to present a segment of the FIAT curriculum and were given feedback on their performance. Further, course participants provided direct commentary on the curricula, highlighting mistakes, errors, and inconsistencies. Upon receiving feedback from the pilot test and instructor training, the NW3C staff modified the FIAT curriculum in time for the first class, which was offered December 1-5, 2003, in Hershey, Pennsylvania.

Evaluation Methodology

The purpose of the FIAT evaluation was to assess what effect the FIAT course had on training participants in terms of their learning, behavior, and work. In particular, this research sought to determine how the participants felt about the training, what new knowledge and skills they may have acquired, and how the training may have changed their work activities.

Evaluation Questions

As discussed earlier in this report, the evaluation team modified and expanded the Kirkpatrick (1998) training evaluation model (see Chapter 4). To summarize, we used the following four levels of training evaluation: Level 1: Reaction—measures what the participants felt about the training; Level 2: Knowledge—measures what knowledge the participants gained from the experience; Level 3: Behavior change—measures how the participants changed work activities as a result of attending the training; and Level 4: Organizational impact—assesses the impacts of the training on the trainees’ organizations. Using this model as a framework, this project examined the following evaluation questions:

- How did the training participants react to the training? What was their assessment of FIAT in terms of meeting course objectives, delivering clear and logical concepts, and providing participants the opportunity to contribute and to network? What was their overall opinion of the instructors and the course? Would they recommend the course to others?
- What did the training participants gain in terms of information and skills?
- Have the training participants experienced any behavior changes due to the training? What effect has the training had on the daily work of the training participants?

Exhibit 7-1 illustrates the FIAT evaluation questions and data collection tools grouped by each evaluation level.

Exhibit 7-1: Evaluation Questions and Data Collection Tools Grouped by Evaluation Level

Evaluation Level	Evaluation Questions	Data Collection Tool
I: Reaction	How did the participants react to the training? Were they satisfied?	Student course evaluation
II: Learning	What information and skills were gained?	Pre/post knowledge test
III: Behavior	How have participants’ work behaviors changed due to the training? What effect has the training had on the daily work of the training participants?	Pre/post training self-assessment on job techniques Follow-up participant survey Follow-up supervisor survey

The evaluation team was unable to collect Level 4 data in the FIAT evaluation. We even had difficulties with our preferred version of a Level 3 evaluation. The original evaluation plan called for mixed within- and between-subjects quasi-experimental design with two comparison groups. One comparison group was to be matched with FIAT participants on place of employment, months or years of experience, and prior training experience. The second comparison group was to be drawn from participants who had taken a different foundational intelligence course.

The matched control group was not possible for logistical reasons. Often, participant registration for the FIAT course was last minute. Because of this, there was not enough time to develop matched controls and implement evaluation pre-tests while FIAT attendees were away at training. The second comparison group did not work out because the provider organization of the other intelligence training failed to cooperate. The research team was not able to find another comparable group (delivering similar basic intelligence training) during the data collection phase of the evaluation. Thus, a pre/post test within-subjects design with follow-up was used. While limited by not having a comparison group, the abundance of data collected pre-training, post-training, and six months or more post-training still provided multiple measures of training effectiveness.

Data Collection Methods and Framework

This section outlines the evaluation plan and data collection tools used in evaluating the FIAT training program. The discussion focuses on the participants, the evaluation design, and the data collection framework.

Evaluation Plan

Participants and Training Sites

The purpose of the FIAT course is to provide basic, entry-level training to intelligence analysts in the early stages of their careers. A total of 136 participants for this study were drawn from seven FIAT trainings between February and August 2004. Fifty-five percent (n=71) of the participants were new analysts, 20.2 percent (n=26) had less than one year's experience, and 24.8 percent (n=32) had more than one year's experience.

The training locations were Ft. Lauderdale, Florida; Boston, Massachusetts; Sacramento, California; Atlanta, Georgia; Denver, Colorado; Richmond, Virginia; and Salem, Oregon (see Exhibit 7-2). Data were collected, but not used, for three additional trainings. Data for the initial FIAT classes in December 2003, in Hershey, Pennsylvania, and Phoenix, Arizona, had to be eliminated because the training curriculum and exercises were significantly revamped after these trainings. These two trainings were the first FIAT classes offered; understandably, they served as additional "pilot tests" for developing the curriculum, perfecting delivery for the instructors, and implementing data collection. In addition, data for training in Springfield, Missouri (from March 29-April 5, 2004) could not be used because participants did not receive a complete cycle of data collection during the training. Training staff inadvertently failed to photocopy a key piece of the knowledge pre-test prior to the training; consequently, the evaluation team did not use data from that training.

Exhibit 7-2: FIAT Classes Included in Evaluation

<i>Date</i>	<i>Location</i>	<i>Number of Participants</i>
February 16-20, 2004	Ft. Lauderdale, Florida	17
March 1-5, 2004	Boston, Massachusetts	24
April 5-9, 2004	Sacramento, California	19
May 17-21, 2004	Atlanta, Georgia	14
June 21-25, 2004	Denver, Colorado	27
July 12-16, 2004	Richmond, Virginia	22
August 2-6, 2004	Salem, Oregon	13

Design

The design for the FIAT evaluation is a pre/post test within-subjects design with follow-up conducted over a 19-month period (see Exhibit 7-3). The pre-tests involved a knowledge test on the training materials and a self-assessment of comfort level in working in one of six main intelligence data areas. These were given to each participant at the beginning of the training. The post-test involved two phases. The first phase consisted of a student course evaluation, a second knowledge test, and self-assessment given to each participant at the immediate conclusion of the training. The second phase consisted of follow-up surveys with training participants and their supervisors about six months subsequent to the training.

Exhibit 7-3: Pre/Post Test Within-subjects Design of the FIAT Training

Pre-test	Training	Post-test	2nd Post-test
O _a O _b		O _c O _d O _e	O _f O _g
T ₁	T ₂	T ₃	T ₄

Where:

O_a = Pre-knowledge test

O_b = Pre-self assessment on job techniques

O_c = Post-knowledge test

O_d = Post-self assessment

O_e = Student course evaluation

O_f = Follow-up survey with training participants

O_g = Follow-up survey with supervisors of participants

Data Collection Framework

As noted earlier, the framework for data collection used in this evaluation was based on Kirkpatrick's (1998) model for evaluating training programs—specifically, the successive levels of (1) reaction, (2) learning, and (3) behavior. Data collection tools involved seven sources of data: pre-training knowledge tests, pre-training self-assessments on job techniques, student course evaluations, post-training knowledge test, post-training self-assessment on job techniques, follow-up survey with participants, and follow-up survey with supervisors.

Level 1: Reaction

The key evaluation questions for the Level 1 evaluation of FIAT were: How did the participants react to the training? How satisfied were they? Answering this involved student course evaluation surveys focusing on the participants' reaction to the training. The survey instrument was developed by NW3C staff and is standard in all NW3C trainings (see Appendix 7-B for an example of the survey).

Level 2: Learning

The key evaluation question for the Level 2 evaluation of FIAT was: What information and skills did the participants gain? Answering this involved conducting a pre and post-test survey focusing on the knowledge obtained during the training. This survey was administered

immediately prior to and immediately following the training. The knowledge test was developed by NW3C staff (see Appendix 7-B for an example of the test).

Level 3: Behavior

The key evaluation question for the Level 3 evaluation of FIAT was: How have participants' work behaviors changed due to the training? Answering this involved administering a pre and post-training self-assessment on the main subject areas of the training, and follow-up surveys with training participants and their supervisors. NW3C staff developed the self-assessment tests, and the evaluation team developed the follow-up survey instruments (see Appendix 7-B for examples of the follow-up survey instruments for participants and supervisors).

Study Strengths and Weaknesses

A key strength of the evaluation design was the extensive cooperation of the NW3C staff, trainers, and training participants. The NW3C staff was open and responsive to this evaluation, developing data collection instruments as needed, and allowing the evaluation team the freedom to observe and collect whatever data was possible. An additional strength included a design that allowed for changes in behavior and knowledge to be measured. The pre/post design with a follow-up at six months post-training, allowed the evaluation team to assess what the attendees learned from the training and how they implemented this learning on the job.

The main weakness with this evaluation was the inability to secure either of the two comparison groups originally proposed in the evaluation plan. The pre/post design provided an adequate, but not ideal, alternative because it assumes that changes in knowledge and behavior can be attributed to the training. It may be that participants gained these skills from additional sources, such as colleagues, supervisors, or additional training.

Another limitation was that the follow-up surveys had relatively low response rates, thus calling into question whether those who did not respond were somehow different than those who did.

Evaluation Findings

Central to the FIAT evaluation design was assessing what participants learned at the training and how they used the new knowledge and skills in their work. As discussed earlier, the

NW3C and partner agencies (IALEIA, LEIU, and RISS) created FIAT to provide affordable, quality analytic training in a market that was lacking both. The overarching goal was to make FIAT a standard bearer intelligence course, meeting or exceeding training standards set by the National Criminal Intelligence Sharing Plan developed by the GIWG. FIAT evaluation findings are based on the seven sources of data shown in Exhibit 7-3:

- pre and post self-assessments of job techniques
- pre and post knowledge tests
- student course evaluations
- follow-up surveys with course participants
- follow-up surveys with supervisors of course participants.

Participant Reaction

In assessing participant reaction (Level 1 of the training evaluation), the central questions were: What were the participants' assessments of the training? What did they think of the training course content, instructors, opportunities for class participation, and networking with other analysts? Findings were derived from the student course evaluations.

Data collection for student course evaluations relied upon NW3C's own procedures for collecting student feedback on trainings. NW3C routinely administers student course evaluations. The student course evaluation for FIAT was developed by NW3C and administered at the end of the training by NW3C staff. The survey questions dealt with a range of issues, including course presentation, participant interaction, and instructor ratings. NW3C staff collected the surveys from the training participants and sent hard copies to the evaluation team.

Reaction to Training

The FIAT training uses a combination of lecture, question and answer sessions, PowerPoint presentations, and individual and group work exercises to teach participants how to conduct intelligence analysis. Of the 136 FIAT participants during the evaluation period, 131 completed student course evaluations. Overall evaluation of FIAT was high; 65.4 percent of attendees thought the course was "excellent," 32.3 percent thought it was "good," and 2.4 percent thought it was "average." When asked if they would recommend the course to others, nearly 98.0 percent said "yes." Specific feedback was also very positive, and included the following comments:

- “Great for beginning analysts.”
- “It gives a basic view of understanding to entry level analysts. Very organized and well done!”
- “Excellent course on analysis training.”
- “I thought it was an excellent course on the foundations of analysis. It held my attention.”
- “Great foundations course for beginning analysts.”
- “Provides essential fundamentals of analysis.”
- “A good beginner course. Would like follow-up classes on specific topics covered in this course.”
- “An excellent introduction to analysis and/or brush up on skills or areas needing improvement.”
- “It was very informative.”
- “Very hands on and useful. Practical information to use immediately. Great basic tools.
- “It’s a great foundation to build on.”
- “For beginners – very basic, good information.”
- “Good information, especially for a new analyst with no experience.”

Participants were asked to evaluate FIAT in three additional areas: course presentation, interaction, and instructors. In general, participants gave FIAT high marks in all three categories (see Exhibit 7-4). The course presentation category included: “Identified course objectives were met,” “Concepts were clear,” and “Course content was organized logically.” Of the 131 responses, the vast majority of participants either strongly agreed or agreed, and no participants disagreed or strongly disagreed, with all three of these statements. The interaction category included: “I felt comfortable asking questions,” “I was given the opportunity to contribute during class,” and “I had opportunities to network with other participants.” The majority of participants either strongly agreed or agreed, while a small minority of participants were either neutral or disagreed. No participants strongly disagreed.

Exhibit 7-4: Participant Reaction to FIAT Training

Student Course Evaluation Question	n	Assessment				
		Strongly Agree %	Agree %	Neutral %	Disagree %	Strongly Disagree %
Identified course objectives were met	131	61.1	37.4	1.5	0	0
Concepts were clear	130	53.8	43.1	3.1	0	0
Course content was organized logically	130	62.3	35.4	2.3	0	0
I felt comfortable asking questions	130	69.2	26.9	3.1	0.8	0
I was given the opportunity to contribute during class	130	76.2	20.0	3.1	0.8	0
I had opportunities to network with other participants	131	80.9	17.6	0.8	0.8	0

Participant feedback on the course and its content was very positive. Comments included:

- “This is the first class where I got to participate in exercises and get real hands on experience.”
- “I thought this course was great! I think one of the most valuable tools was the interactive group activities. Prior to the course I had very little knowledge regarding charting. I think what I have learned will prove to be an asset throughout my career as an analyst. Excellent!”
- “Definitely helps break down complicated cases into workable form.”
- “Very good exercises. Thank you!”
- “Gives actual knowledge/working knowledge and experience to otherwise empty examples.”
- “Really liked the development/use of cases to reinforce things. Very glad you brought this to Denver for no cost – I would never get this type of training otherwise! This course will help me with all kinds of investigations involving the homeland defense stuff I do.”
- “The hands on applicability of the exercises solidified the lessons superbly.”
- “Excellent presentation of background and fundamentals of work area.”

Reaction to Trainers

The FIAT training used a combination of NW3C staff to instruct individual modules and outside trainers who were experienced practitioners in intelligence analysis and had completed the FIAT instructor development program. The FIAT courses during the evaluation period were taught by 16 different instructors, with the majority teaching only once. On average, two or three instructors shared teaching duties during the week-long course.

In general, the instructors were highly regarded. The vast majority of participants rated the instructors as either “good” or “excellent.” Exhibit 7-5 shows the top seven ratings by instructor. No participants rated the instructors as “poor,” and few found them to be “average” or “fair.” Comments about the instructors included:

- “All instructors were very professional and provided a lot of experience history and it came through in their presentations. I was very impressed with the layout and content of the class, most people would classify as “boring” information – the instructors made it not boring.”
- “All instructors did a really good job. Keep up the good work.”
- “Superb instructors!”
- “Both instructors showed to have a good working relationship and used that asset to get a good learning environment.”
- “All of the instructors... really made it very hands on and easy to read all topics [from the] beginning all the way to the end of the class.”

Exhibit 7-5: Participant Reaction to FIAT Instructors

Overall Effectiveness	n	Assessment				
		Excellent %	Good %	Average %	Fair %	Poor %
Instructor 1	130	58.5	36.2	5.4	0.0	0.0
Instructor 2	130	49.2	31.5	12.3	6.9	0.0
Instructor 3	98	59.2	35.7	5.1	0.0	0.0
Instructor 4	32	81.3	15.6	3.1	0.0	0.0
Instructor 5	19	68.4	26.3	5.3	0.0	0.0
Instructor 6	19	73.7	26.3	0.0	0.0	0.0
Instructor 7	19	94.7	5.3	0.0	0.0	0.0

Knowledge and Skills Gained

At the second level of evaluating the FIAT training—an assessment of learning—the central question was: What information and skills were gained? Findings were derived from the intelligence analysis knowledge test that participants took at the beginning and conclusion of the training. While measuring participants’ reactions (Level 1) is important for assessing their “buy-in” to a course, the Level 2 measures provide a useful assessment of what the participants learned from the course. The pre/post design provided a baseline measure of participants’ intelligence analysis knowledge from which improvements could be determined.

Each participant was given a knowledge pre-test at the beginning of the training and the same test at the conclusion of the training five days later. NW3C staff administered both the pre and post-tests. The purpose of the knowledge test was to assess analyst skills learned as a result of the FIAT training. At ILJ's request, the test was developed specifically for this evaluation by the curriculum developers at NW3C responsible for creating the FIAT course. Initially, the test consisted of 25 multiple choice questions; however, after the Sacramento training, the test was modified to 20 questions because the course developers felt that “...a few of the questions were still more confusing than evaluative” (personal communication, May 11, 2004). The questions were based on modules taught in the FIAT course and included questions such as:

1. In a link chart, circles show _____, while rectangles show _____.
 - a. locations ... events
 - b. people ... organizations
 - c. strong associations ... weak associations
 - d. activity ... results
 - e. criminals ... law enforcement

2. Logic that moves from the specific to the general is
 - a. inductive
 - b. deductive
 - c. faulty
 - d. imperative
 - e. specific

3. If your agency receives federal funds for intelligence analysis and storage, information on religious groups cannot be gathered and stored unless directly connected to a criminal investigation because of
 - a. FAACP
 - b. 28 CFR part 23
 - c. FFIA
 - d. all of the above
 - e. none of the above

Of the 136 FIAT attendees, 135 completed pre-tests, and 129 completed post-tests. The tests were matched and analyses were conducted on the matched pre/post tests.

The results of the knowledge test showed that FIAT participants improved in their knowledge and analysis techniques as a result of the course. Statistical analyses indicate a significant improvement in the test scores of participants, measured as percent correct,¹⁴ between the post-test ($M=79.87$, $SD=14.15$) and the pre-test ($M=68.59$, $SD=12.22$), $t(126)=9.728$, $p<.05$. Separate analyses—conducted on participants who were new to intelligence analysis, had less than one year experience, and had more than one year experience—indicated that each of these groups improved in their knowledge and skills as a result of the FIAT training (see Exhibit 7-6).

Exhibit 7-6: Matched Pairs T-Test Results of Intelligence Analysts Who Were New, Less-Than-1-Year Experience, and More-Than-1-Year Experience

Analyst Experience	Pre-knowledge Test % Correct		Post-knowledge Test % Correct		<i>df</i>	<i>t</i>
	<i>Mean</i>	<i>SD</i>	<i>Mean</i>	<i>SD</i>		
New	64.68	11.25	78.58	9.48	58	11.526**
Less than 1 year	70.83	10.26	80.75	15.60	23	3.03*
More than 1 year	75.25	10.68	82.75	14.71	27	3.370*

* $p<.05$
 ** $p<.001$

¹⁴ Percent correct rather than number correct was used because NW3C staff altered the pre/post knowledge test midway through the evaluation, decreasing the number of questions from 25 to 20.

The results indicated that all training attendees learned more about intelligence analysis from the FIAT course, regardless of level of experience. New analysts showed the most improvement, raising their scores by more than 14 percent from the pre to the post-test. Both the analysts with less than one year experience and those with more than one year experience showed improvements in their test scores as well. Analysts with less than a year's experience improved nearly 10 percent and those with the most experience improved more than 6 percent.

Participant feedback about the knowledge and skills gained was equally positive.

Comments included:

- “The practical exercises were great. I feel comfortable applying knowledge I have learned on practical exercises.”
- “I liked that the class was hands on. Most trainings I’ve been to have been solely informational, not practical exercises. Also liked the way the class was broken up – with some lecture followed by exercises. It also makes interaction [possible] with people from other agencies.”
- “Good basis of understanding on financial crimes.”
- “The association links will definitely help me in my work in financial crime.”
- “Overall it was an excellent course. A great way to strengthen your thinking process when trying to associate occurrences, people, etc. and hopefully put together facts needed to assist in closing a case. Thank you.”

Behavior Changes

The third level of evaluating the FIAT training programs involved assessing behavior change. The central question was: How have the participants' work behaviors changed as a result of the training? Findings were derived from self-assessments of job techniques completed by training participants before and at the completion of training; a follow-up survey sent out to participants about six months post-training; and a survey completed by participants' supervisors about six months after the training.

Self-assessments of Job Techniques

The purpose of this assessment was to gauge comfort level with the six main types of intelligence analysis commonly performed by analysts in the FIAT course. The topics included crime pattern, associations, flow, communications, financial, and strategic analysis.

Participants were asked to specify their level of experience in the intelligence analysis field, and then to rate their comfort level with each topic using a scale from 0 (avoid!) to 3

(completely comfortable). The NW3C curriculum development team developed the self-assessment specifically for this evaluation. Of the 136 participants, 106 were matched on pre-post training comfort level with crime pattern, flow, and communication analysis; 107 participants were matched on association and strategic analysis; and 108 participants were matched on financial analysis. Data for the other participants were either missing or incomplete.

Before the start of FIAT, the majority of participants felt either “fairly comfortable,” or “not very comfortable” with each of the six intelligence analysis subject areas. A minority of participants were “completely comfortable,” or wanted to “avoid” the subject area altogether (see Exhibit 7-7). The same assessment given at the conclusion of the training indicated that no participants wanted to avoid the subject area, and most felt either “completely comfortable” or “fairly comfortable” with the subject matter (see Appendix 7-C for a table of scores).

Exhibit 7-7: Matched Pairs T-test Results of Pre/Post FIAT Course Comfort Level with Intelligence Analysis Subject Matter

Subject Area	Pre-training Self-Assessment		Post-training Self-Assessment		df	t
	Mean	SD	Mean	SD		
Crime Pattern Analysis	1.66	.755	2.34	.55	105	9.654**
Association Analysis	1.74	.769	2.55	.570	106	10.974**
Flow Analysis	1.58	.754	2.49	.573	105	11.668**
Communication Analysis	1.53	.771	2.51	.573	105	11.296**
Financial Analysis	1.65	1.65	2.20	.592	107	17.427**
Strategic Analysis	1.43	.741	2.13	.568	106	8.653**

**p<.001

Statistical analyses indicate a significant improvement in the pre/post self-assessment ratings of participants in each of the six subject areas. Separate analyses conducted on participants who were new to intelligence analysis, had less than one year experience, and had more than one year experience, indicate that each of these groups also felt more comfortable with

the subject matter in each of the six areas as a result of the FIAT training (see Appendix 7-D for a table of t-test scores by level of experience and subject matter). The FIAT training improved both the knowledge and skills of participants and their comfort level with the material. This held true regardless of how much experience the participants had in the intelligence analysis field prior to the training.

Follow-up Surveys with Participants and Supervisors

Additional data sources consisted of follow-up surveys with the training participants and their supervisors. These surveys were conducted by the evaluation team. In this level, the evaluation purpose of both the participant and the supervisor surveys was to determine if the participants retained their new knowledge and skills in the six key intelligence analysis areas; and included questions about promotions and career advancements, what attendees were doing differently as a result of the training, whether the training met their expectations, and the most and least significant thing learned during the training. The supervisor survey mainly addressed the issue of new knowledge and skills and how the training affected the training attendee's work. The surveys were personalized for each supervisor.

NW3C staff provided the evaluation team with participant contact information. From this list, the evaluation team sent surveys to training participants by email, fax, and through the mail. Follow-up emails were sent and telephone calls made to each non-responsive participant. These non-responsive participants were contacted at least five times in an effort to obtain a response to the follow-up survey. To provide participants with the maximum time to assess knowledge acquisition and behavioral changes, participant surveys were sent out nine to 12 months post-training.

Administration of the participant surveys was compromised by an internal error with the ILJ computer server. An unknown number of surveys were returned via email to the evaluation team but were lost due to ILJ server changes happening concurrently with survey administration. The evaluation team was only made aware of these lost surveys upon telephone follow-up with participants weeks after the initial survey administration. Some participants were able to resend their survey, others were not (or were not willing).

Of the 136 FIAT participants, 59 returned the survey, for a participation rate of 43.3 percent. Most of the responding FIAT participants (89.8 percent) held the same job nearly one

year after attending the training. Of the participants who had been promoted (16.9 percent) since the training, 60 percent believed it was due, at least in part, to their FIAT training. In addition, nearly a year after attending the FIAT course, the vast majority of respondents believed they had gained both knowledge and skills in the six subject areas on which the course had focused (see Exhibit 7-8).

Exhibit 7-8: Did Training Participant Respondents Gain New Knowledge or New Skills in the Six Intelligence Analysis Subject Areas 9-12 Months After the Training?

Subject Area	Knowledge		Skills	
	Yes %	No %	Yes %	No %
Crime Pattern Analysis	88.1	11.9	89.5	10.5
Association Analysis	88.1	11.9	86.0	14.0
Flow Analysis	91.5	8.5	87.9	12.1
Communication Analysis	89.8	10.2	87.7	12.3
Financial Analysis	86.4	13.6	82.5	17.5
Strategic Analysis	89.8	10.2	86.2	13.8

When asked to describe what they learned that was new and how they were using this knowledge, training participant respondents stated:

- “Being new to the intelligence field, attending the course gave me insights on how to correlate data and provide products to our customers.”
- “Much of the history was new to me. Also, all of the flow and financial analysis materials were new and very useful.”
- “I have used the association analysis in a recent case and it worked well for court.”
- “[I] learned different ways to “link” crimes thru graphical analysis.”
- “I learned how to develop and present complex cases with large volumes of information and how this relates to people involved at varying levels, so that others can more easily understand the overall picture.”

- “One of the biggest things I learned was not to ignore any piece of information—that even a tiny clue can lead to further clues that can lead to much needed information.”
- “How to effectively construct an initial blueprint of all analysis to be completed during an investigation and being able to use each procedure in conjunction with the others.”
- “The history of intel analysis gave a good intro as to why we do analysis and how it benefits law enforcement and society.”
- “Recently put together a crime pattern analysis for Commissioner’s Office, and ultimately made LEO response recommendations regarding this crime pattern.”

When asked what, if anything, training participant respondents were doing differently at their job because of the FIAT training, they stated:

- “The investigators rely upon me more frequently for information.”
- “I’m practicing skills... on small cases at this time.”
- “I am more diligent in collection of intelligence-related material.”
- “[I] have created a timeline for a[n] investigation.”
- “I prepare my case presentation based upon much more of what I learned at this course. It has made these reports far easier to explain for me and to understand for those receiving them.”
- “Viewing my cases through logical and critical thinking. Creating various flowcharts.
- “My analysis of the information now feels more complete and I am more confident with my suggestions to our investigators.”
- “To dig a little deeper and leave no stone unturned.”
- “Have made and used chart to assist in investigation.”
- “My charts are better focused and more easily understood.”
- “I’ve begun to consider additional sources of info and additional types of analysis to be applied to my cases.”
- “I have used link charts for presenting to grand juries and net worth analysis in determining ability to pay restitution.”

The majority of survey respondents shared their training notebooks with their co-workers (69.0 percent), referred to the notebooks since returning (74.6 percent), and found the notebooks useful in their work (86.5 percent). As a gauge of training value, participants were asked if, in retrospect, the training both met their expectations and if they would recommend the training to

colleagues. Nearly all the participants felt the training met their expectations (93.1 percent), and 89.5 percent said they would recommend the training.

The evaluation team also experienced some problems with the supervisor survey administration. Initially, at the request of the evaluation team, NW3C included supervisor contact information on the FIAT registration forms. However, this form was changed following the Boston training. Thus, supervisor contact information was incomplete for a majority of the participants. To obtain this information, the evaluation team requested it on the participants' follow-up survey. Because of the computer system problems (noted earlier) in sending the participant survey by email, the supervisor survey was sent via mail. These surveys were sent out 12 months following training. Of the 136 participants, the evaluation team had contact information for 59 supervisors. Of these 59 supervisors, 33 returned the surveys. Supervisors (like the participants) were contacted a minimum of five times by telephone and email.

Of the 33 supervisors who returned the survey, the vast majority felt that the employee who attended the training improved in both knowledge and skills in each of the six training subject areas (see Exhibit 7-9). When asked if they saw a positive change, negative change, or no impact on the work of training participants, 93.9 percent of supervisors reported a positive change, while 6.1 percent reported no impact. No supervisors reported a negative change in the work of the participants as a result of the training. Nearly all of the supervisors (93.5 percent) commented that they would send other staff to the training. This is especially noteworthy, since the 40-hour FIAT course keeps participants away from work for an entire week.

When asked how the training participants' work had changed as a result of FIAT, supervisor statements included:

- “[We] restructured the intelligence unit based on information gained by attending FIAT.”
- “She has greater confidence in her abilities.”
- “[Participant] has been able to graph and link investigatory issues for a clearer presentation to grand juries.”
- “Able to supervise the analysts with more understanding of their duties.”
- “Use of timeline analysis and increased and improved notetaking during interviews.”

- “He has been able to develop cases more expeditiously, and this has resulted in more efficient use of available time.”
- “He needs less guidance and direction.”

Exhibit 7-9: Does Supervisor Think Training Participant Gained New Knowledge or New Skills in the Six Intelligence Analysis Subject Areas 9-12 Months After the Training?

Subject Area	Knowledge		Skills	
	Yes %	No %	Yes %	No %
Crime Pattern Analysis	97.0	3.0	90.9	9.1
Association Analysis	90.9	9.1	90.9	9.1
Flow Analysis	90.9	9.1	93.9	6.1
Communication Analysis	84.8	15.2	87.5	12.5
Financial Analysis	90.9	9.1	90.9	9.1
Strategic Analysis	84.8	15.2	84.8	15.2

Discussion

The NW3C and partner agencies (IALEIA, LEIU, and RISS) answered the call by the GIWG to provide 40 hours of core training for intelligence analysts. They strived to be the leader in providing a model training course that sets the standard for intelligence analysis. This evaluation tested our expanded version of Kirkpatrick’s training evaluation model by evaluating the FIAT training. The evaluation sought to answer the questions of what participants thought of the training, what they learned, and how they used this information in their work. We believe this evaluation provides meaningful feedback to NIJ, NW3C, intelligence analysts, and the greater law enforcement community.

Strengths of the Course

Important measures of a training program’s success are whether participants (1) learned according to the training objectives and (2) used the new knowledge and skills in their day-to-day work. This evaluation collected seven sources of data, including pre and post-knowledge

tests and self-assessments, student course evaluations, and follow-up surveys with both participants and their supervisors approximately one year after the FIAT training. Findings indicate that all participants, from those new to the field to those with more than one year of experience, learned significantly more about intelligence analysis from the FIAT course. In addition, they became more comfortable with each of the six main subject areas, as measured in a pre/post training self-assessment. Follow-up surveys with both participants and supervisors indicated that most felt they had gained both knowledge and skills and were using their new found skills in their daily work.

Keys to successful training include designing a high quality curriculum and finding capable instructors. The NW3C staff and SME curriculum developers determined early on that they wanted FIAT to be a foundational course, providing basic training in intelligence analysis for new and inexperienced analysts. As with any profession, intelligence analysis has a set of skills that are unique and has its own history. The experts involved in creating FIAT wanted to ground the training in the history of intelligence analysis, while balancing the need to develop competency through a combination of lecture, group work, and exercises.

NW3C used a transdisciplinary, problem-based learning approach in creating the FIAT curriculum. The FIAT curriculum transmits a set of six intelligence analytic skills (discussed earlier) to participants via lecture, interactive discussion, and a series of exercises that successively build upon each other. Thus, by the end of the 40-hour training, participants are able to analyze criminal cases with varying degrees of complexity.

Instructors are instrumental to the success or failure of training. Instructors not only control curriculum content, but their presentation style can either engage or turn off participants. To ensure competent, thorough, and consistent training, NW3C and the FIAT SMEs created a curriculum instruction guide that all instructors were required to use. Most instructors were also trained in an instructor develop program. All FIAT instructors were either employees of NW3C or experienced practitioners in the field of intelligence analysis.

Recommendations for Change

Many participants remarked that FIAT was an excellent beginner course and provided a great foundation. They also enjoyed the exercises and found the instructors to be very

knowledgeable. Even when asked directly for negative feedback on the course, few participants had a response. The few who did offered only mild criticisms, including:

- “[Spend less time on] the matrix—currently it is not that important, good to explain but not to spend a lot of time on it.”
- “Some of the presenters were difficult to listen to, therefore, my attention strayed.”
- “Crime pattern [analysis] should have been directed at more diverse crimes, fraud, financial.”
- “[Spend less time on] Financial analysis—I think we would submit it to NW3C to do.”

Participants did offer some recommendations on ways in which the training could be improved. For the most part, the recommendations addressed structural aspects of the training, such as spending less time on history and more time on group work. Specific recommendations included:

- “Make sure the presenters are interesting to listen to and that they have a passion for the topic they are presenting.”
- “More exercises. I enjoyed working with other agencies and trying to solve problems together.”
- “Dedicate more time to financial crime.”
- “Cut down a little on the morning material (i.e., history) because it’s a quick way to lose the audience.”
- “If anything, I wish the course had been more detailed and longer. When I went in I knew so little, I felt rushed.”
- “Maybe a little less lecture and more hands on. I found it more interesting and challenging to the mind to [do] actual work on some of the scenarios that were given in groups.”
- “More focus on practice exercises used for each of the courses.”
- “The instructor for the Code of Federal Regulations (28 CFR Part 23) did an excellent job. However, due to the importance of this subject additional time should be allotted in order to ensure that everyone has a clear understanding of the statutory requirements for an intelligence exchange.”
- “More time for exercises; eliminate sections that aren’t used as often in “real life” to make more time for using the tools like association analysis, flow, financial, communication analyses.”
- “Update the communications block of instruction.”

The FIAT training clearly fills a void. As the GIWG discovered, there is a lack of quality intelligence analysis training programs. In addition, few of the trainings meet the core recommendation of 40 hours, nor do they provide train-the-trainer opportunities as recommended by the GIWG. Efforts to meet the need for quality training programs are hampered by the lack of professional certification or set industry standards covering the work of intelligence analysts. NW3C and its partner organizations would like to see this change. One participant stated that they would recommend FIAT to co-workers because it would "...get everyone 'on the same page' about where we should be in our field, what resources we have, how we should be doing our jobs (i.e., steps we take in different types of analyses)."

Additional recommendations by the evaluation team include:

- NW3C and its partner organizations should strive to maintain consistent and high-quality instructors for FIAT. Instructors with poor ratings should be dropped.
- Instructors should be paid for their time. The old adage, "You get what you pay for," means quality but also commitment. When you pay a subcontractor, you get a commitment. When that person is working for free, this is not always the case.
- The FIAT training should continuously be improved through evaluation. A cycle of instruction, evaluation, and adjustment to curriculum, instructors, or methods is vital for keeping training fresh, current, and successful. ILJ's enhanced Kirkpatrick training evaluation model provides a good model for training providers to use to evaluate their courses.
- NW3C might also strive to conduct a Level 4 training evaluation to show a return on investment. Additional questions would include measures of cost benefit or cost effectiveness of the training. For example, can participants "buy" the same training elsewhere for less? Answering these important questions would require use of comparison groups and carefully planned pre and post-training cost benefit measures.

APPENDIX 7-A

FIAT Development SME Participants

Don Brackman	National White Collar Crime Center
Dick Darnell	RISS/Mid-States Organized Crime Information Center
Lloyd Hoffman	Retired—Defense Intelligence Agency
Robert Leonard	CIA Sherman Kent Instructor
Robert Lookabill	Defense Intelligence Agency
Gary Lusher	National White Collar Crime Center
Kristie Manzie	Florida Department of Insurance Fraud
Ritchie Martinez	Arizona DPS/IALEIA President
Bob Morehouse	California DOJ/LEIU
Marilyn Peterson	New Jersey Office of Counter-Terrorism
Joe Regali	RISS/New England State Police Information Network
Gary Stoops	National White Collar Crime Center
Danny Taylor	National White Collar Crime Center
Greg Thomas	Pennsylvania State Police

APPENDIX 7-B

NW3C FIAT Course Training Evaluation Materials

Participant Reaction Survey

Knowledge Test

Follow-up Surveys: Participants and Supervisors

Foundations of Intelligence Analysis Training Evaluation

COURSE PRESENTATION

1) Overall course objective was met:

- A) Strongly Agree
- B) Agree
- C) Neutral
- D) Disagree
- E) Strongly Disagree

2) Concepts were clear:

- A) Strongly Agree
- B) Agree
- C) Neutral
- D) Disagree
- E) Strongly Disagree

3) Course sequence was logical:

- A) Strongly Agree
- B) Agree
- C) Neutral
- D) Disagree
- E) Strongly Disagree

INTERACTION

4) I felt comfortable asking questions:

- A) Strongly Agree
- B) Agree
- C) Neutral
- D) Disagree
- E) Strongly Disagree

5) I was given the opportunity to contribute during class:

- A) Strongly Agree
- B) Agree
- C) Neutral
- D) Disagree
- E) Strongly Disagree

6) I was provided with opportunities to network with other participants:

- A) Strongly Agree
- B) Agree
- C) Neutral
- D) Disagree 0
- E) Strongly Disagree 0

INSTRUCTORS

7) Please rate the overall effectiveness of the instructors:

Instructor [insert name]

- A) Excellent
- B) Good
- C) Average
- D) Fair
- E) Poor

8) Please note the strengths of each instructor:

Instructor [insert name]

- A) Spoke very clearly
 - B) Knowledge
 - C) Knowledge easily relayed to us
 - D) Very good knowledge of topics he was teaching
 - E) A lot of personal experience
-
- A) Good personality
 - B) Great at taking questions further into Law Enforcement aspects
 - C) Encouraged participation
 - D) Knowledge and communication skills were exceptional
-
- A) Dynamic and excellent speaker
 - B) Excellent pace and good presentations
 - C) Comfortable with self and subject matter
 - D) Excellent public speaker
 - E) Kept a constant energy flowing

11) Please Rate the Overall quality of the FIAT course:

- A) Excellent
- B) Good
- C) Average
- D) Fair
- E) Poor
- F) No Answer

12) Would you recommend this course to others:

- A) Yes
- B) No
- C) No Answer

Overall Suggestions and feedback:

Foundations of Intelligence Analysis Training (FIAT)

Test Your Knowledge

Circle the correct answer.

1. **The intelligence process includes planning, collection, collation, _____, analysis, dissemination and feedback.**

- a. storage
- b. evaluation
- c. retention
- d. computerization
- e. investigation

2. **The primary use/uses of law enforcement intelligence is/are to:**

- a. support decision-making
- b. prioritize and select cases
- c. investigate organized crime and terrorism
- d. a and b
- e. none of the above

3. **Crime patterns can be seen in times, locations, _____**

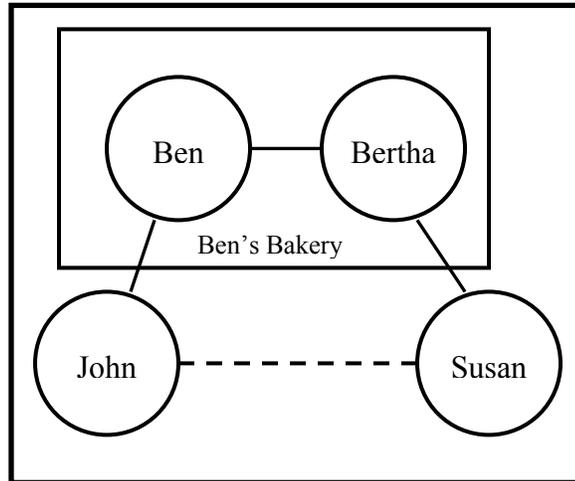
- a. weapons used and methods of operation
- b. names of perpetrators and names of victims
- c. vehicles used and numbers of perpetrators
- d. all of the above
- e. none of the above

4. **In a link chart, circles show _____ while rectangles show _____.**

- a. locations ... events
- b. people ... organizations
- c. strong associations ... weak associations
- d. activity ... results
- e. criminals law enforcement

5. **The link chart at the right tells you:**

- a. Susan is related to Bertha and Ben.
- b. John has a known association with Ben, a known association with Bertha, and a suspected association with Susan.
- c. John has a known association with Ben, no known association with Bertha, and a suspected association with Susan.
- d. Ben, Bertha and John are involved in a criminal activity.
- e. None of the above.



6. **Flow analysis types include:**

- a. event flow
- b. commodity flow
- c. visual investigative analysis
- d. all of the above
- e. none of the above

7. **Communication analysis looks at:**

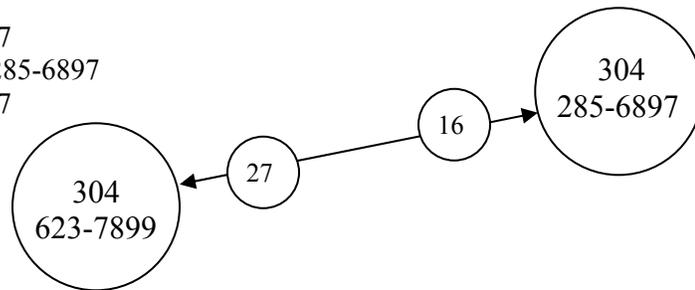
- a. what number is connected to what number
- b. what dates and times the connections were made
- c. what was said/written in the communications
- d. a and b
- e. all of the above

8. **Geographic distribution analysis allows us to:**

- a. plot locations of crimes
- b. look for patterns in the incidents
- c. look at distances between incidents
- d. a and c
- e. all of the above

9. **The chart at the right tells you:**

- a. 623-7899 made 27 calls to 285-6897
- b. 623-7899 talked for 27 minutes to 285-6897
- c. 623-7899 made 16 calls to 285-6897
- d. None of the above



10. **In bank record analysis, patterns can be seen in:**

- a. check amounts
- b. check dates
- c. who endorsed the checks
- d. a and b
- e. all of the above

11. **Which statement(s) below is/are true?**

- a. premises lead to inferences
- b. the inferences we want to test become our hypotheses
- c. as we analyze the information in light of our hypotheses we draw conclusions
- d. from our conclusions we make recommendations
- e. all of the above are true

12. **Shows the movement of stolen property, money, drugs or smuggled goods.**

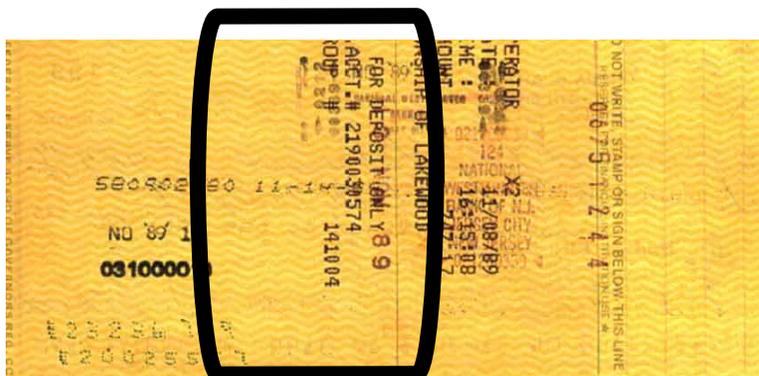
- a. a commodity flow chart
- b. a link chart
- c. an event flow chart
- d. a financial profile
- e. none of the above

13. **Thinking “outside-of-the-box” best describes**

- a. critical thinking
- b. creative thinking
- c. inductive thinking
- d. deductive thinking
- e. none of the above

14. **The information circled on the back of this check refers to.....**

- a. the bank that paid the check.
- b. the bank where the check was deposited.
- c. person who cashed the check.
- d. The business that wrote the check



15. **Strategic analysis products might include:**

- a. activity flow charts
- b. link charts
- c. threat assessments
- d. estimates / forecasts
- e. all of the above

16. **Recommendations generated from conclusions should take all but the following into account:**

- a. the underlying facts
- b. the conclusions/hypotheses drawn
- c. the previous opinion on the subject
- d. resources available to be used in the response
- e. b and c

17. **Indicators are developed from:**

- a. our projections
- b. past history
- c. academicians
- d. FBI documents
- e. recommendations

18. **If your agency receives federal funds for intelligence analysis and storage, information on religious groups cannot be gathered and stored unless directly connected to a criminal investigation because of**

- a. FAACP
- b. 28 CFR part 23
- c. FFIA
- d. all of the above
- e. none of the above

19. Logic that moves from the specific to the general is

- a. inductive
- b. deductive
- c. faulty
- d. imperative
- e. specific

20. A SAR is:

- a. System Analysis Report
- b. An investigative report
- c. A slang term for a white collar criminal
- d. A Suspicious Activity Report
- e. None of the above

Foundations for Intelligence Analysis Training (FIAT)

Participant Follow-up Survey

The Institute for Law and Justice is conducting an independent evaluation of the FIAT course you took in Ft. Lauderdale, FL from February 16 – 20, 2004. The information you provide us is completely confidential and your name will not be associated with your answers in any way. We ask that you put your name on this sheet for tracking purposes only.

NAME:

1. What is your job title?
2. Is this the same position you held when you took the FIAT course in February 2004?
YES NO
3. Have you received a promotion or other career advancement since attending FIAT?
YES NO

If yes, do you think this is related to your attending FIAT? How?

4. As a result of attending the FIAT course, did you gain any new **KNOWLEDGE** in the areas of:

Crime Pattern Analysis	YES <input type="checkbox"/>	NO <input type="checkbox"/>
Association Analysis	YES <input type="checkbox"/>	NO <input type="checkbox"/>
Flow Analysis	YES <input type="checkbox"/>	NO <input type="checkbox"/>
Communication Analysis	YES <input type="checkbox"/>	NO <input type="checkbox"/>
Financial Analysis	YES <input type="checkbox"/>	NO <input type="checkbox"/>
Strategic Analysis	YES <input type="checkbox"/>	NO <input type="checkbox"/>

5. As a result of attending the FIAT course, did you gain any new **SKILLS** in the areas of:

Crime Pattern Analysis	YES <input type="checkbox"/>	NO <input type="checkbox"/>
Association Analysis	YES <input type="checkbox"/>	NO <input type="checkbox"/>
Flow Analysis	YES <input type="checkbox"/>	NO <input type="checkbox"/>
Communication Analysis	YES <input type="checkbox"/>	NO <input type="checkbox"/>
Financial Analysis	YES <input type="checkbox"/>	NO <input type="checkbox"/>
Strategic Analysis	YES <input type="checkbox"/>	NO <input type="checkbox"/>

6. Please briefly describe what you learned that was new and how you have used this knowledge:

7. Have you done anything differently on the job as a result of attending the FIAT?

8. In retrospect, did the FIAT course meet your expectations?

YES NO

Why or why not?

9. What was the **MOST SIGNIFICANT** thing you learned at the FIAT?

10. What exercise, activity, or lecture topic do you think was a **waste of your time**?

11. Have you shown your notebook or shared what you learned with your co-workers?

YES NO

12. Have you referred to the notebook since the FIAT training?

YES NO

Has it been useful? YES NO

13. What changes do you think could be made to the FIAT course to make it more effective?

14. Would you recommend the FIAT course to your coworkers?

YES NO

Why or why not?

15. Have you attended any other intelligence trainings or other courses **since** you took the FIAT course? Please describe the subject matter, number of hours, and provider.

Topic	Number of Hours	Training Provider

16. What is your supervisor's name?

Email:

Telephone:

THANK YOU for your participation in this survey. If you have any questions, comments, or concerns, please feel free to contact Kelly Bradley at the Institute for Law and Justice at (703) 684-5300, or by email at kelly@ilj.org.

Please return to Chera Baity at the Institute for Law and Justice

EMAIL: Chera@ilj.org

FAX: (703) 739-5533

MAIL: Institute for Law and Justice
1018 Duke Street
Alexandria, VA 22315

Foundations for Intelligence Analysis Training (FIAT)

Supervisor Survey

The Institute for Law and Justice, in conjunction with the National Institute of Justice, U.S. Department of Justice, is conducting an evaluation of the FIAT course NAME OF PARTICIPANT took earlier in Richmond on July 12-16, 2004. The purpose of this evaluation is to determine what, if any, impact the FIAT course had on the job performance of those who attended the training. As her supervisor, your feedback is very important in helping us assess if FIAT is meeting the foundational training needs of intelligence analysts. The information you provide us is completely confidential and your name will not be associated with your answers in any way. Your candor is appreciated.

PLEASE RETURN BY MAY 12, 2005

1. What is your job title? _____
2. After attending the FIAT course, did NAME OF PARTICIPANT gain new **KNOWLEDGE** in the following areas?

Crime Pattern Analysis	YES <input type="checkbox"/>	NO <input type="checkbox"/>
Association Analysis	YES <input type="checkbox"/>	NO <input type="checkbox"/>
Flow Analysis	YES <input type="checkbox"/>	NO <input type="checkbox"/>
Communication Analysis	YES <input type="checkbox"/>	NO <input type="checkbox"/>
Financial Analysis	YES <input type="checkbox"/>	NO <input type="checkbox"/>
Strategic Analysis	YES <input type="checkbox"/>	NO <input type="checkbox"/>

3. After attending the FIAT course, did this employee gain new **SKILLS** in the following areas?

Crime Pattern Analysis	YES <input type="checkbox"/>	NO <input type="checkbox"/>
Association Analysis	YES <input type="checkbox"/>	NO <input type="checkbox"/>
Flow Analysis	YES <input type="checkbox"/>	NO <input type="checkbox"/>
Communication Analysis	YES <input type="checkbox"/>	NO <input type="checkbox"/>
Financial Analysis	YES <input type="checkbox"/>	NO <input type="checkbox"/>
Strategic Analysis	YES <input type="checkbox"/>	NO <input type="checkbox"/>

4. Do you feel the FIAT training had **positive**, **negative**, or **no impact** on the work of NAME OF PARTICIPANT? Please offer a brief explanation.

5. Since this employee's return to the job after the FIAT course, is there anything that she has done differently as a result of this training?

6. How does this employee's work compare to her colleagues? Has this changed since the training?

7. If given the opportunity, would you send other staff to this training?
YES NO

8. Based on your knowledge of the FIAT course, is there anything you feel should be changed?

THANK YOU for your participation in this survey. If you have any questions, comments, or concerns, please feel free to contact Kelly Bradley at the Institute for Law and Justice at (703) 684-5300, or by email at kelly@ilj.org.

APPENDIX 7-C

Pre/Post FIAT Participant Self-assessment of Course Comfort

Level with Intelligence Analysis Subject Material

Subject Area	n		Completely Comfortable %		Fairly Comfortable %		Not Very Comfortable %		Avoid! %	
	<i>Pre</i>	<i>Post</i>	<i>Pre</i>	<i>Post</i>	<i>Pre</i>	<i>Post</i>	<i>Pre</i>	<i>Post</i>	<i>Pre</i>	<i>Post</i>
Crime Pattern Analysis	126	115	12.7	36.5	43.7	46.6	40.5	3.5	3.2	0
Association Analysis	126	116	16.7	57.8	44.4	38.8	35.7	3.4	3.2	0
Flow Analysis	126	115	11.1	50.4	39.7	46.1	45.2	3.5	4.0	0
Communication Analysis	126	115	11.1	53.9	39.7	42.6	43.7	3.5	5.6	0
Financial Analysis	127	116	13.4	30.2	44.9	61.2	34.6	8.6	7.1	0
Strategic Analysis	127	115	9.4	22.6	33.9	67.8	51.2	9.6	5.5	0

APPENDIX 7-D

Matched Pairs T-test Results of Pre/Post FIAT Course Comfort Level with Intelligence Analysis Subject Material by Experience Level

New Analysts

Subject Area	Pre-training Self-assessment		Post-training Self-assessment		<i>df</i>	<i>t</i>
	<i>Mean</i>	<i>SD</i>	<i>Mean</i>	<i>SD</i>		
Crime Pattern Analysis	1.63	.784	2.37	.560	53	6.965**
Association Analysis	1.65	.821	2.55	.571	54	8.167**
Flow Analysis	1.44	.788	2.45	.571	54	8.887**
Communication Analysis	1.56	.816	2.54	.539	53	8.009**
Financial Analysis	1.60	.830	2.15	.621	54	4.500**
Strategic Analysis	1.41	.813	2.06	.596	53	4.888**

**p<.001

Less than 1 Year's Experience

Subject Area	Pre-training Self-assessment		Post-training Self-assessment		<i>df</i>	<i>t</i>
	<i>Mean</i>	<i>SD</i>	<i>Mean</i>	<i>SD</i>		
Crime Pattern Analysis	1.55	.671	2.14	.468	21	4.161**
Association Analysis	1.55	.671	2.45	.596	21	5.684**
Flow Analysis	1.52	.680	2.38	.590	20	5.403**
Communication Analysis	1.38	.740	2.38	.669	20	4.583**
Financial Analysis	1.45	.800	2.09	.526	21	4.537**
Strategic Analysis	1.32	.568	2.05	.375	21	5.405**

**p<.001

More than 1 Year's Experience

Subject Area	Pre-training Self-assessment		Post-training Self-assessment		<i>df</i>	<i>t</i>
	<i>Mean</i>	<i>SD</i>	<i>Mean</i>	<i>SD</i>		
Crime Pattern Analysis	1.86	.756	2.46	.576	27	4.688**
Association Analysis	2.11	.629	2.68	.548	27	4.382**
Flow Analysis	1.96	.637	2.68	.548	27	4.954**
Communication Analysis	1.64	.731	2.61	.567	27	5.791**
Financial Analysis	1.93	.766	2.43	.573	27	4.145**
Strategic Analysis	1.61	.737	2.36	.621	27	5.281**

**p<.001

Chapter 8

Simon Wiesenthal Center's National Institutes Against Hate Crimes and Terrorism Training

Simon Wiesenthal Center

The Simon Wiesenthal Center, Inc. (SWC), located in Los Angeles, California, was founded in 1977 as a center for Holocaust remembrance and the defense of human rights. Today it has more than 400,000 member families and offices throughout the U.S., Argentina, Canada, France, and Israel. It has become one of the world's leading human rights organizations and is recognized as an important source for in-depth information and technical assistance on human rights. The SWC campus in Los Angeles contains a library and research center, and the Museum of Tolerance. The Museum of Tolerance functions as the educational arm of SWC and has worked to support law enforcement and other criminal justice agencies nationwide for nearly ten years.

The high-tech Museum of Tolerance provides an interactive view of the dynamics of racism and prejudice in modern society, including intolerance in America, active international hate groups, and recent genocides including those in Rwanda and Bosnia. The Holocaust Museum leads visitors through the events of World War II, beginning with the rise of Nazism and ending with the liberation and after-effects of the concentration camps. Each year, an average of 350,000 people visit these two museums.

Review of Hate Crimes and Terrorism Literature

Hate crimes have become the object of growing attention for federal reporting systems and state legislatures over the past two decades. These crimes, also known as bias crimes, are not separate types of offenses but are crimes identified by a specific motivation of the offender. There are variations in the definition of hate crimes from state to state.

Hate Crime Defined

What makes a crime a hate crime? A straightforward, widely accepted definition has never been firmly established, in part because "criminal acts motivated by bias can easily be

confused with forms of expression protected by the U.S. Constitution” (Bureau of Justice Assistance, 1997). After the term “hate crime” came into use in the late 1970s, it became “increasingly understood that criminal conduct takes on a new meaning when it involves an act motivated by bigotry” (Jenness 2003). The challenge has been in establishing a coherent legal definition. The legal definition of a hate crime depends on the law being considered, with definitions of protected groups differing from jurisdiction to jurisdiction. What these statutes have in common is that they are “laws that criminalize, or further criminalize, activities motivated by bias towards individuals or groups because of their real or imagined characteristics” (Jenness 2003).

Federal Legislation

In 1990, Congress passed the Hate Crime Statistics Act (HCSA) requiring the Attorney General to collect data about hate crimes, which is done by the Federal Bureau of Investigation (FBI) as part of the Uniform Crime Reporting (UCR) Program. Specifically, the HCSA calls for acquiring data on crimes that “manifest prejudice based on race, religion, sexual orientation, or ethnicity” (Hate Crimes Statistics Act 1990). This definition was in keeping with the data collection efforts of the Anti-Defamation League of B’nai B’rith (ADL) over the preceding decade. The ADL was the central player in national efforts to monitor hate crimes in the 1980s. In addition to tracking hate crime incidents against individuals based on their race, religion, and ethnicity, the organization also drafted model legislation (Jenness 1999). It was through the work of the National Gay and Lesbian Task Force (NGLTF) that “sexual orientation” was added as a protected group for the final wording of the 1990 Act. The HCSA was then expanded in 1994 to include disability.

The Hate Crimes Sentencing Enhancement Act (HCSEA), enacted as part of the 1994 Violent Crime and Law Enforcement Act, further expanded the federal definition of a hate crime. Under HCSEA, a hate crime is “a crime in which the defendant intentionally selects a victim, or in the case of a property crime, the property that is the object of the crimes, because of the actual or perceived race, color, religion, national origin, ethnicity, gender, disability, or sexual orientation of any person.” The Act directs the U.S. Sentencing Commission to provide a sentencing enhancement of “not less than 3 offense levels for offenses that the finder of fact at trial determines beyond a reasonable doubt are hate crimes” (HCSEA 1994).

While the HCSA and HCSEA address hate crimes committed against a variety of groups, other pieces of federal legislation also address hate crimes directed at specific groups. Under the Violence Against Women Act of 1994 (VAWA), “all persons within the United States shall have the right to be free from crimes of violence motivated by gender.” In response to a series of attacks against houses of worship—a reminder of the American struggle with religious intolerance—the Church Arsons Prevention Act was passed in 1998. This expanded existing Federal jurisdiction and facilitated criminal prosecutions for attacks against houses of worship.

State Legislation

In 1997, BJA released *A Policymaker’s Guide to Hate Crimes*, which assessed the existing laws and strategies designed to fight and prevent bias-motivated offenses. BJA noted that there are three basic approaches to hate crime legislation: prohibition of specific acts of intimidation, prohibition of general behavior motivated by bias, and enhancement of penalties for criminal acts motivated by bias (Bureau of Justice Assistance 1997). The major legislative and public policy initiatives identified at that time included (1) enactment in 39 states of laws that address hate crime, many of them based on the ADL’s model law; and (2) statutes in 19 states mandating the collection of hate crime data. Today, nearly every state and the federal government require sentencing enhancements for offenders who commit hate crimes; and as of 2003, the ADL counted 25 states that mandate hate crime data collection (Anti-Defamation League 2003).

State statutes vary most with respect to the groups that are protected. For example, Wisconsin provides for sentencing enhancements if the defendant selected the victim because of “race, religion, color, disability, sexual orientation, national origin or ancestry” (Wis. Stat. 939.645, 1991-1992). Only six states do not have specific protections based on race, religion, and ethnicity, but most of those do have statutes that criminalize the interference with religious worship. Thirty states have statutory protections based on sexual orientation, 30 states also include disability in their statutes, and slightly fewer (27) identify gender as a protected group (Anti-Defamation League 2003). Another difference in state hate crime statutes is whether there are statutory mandates concerning training for law enforcement personnel. While the ADL notes that some states have administrative regulations that require training, that training is mandated by law in only 12 states.

State hate crime statutes have been challenged in the courts on the grounds that they violate the First Amendment of the U.S. Constitution, which restricts the government's power to make laws impinging on an individual's freedom of speech and expression. State appellate and high courts have generally upheld the constitutionality of the hate crime laws, and the U.S. Supreme Court has largely denied appeals of those decisions. While not all challenges have been the same, generally the finding is that "hate crime statutes do not conflict with the tenets of the First Amendment because they do not punish an individual for exercising freedom of expression but rather for motivation for engaging in criminal activity, a fact often considered when evaluating the seriousness of an offense" (Bureau of Justice Assistance 1997).

Statistics

Since the passage of the HCSA, the UCR has become the primary source of statistics on hate crime incidents reported to law enforcement in the U.S. In 2003, approximately 7,500 hate crime incidents, involving more than 9,000 victims, were reported to the FBI. All but four reported incidents were considered single-bias motivated. Racial bias represented the largest percentage (51.4 percent) of offenses, with bias against blacks accounting for 66.3 percent of racial bias victims. Religious bias-motivated offenses comprised 16.4 percent of incidents, and more than two-thirds of the victims were Jewish. Physical or mental disability bias made up the smallest number of reported bias-motivated incidents in the 2003 UCR, accounting for only 43 victims. Sixty percent of all incidents were crimes against persons, including simple assault, aggravated assault, and objects of intimidation. Less than one percent of victims was raped or murdered (Federal Bureau of Investigation 2004).

The FBI is not the only source of hate crime statistics. The ADL, as it did in the years before the HCSA, continues to collect and publish data on hate crime incidents. The Southern Poverty Law Center (SPLC), the National Coalition of Anti-Violence Programs (NCAVP), and other human rights organizations also track hate crime incidents and hate group activities on an annual basis. Because of differences in definitions and tracking mechanisms from the FBI, these organizations often report numbers that differ from those in the UCR. For example, the ADL reported 1,557 anti-Semitic incidents in 2003 (Anti-Defamation League 2004), which is significantly higher than the 927 reported incidents in the UCR. Similarly, the NCAVP counted 1,720 anti-LGBT incidents (National Coalition of Anti-Violence Programs 2005), whereas the

UCR reported 1,239 in 2003. These sorts of differences only serve to highlight the wide variety of ways in which hate crime is defined in this country, as well as the challenges inherent in counting incidents of it.

Challenges to Data Collection

The FBI collects data on crimes reported to the police, whereas private organizations like ADL and SPLC tend to gather information on all incidents that are reported regardless of whether they represent a crime. In addition, there are other challenges associated with collecting and using hate crime data.

Underreporting of hate crime incidents to the police has been the subject of numerous studies. Crime reporting in general is relatively independent of demographic characteristics (Kuehnle & Sullivan 2001). However, several studies in the late 1980s and early 1990s found that lesbians and gays were reluctant to report hate crimes to the police, and that they believed the police would treat them with indifference if they reported a hate crime (Berrill & Herek 1992; Comstock 1989; Finn & McNeil 1987). Similarly, Christopher et al. (1991) concluded that the history of discrimination by law enforcement makes many victims of racially motivated hate crimes reluctant to report to the police.

Research has also explored the failure of law enforcement agencies to correctly identify and categorize hate crime incidents. A New York study found that victims of anti-gay or anti-lesbian violence reported those crimes 37 percent of the time, but only 68 percent of those reports were officially labeled as bias crimes. A study examining bias incidents handled by the Boston Police Department found that less than 10 percent were appropriately identified as bias incidents by the reporting officers (Nolan & Akiyama 1999). “It seems that the failure of crime victims to report bias incidents to police, combined with police misidentification or failure to identify these crimes, add to the already burdensome and complex task of collecting meaningful data on hate crimes” (Nolan & Akiyama 1999).

Terrorism

While there is a lack of agreement on the definition of terrorism, the elements provided in definitions used by the FBI, U. S. Department of Defense, and U.S. Department of State can be summarized to conclude that terrorism is “the deliberate creation and exploitation of fear through

violence or the threat of violence in the pursuit of political change. Terrorism is specifically designed to have far-reaching psychological effects beyond the immediate victim(s) or objects of the terrorist attack” (Hoffman 1999). In light of the literature already discussed, it is logical to conclude that there is an overlap between hate crime and terrorism (Seymour et al. 2002) and that “if hate crimes were to be interpreted in the broadest sense, then they include acts of terrorism” (Weaver & Wittekind 2002). Petrosino (1999) has also noted that “hate crime, on some levels, is becoming indistinguishable from domestic terrorism.”

The question remains, though, whether an attempt to combine terrorism and hate crime in the same package hampers the ability to address either problem. The National Academy of Sciences (NAS) addressed this at a 1997 meeting. Attendees “struggled to identify the connections between hate crime and terrorism to develop a definition of bias-motivated activity that might embrace the concepts of both.” (Seymour et al. 2002). NAS concluded that despite significant similarities between hate crime and terrorism, they warrant a separate, distinct, and individualized approach. Others, however, still suggest expanding the definition of hate crime to include terrorism. If hate crime includes crimes based on prejudice against national origin, then events such as those occurring on September 11, 2001—in which “the intended victims were essentially identified by the perpetrators as being American in terms of national origin” (Weaver & Wittekind 2002)—can be counted among bias-motivated crimes.

Training

As hate crime legislation spread across the country, a number of training and technical assistance programs were started to help criminal justice professionals address hate crime. The following is not an exhaustive list of such resources, but it does indicate the range of agencies and organizations involved in the effort.

The U.S. Department of Justice created a four-part curriculum for law enforcement agencies as part of the Hate Crime Training Initiative. The National Organization for Black Law Enforcement Executives (NOBLE) drafted a training program on racial and religious violence and harassment. NOBLE also produced a curriculum aimed at culture clashes on college campuses. The Southern Poverty Law Center began its Teaching Tolerance program in 1991. The ADL offers the World of Difference Institute to help law enforcement professionals develop awareness and skills for working within a pluralistic society. The International Association of

Chiefs of Police (IACP) and the American Prosecutors Research Institute (APRI) have both produced guidebooks on responding to hate crime. The IACP offers *Responding to Hate Crimes: A Police Officer's Guide to Investigation and Prevention*, which is a compact guide and reference tool when responding to hate incidents and hate crimes. APRI created the *Prosecutors Respond to Hate Crimes Project* and produced a resource guide on hate crimes aimed at local prosecutors.

Teaching Tools for Tolerance

In 1994, SWC started the Tools for Tolerance program to provide professional education and training for students, police, government, education, and social service workers. This experiential program challenges participants on issues of tolerance, diversity, personal values, and responsibility as they relate to the workplace and the community. The Tools for Tolerance program is tailored to meet the needs of different professional groups, including educators, corporate executives, and municipal employees. In 1999, with funding from BJA, the National Institutes Against Hate Crimes and Terrorism was initiated.¹⁵ This program helps criminal justice professionals develop partnerships and strategies to combat hate-motivated violence and terror.

Tools for Tolerance National Institutes Against Hate Crimes and Terrorism

The purpose of the Tools for Tolerance National Institutes Against Hate Crimes and Terrorism (National Institutes) is to help participants develop new perspectives on hate crime and terrorist acts, aid them in forming multi-agency collaborations, and foster the development of strategic action plans for combating hate crime and terrorism. The goal of the National Institutes is to "...further enhance the criminal justice professional's ability to effectively address hate crimes and the wave of terrorist threats in their communities" (SWC National Institutes Against Hate Crimes and Terrorism curriculum 2004).

The objectives of the training are to (1) explore the impact of hate crime and terror on law enforcement, (2) create a structured forum for participation, and (3) introduce new definitions and concepts of punishment. Since 1999, SWC has used the BJA grant funds to

¹⁵ The term "terrorism" was added a couple of years after the Institutes were initiated.

conduct the training, cover travel and accommodations for all participants,¹⁶ provide technical assistance to individuals and teams engaged in enforcement and prevention initiatives, and monitor the activities of National Institute teams (SWC Proposal 2002). The grant is directed specifically for training at two SWC sites: the New York Tolerance Center and the Museum of Tolerance located on the campus of SWC world headquarters in Los Angeles. The New York Tolerance Center opened in 2004 to provide an East Coast alternative to the Museum of Tolerance. It functions as a multi-media professional development and training facility for educators, law enforcement, and state and local government officials.

National Institutes Course Structure

At least four times each year, SWC hosts the National Institutes. By 2005, 24 institutes had been conducted and two plenary sessions had been held.¹⁷

Target Population

Each National Institute consists of jurisdictional teams of police officers, judges, district attorneys, public defenders, probation officers, social service providers, and educators. This team composition is mandatory and considered vital to the training. The training targets a national audience, with teams from all over the country attending the National Institutes, creating a network of criminal justice professionals whose common interest is to fight hate crime and terrorism.

Curriculum

The National Institutes is a four-day training program consisting of three training modules: experience and environment, new perspectives, and innovative strategies. The first module involves a tour of the interactive Museum of Tolerance and the Holocaust Museum, where participants explore the history of American racism and bigotry and the inhumanity of the Holocaust. The new perspectives module focuses on definitions of hate crime and terrorism, motivations of offenders, and the ethical implications for justice workers. The innovative

¹⁶ Covering all travel and accommodation expenses for participants helps with attendance because many local and state jurisdictions have limited travel budgets for out-of-state travel.

¹⁷ The plenary sessions, sponsored by the SWC, bring together previous National Institutes attendees. They are given an opportunity to network with other jurisdictions and relate the successes and challenges they have encountered since attending the training. In addition, SWC staff and training facilitators provide a training review to rejuvenate the attendees in their efforts to deal with crimes of hate and bias.

strategies module directs participants to discuss and debrief about problems in their cities and devise a strategy of action directed at the problems in their jurisdictions.

The format for each four-day National Institute is the same except for the scheduled guest speakers who provide personal testimony about their experiences with hate and prejudice. On day one of the training, there is a welcome and orientation by SWC staff, a guided tour of the Tolerance Center, and a personal testimony session. For Institute #17, Dr. Terrance Roberts, an original member of the Little Rock Nine, gave his personal testimony. The 10th Anniversary of the Rwandan Genocide occurred during the time of Institute #18, and the SWC featured a special presentation and speakers on the genocide in Rwanda. Institute #19 occurred during the 50th anniversary of the landmark *Brown vs. Board of Education* decision, and the training featured a special speaker and panel that addressed some of the issues leading up to the *Brown* decision. The personal testimony session is followed by a tour of the Holocaust Museum, which gives visitors a graphic depiction of the atrocities of the Holocaust. After the tour, Elaine Geller, a Holocaust survivor spoke to one class about her experience as a child in the Holocaust. At the end of day one, there is a jurisdictional team meeting where participants meet with one another and discuss issues facing their community.

Day two of the training features a lecture on consensus building and terrorism conducted by David Lapin and Desi Rosenfeld of Strategic Business Ethics, Inc. (SBE). The objectives of the workshops are to stimulate, challenge, and question participants' thinking about hate crimes and terrorism. The lecture covers the definition of hate crime and terrorism; the "spectrum of tolerance;" and the culture, values, and ethics of offenders. In addition, SBE coordinates breakout sessions where students apply the information they are taught. The day ends with a lecture on constitutional law and an introductory lecture on terrorism. Participants are given strategies on how to involve the community and increase collaboration among criminal justice entities in order to combat terrorism.

Day three opens with a workshop led by SBE on the development of a strategic theory and strategic action plan. The strategic theory is a mission statement created in tandem by all the jurisdictions to guide them in their work. The strategic action plan is an individual team "blueprint for action." It is a plan developed by different teams to be implemented once the training is concluded. Teams focus on an area of concern in their jurisdictions and build a plan

around it based on the concepts they have been taught at the National Institutes. Some examples include hate crime education for law enforcement or anti-hate crime programs in schools. At the end of day three, participants are given a tutorial on how to access and use the Simon Wiesenthal tracking database. The database was created solely for National Institutes participants so that teams can correspond with one another and receive assistance on their strategic action plans from SBE and SWC staff.

The last day of the training features a demonstration of a racial profiling training CD developed by SWC for law enforcement. The demonstration involves interactive class exercises and a question and answer period. Personal testimony from a reformed white supremacist and a presentation by a SWC senior researcher on contemporary hate crimes is also featured. An Institute graduation concludes the training.

Evaluation Methodology

The purpose of this evaluation was to assess what effect the National Institutes training had on program participants in terms of their learning, behavior, and work. In particular, it sought to determine how training participants felt about the training, what new knowledge and skills they may have acquired, and how the training influenced their efforts to deal with crimes of hate and terror in their communities.

Evaluation Questions

As discussed earlier, in conducting this evaluation, the researchers used our expanded model of Kirkpatrick's framework. The following evaluation questions were examined (see also Exhibit 8-1):

- How did the training participants react to the training? What did they think of the content, different modules, speakers, facilitators, SWC staff, and facility?
- What did the training participants gain in terms of information and skills?
- Have the training participants experienced any attitude and behavior changes due to the training?
- What effect has the training had on the training participants' plans and strategies for combating hate crime and terrorism?
- Can the National Institutes training be improved? If so, how?

**Exhibit 8-1: SWC Training Evaluation Questions and Data Collection Tools
Grouped by Evaluation Level**

Evaluation Level	Evaluation Questions	Data Collection Tool
1: Reaction	How did the participants react to the training? Were they satisfied?	Student course evaluation Reaction interviews with participants
2: Learning	What information and skills were gained?	Pre-training interviews with team leaders Reaction interviews with all participants Follow-up survey with team leaders
3: Behavior Change	How have participants' attitudes and behavior changed due to the training?	Pre-training interviews with team leaders Reaction interviews with participants Follow-up survey with team leaders Analysis of strategic action plans
4: Organization Impact	What effect has the training had on strategies for combating hate and terror crimes? Can the training be improved? If so, how?	Follow-up survey with team leaders Descriptive case studies

Data Collection Methods and Tools

This section outlines the evaluation plan used in evaluating the SWC National Institutes program. The discussion focuses on the participants, evaluation design, data collection framework, and data collection tools.

Evaluation Plan

Participants

An important focus of the National Institutes training was on building teamwork. The SWC facilitators believe that the most effective means of achieving jurisdictional and organizational change is to require collaboration among parties that do not often work together.

To facilitate collaboration, each jurisdiction attending the training must send a group composed of a judge; prosecuting attorney; public defender; probation official; law enforcement official; and a school official, victim/witness advocate, or human relations commissioner. Before attending the training, team members often do not know each other. The training provides them an opportunity to establish a working relationship, which is further enhanced when they collaborate to create a strategic plan. As a result of this process, the strategic plans are richer for having input from people with very different roles and perspectives on the criminal justice system; and each member of the team is more likely to “buy into” the project.

A total of 105 participants for this study were drawn from three National Institutes trainings: #17 (conducted in January 2004), #18 (March 2004), and #19 (May 2004). Institute #17 had 40 participants comprising six teams. Institute # 18 had 39 participants comprising six teams, and Institute #19 had 26 participants comprising four teams (see Exhibit 8-2). The teams came from all across the United States, including Alabama, California, Florida, Idaho, Michigan, Montana, New Jersey, Ohio, Pennsylvania, West Virginia, and Wisconsin.

Exhibit 8-2. Breakdown of Institute Participants by Discipline

Institute	#17	#18	#19	Total
Corrections/Probation	6	3	2	11
Human Relations Commission	2	1	0	3
Judge	5	5	2	12
Police	6	6	13	25
Prosecution/Attorney General’s Office	6	14	4	25
Public Defender	6	4	2	12
Schools	1	1	2	4
Victim Services	4	2	1	7
Other*	4	3	0	6
Total	40	39	26	105

*Includes representatives from community organizations and government agencies such as the employment opportunity office.

Design

The SWC Institutes evaluation was a pre/post reflexive design (see Exhibit 8-3) conducted over a 16-month period. The pre-test involved interviewing SWC-identified team

leaders before they attended the National Institutes training. The post-test involved two stages. The first stage consisted of student evaluation surveys given to each participant at the immediate conclusion of the training, reaction interviews with all participants, and a review of each team’s strategic action plan. The second stage involved follow-up surveys with team leaders, a review of the progress on each team’s strategic action plan, and descriptive case studies of two different teams that attended the training.

A reflexive design was used because the National Institutes training is unique training that does not lend itself to any natural comparison groups: jurisdictions were required to form multi-agency teams specifically for the training. The level of effort and expense that would have been required to (1) develop an adequate comparison group and (2) obtain cooperation from its members was beyond the scope of this evaluation. Consequently, use of an experimental design was not possible. The design employed, however, permits determination of whether participation in the National Institutes met the goals set forth by SWC, and if this was over and above what would have occurred if these teams had not attended the training. For instance, through interviews it is possible to assess how a jurisdiction defines and deals with hate crime and terrorist acts both before and after attendance at the training.

Exhibit 8-3: Reflexive Design of the National Institutes Training

Pre-test	Training	Post-test	2 nd Post-test
O _a		O _b O _c O _d	O _d O _e O _f
T ₁	T ₂	T ₃	T ₄

Where:

- O_a = Interviews with team leaders
 - O_b = Reaction interviews with all course participants
 - O_c = Student course evaluation
 - O_d = Strategic action plans
 - O_e = Case studies
 - O_f = Follow-up survey of team leaders
-

Data Collection Framework

Each of the four levels of training evaluation noted earlier is discussed below in the context of this evaluation.

Level 1: Reaction

The key question for Level 1 was: How did the National Institutes training participants react to the training? Answering this involved administering student course evaluation surveys that asked participants about each individual training component and their overall reaction to the training. The survey instruments were developed by SWC and were standard procedure for all of the National Institutes (see Appendix 8-A for examples). Evaluation team staff also conducted reaction interviews by phone with a sample of training participants.

Level 2: Learning

The key evaluation question at Level 2 was: What information and skills were gained by National Institutes training participants? Answering this involved conducting pre-training interviews with team leaders, post-training reaction interviews with participants, and follow-up surveys with team leaders. The research team developed these survey instruments (see Appendix 8-A for examples).

Level 3: Behavior Changes

At Level 3, the key evaluation question was: How have National Institutes participants' attitudes and behaviors changed because of the training? Answering this involved conducting pre-training interviews with team leaders, follow-up surveys with team leaders, and analyzing each team's strategic action plan. The research team developed the interview protocols and follow-up survey instrument. The SWC developed the strategic action plan format completed by each team.

Level 4: Organizational Impact

The key evaluation questions at Level 4 was: What was the training's impact on the participants' agencies after the training? As discussed earlier, evaluations at Level 4 are concerned with measuring training's impact on participants' organizations, which is often difficult to assess. Addressing this involved conducting two descriptive case studies. The research team developed the case study protocol.

Data Collection Tools and Methods

The six primary sources of data collected in this evaluation are discussed below.

Pre-training Interviews

Team leaders from each jurisdiction were contacted for pre-training interviews prior to the start of the National Institutes training. The purpose of these interviews was to (1) discover how the teams defined hate crime and terrorism; (2) learn more about what hate crime and terrorism issues were prevalent in the communities represented at the training; and (3) learn what the teams hoped to gain from the training, and what provided the catalyst for them to get involved with the SWC. The eleven interview questions were:

- How did you hear about the training?
- How did you go about selecting members for your team?
- What are some tolerance and/or hate issues facing your community?
- Is there a specific incident or event that prompted you to get involved in the training?
- How do you define hate crime?
- How do you define terrorism?
- How big a problem is hate crime and terrorism in your local jurisdiction?
- How has the impact of hate crimes and terrorism changed your role as a (law enforcement officer, prosecutor, etc.)?
- Currently, what is your agency's approach in dealing with hate crime and terrorism?
- Have you ever been to any training on hate crime and/or terrorism? If yes, what do you think is different about the National Institutes Against Hate Crime and Terrorism?
- What do you hope to gain from this training?

To conduct the interviews, SWC staff first contacted all participants by email to explain the evaluation, introduce the evaluators, inform them that the evaluators would be attending the training, and seek their cooperation. Following this email, the evaluators sent an introductory email to the self-identified team leaders asking them to participate in the pre-training interview and seeking a specific time and date for conducting the interview. Interviews were conducted by telephone and averaged approximately 20 minutes each.

Team leaders were selected as the source of pre-training data for two reasons. First, they were instrumental in creating the teams; and second, the participant list often changed until just before the training. The fluidity of the participant list made it virtually impossible for the evaluators to interview each member of the team. Of the team leaders contacted, most agreed to participate in the interview. Of the 12 team leaders from Institutes 17 and 18, ten completed a pre-training interview.¹⁸

Student Course Evaluations

A second source of data, student course evaluations, relied on the SWC's own procedures for collecting student feedback. The SWC uses student course evaluations in each of its courses to obtain student reactions to a range of training elements, including materials, presenters, facilitators, and the Museums of Tolerance and Holocaust. These surveys were administered to each participant at the conclusion of each day, and at the end of the entire training (see Appendix 8-A). With the exception of four missing surveys from the first day of Institute 18, student course evaluations were collected after each module and for the entire National Institutes training from each attendee. SWC staff collected the forms from the training participants and sent the researchers the evaluations (hard copies of the completed forms and in data entry format).

Reaction Interviews

A third source of data consisted of interviews with course attendees after they returned from the National Institutes training. The purpose of these interviews was to assess what concepts, definitions, and strategies the participants learned in order to address issues of hate and terror in their communities. The following topics were covered:

- Key knowledge, skills, or attitudes learned at the training
- Effects the training had on their job roles and responsibilities
- Team meetings and follow-up
- Strategic theory
- Strategic action plan
- Tools from the training to effect change
- New concepts learned

¹⁸ For Institutes 17 and 18, one team leader agreed to an interview but was never available to complete it, and the other was never contacted due to an oversight by the research team. Pre-interviews were not conducted with Institute #19 because the evaluators did not receive the contact list before the training started.

- Improvements for the training
- Benefits of jurisdictional meetings
- SWC National Institutes tracking database
- Changes needed in participants' organizations

The evaluation team contacted each participant by email to set up an appointment time for conducting the interview. Participants who did not respond to this email received up to five subsequent calls and emails in an effort to contact them. Participants who did respond to the email were called at the appointed time for the interview. The interviews lasted for an average of 30 minutes and were completed within two to six months after the training.

The original goal was to complete these interviews within one month of the training. However, it proved challenging to connect with the training participants by either email or telephone. Consequently, this original goal was extended to six months. Even with this extension, the overall response rate was low at 62 percent. Nevertheless, at least 50 percent of attendees from each discipline were surveyed. The breakdown of participation was as follows: interviews were conducted with 73 percent of corrections/probation officers, 100 percent of human relations commissioners, 58 percent of judges, 56 percent each from law enforcement and prosecution, 50 percent of public defenders, 75 percent of school officials, 71 percent of victim services personnel, and 83 percent of the community group and other personnel.

Usual explanations for low response rates include a lack of enthusiasm about the subject matter or general dislike for a program, survey materials or questions so lengthy and complex that they are burdensome on the respondents, or an inability to contact respondents due to their relocating without providing forwarding contact information. In the case of this training, however, the attendees were enthusiastic about the training and committed to the subject matter, the survey questions were purposely kept brief, and none of the attendees had relocated. The most likely explanation for this low response rate is that many of the attendees were over-extended in their work commitments and simply could not make the time to schedule an interview.

Strategic Action Plans

A fourth data source was the strategic action plans developed by each team during the training. These plans were instrumental to the training and required by the SWC and National Institutes instructors for completion of the training.

The purpose of the plans was to “...formulate new strategic approaches to combat hate crimes and terrorism based on a fresh understanding of the unique elements that differentiate them from other criminal acts” (SWC Museum of Tolerance brochure). To formulate the plans, teams first identified one of four possible pre-determined targets (school children and parents, offenders, criminal justice system colleagues, or every member of the public). Next, each team member outlined both their capabilities and “passions.” That is, they were asked to identify “what are you capable of *achieving* with your strength to effectively address your target? And what personal passions do you bring to your work in this area?” (Strategic Business Ethics 2004, p. 29, emphasis in original). Finally, team members created their action plan by brainstorming on what each discipline can do to “...enhance your discipline’s contribution to the Hate Crime and Terrorism battle in light of your Strategic Theory” (Strategic Business Ethics 2004, p. 30), and then outlining what needs to be done, how it is going to get done, by when, and by whom.

During the course of the training, the teams were given an opportunity to develop their plans and share them with the SWC staff, training instructors, and other participants. SWC staff collected the plans from each team and provided them to the evaluation team. The response rate for the strategic action plans was 100 percent. The research team used the strategic action plans as a measure for assessing the progress these teams made toward their self-defined goals.

Follow-up Surveys

A fifth source of data consisted of follow-up surveys conducted with the team leaders from the Institutes. The purpose of the follow-up survey was to determine if teams had implemented their strategic action plans and to document any issues or challenges in implementing individual action items (see Appendix 8-A). Questions were also asked about whether teams were properly structured, cohesiveness among team members, and whether adequate time was devoted to working the plan. To give each team maximum opportunity to implement their strategic action plans, these surveys were sent 12 to 16 months post-training.

To conduct the surveys, the evaluation team sent each team leader the survey by email. Participants who did not respond received at least five follow-up telephone calls and emails, requesting their assistance in completing the survey. Of the 16-team leaders from Institutes 17 through 19, twelve participants (75 percent) completed the surveys.

Case Studies

A sixth source of data was from descriptive case studies of two selected jurisdictions that sent teams to receive training during the evaluation period—Madison, Wisconsin, and Monmouth County, New Jersey.¹⁹ The case study locations were chosen because they had been praised consistently by other participants for having established model programs in their respective jurisdictions. In addition, teams from Madison had been attending the training regularly for several years and had recently created a Hate Crimes Task Force composed of members who attended the National Institutes training. Monmouth County was selected as a case study site because its innovative strategies were being emulated by other sites, including Madison (see Appendix 8-B for case study reports).

The case studies allowed for a more in-depth analysis of activities that teams implemented after attending the National Institutes. The purpose of the case studies was to provide a more thorough understanding of how the training was being used and what the program participants did after returning from the National Institutes training. Of particular importance were reactions to the training, effects the training may have had on trainees' responsibilities, and policy or program changes made as a result of the training. Capturing this information is central to assessing the outcome of the SWC Institutes program. The logic for the case study is found in Yin (2003, p. 1): "In general, case studies are the preferred strategy when "how" or "why" questions are being posed, when the investigator has little control over events, and when the focus is on a contemporary phenomenon within some real-life context."

The case studies were developed by conducting semi-structured interviews with key stakeholders and through field observation. Semi-structured interviews were used because they provided uniformity while also allowing the interviewers to digress from the interview protocol and ask probing questions as warranted. The interviews were approximately 60 minutes in

¹⁹ Two ILJ researchers visited Monmouth County, New Jersey, from December 7-9, 2004, and visited Madison, Wisconsin, on April 20-21, 2005.

length. The shortest interview was 30 minutes and the longest was more than two hours.

Sources of data included:

- Interviews with 30 team members, their colleagues and supervisors, commission and task force members, and community activists
- Direct observation of training at the Monmouth County Police Academy and at the New Jersey Department of Corrections Training Academy at Sea Girt, the Brookdale Community College's Juvenile Bias Program, and a meeting of the Madison Hate Crimes Task Force.

The analyses of the interview, strategic action plans, and case study data consisted of discovering themes and categories developed during the interviews with the key stakeholders. Interview field notes were transcribed and checked for accuracy. The case studies were written up as separate, stand-alone reports and submitted to the sites for verification of accuracy. During the data collection phase, tentative categories and themes were developed based on categories derived from the interview protocol and from ideas formed during the interviewing process. The tentative categories were formed based on a combination of inductive mechanisms (e.g., immersion in the data) and deductive mechanisms (e.g., categorical schemes suggested by program theory and policy) as recommended by Berg (2001). The tentative categories and themes provided a framework for additional questions to be asked of respondents during the on-going interview process. The field notes generated from the interviews were coded individually by each interviewer, with the suitability of categories checked during case analysis meetings. Case analysis meetings are a form of qualitative analysis that involves back and forth discussion among colleagues about the data being collected and the conclusions being drawn, for the purpose of allowing opposing points of view to be examined (Miles & Huberman 1994).

Strengths and Weaknesses

As with any evaluation, there are design strengths and weaknesses. Strengths of this National Institutes training evaluation included access to the SWC staff and facilities, National Institutes trainers, and training participants. The SWC staff was incredibly open and responsive to this evaluation, allowing the evaluation team the freedom to observe and collect whatever data were possible. Additional strengths included a design that permitted capturing as much data as possible about the individualized outcomes of the National Institutes teams. That is, while each team participates in the same training, each produces different strategic action plans with

different goals and tasks. This evaluation provided enough flexibility to capture that information.

Potential weaknesses included bias in design choice and methods, validity issues with the data, and low response rates. The principal source of bias with a reflexive design is accounting for changes in participant behavior for reasons unrelated to the intervention (in this case, attendance at the SWC Institute program). For instance, some jurisdictions might have decided to form multi-agency collaborative teams to develop a strategic action plan for combating hate crimes and terrorist acts independent of attending the SWC training. The best remedy for this source of bias is to obtain multiple outcome measures (Rossi, Lipsey, & Freeman 2004). In addition, although there were multiple data sources, the bulk of our findings relied upon interview and survey data. Both the interview and the survey data suffered from relatively low response rates, thus calling into question whether those who did not participate were somehow different from those who did participate in this study. The key validity issue with qualitative data is trustworthiness of the research (Arksey & Knight 1999). Valid qualitative research can be defined as the “...correctness or credibility of a description, conclusion, explanation, interpretations, or other sort of account” (Maxwell 1996, p. 87).

As explained earlier, the research team had relatively low rates of participation in the follow-up interviews with training participants, despite repeated efforts to contact them by both telephone and email. These marginal rates of participation presented challenges in measuring the effects of the training.

Evaluation Findings

Central to the evaluation design of the National Institutes training is assessing program effectiveness in changing participants’ behavior. This training is a unique program that requires participants to seek change within their jurisdictions to address hate crime and terrorist acts; thus, a key question was: How was the training being used? That is, what did participants do after they returned home from the Simon Wiesenthal Center in Los Angeles? Training participants were expected to be proactive in applying their unique, team-developed solutions to the problems being faced.

The National Institutes staff recognizes that effective training is not one-size-fits-all. The training teaches a set curriculum on hate crime but individualizes the training for each attending team. The curriculum covers different types of hate crime and acts of terrorism and explores the ethical dimension of hate crime and terror threats. It is individualized in that each team develops its own strategic action plan based on the collaborative strategies learned in the training.

The following sections discuss the results of this evaluation with respect to participant reactions to the training, knowledge and skills gained, attitude and behavior changes, results of the training, and recommended improvements.

Participant Reaction

Participants were asked for their opinions about the different training modules, course content, trainers, guest speakers, and the Museum of Tolerance. Findings were derived mainly from the student course evaluations and study team reaction interviews.

Reaction to Training Modules

As noted earlier, National Institute participants were provided a combination of training modules, guest speakers, and Museum of Tolerance resources.²⁰ The training modules covered the following areas:

- *Experience and Environment*, which addresses both the historical and contemporary dynamic relationship between stereotypes, prejudice, discrimination, oppression, and genocide
- *New Perspectives*, which addresses the definition of hate crime, its differences from other types of crimes, and the impact of the threat of terrorism and hate crimes on the criminal justice system
- *Innovative Solutions*, which includes multi-disciplinary peer discussions assessing the implications of the new perspectives, the application of new perspectives in each discipline (law enforcement, probation, etc.), and new collaborative responses to the challenges of hate crimes in the teams' jurisdictions.

The *Experience and Environment* training module focused on the Museum of Tolerance experience, in which participants were systematically walked through a series of exhibits that discussed stereotypes and prejudice and illustrated how hate has been and still is propagated throughout the world. This module was led by SWC staff.

²⁰ See Simon Wiesenthal Center proposal submitted to the Bureau of Justice Assistance (May 2, 2002) and www.museumoftolerance.com for more information.

The *New Perspectives* module attempted to redefine and reframe how acts of hate and terror were defined by participants. This module was led by a team from Strategic Business Ethics, Inc. (SBE). The module included lessons on culture, values and ethics, consensus building, and collaboration. The session included an extensive review of the definition of hate crime and offenders’ motivations, as well as a discussion of appropriate forms of punishment for these violators. Different hate typologies and the “culture-bias-ethics paradigm” were also reviewed. Using case study examples, participants learned about the spectrum of tolerance and saw how hate crime offenders are driven by instincts.

The *Innovative Strategies* module provided participants both time and resources to develop a collaborative response to problems of hate and terror in their own communities. In particular, this module (also led by the SBE team) brought participants together through workshops and planning sessions. During the first day of training, participants met with other team members from their jurisdiction to discuss the problems and issues they faced. On the third day, the participants developed a strategic theory and action plan.

Participant reaction to the training modules was quite high (see Exhibit 8-4). The museum experience was rated highest, with a mean score of 3.86 on a 4.0 scale. The *New Perspectives* module was also highly regarded, especially the relevance (3.60) and the “new” thinking aspect of the material (3.55). The jurisdictional team meetings, part of the *Innovative Strategies* module, received the lowest ratings (3.13), which was still a relatively high score.

Exhibit 8-4: Participant Reaction to the National Institute Training Modules

Training Module	n	Mean Score Scale 1 (poor) to 4 (excellent)
Experience & Environment		
Museum experience	99	3.86
New Perspectives: Consensus Building		
“New” thinking	97	3.55
Relevance of subject	93	3.60
Relevance to work	94	3.15
Quality of written materials	99	3.45
Quality of visual materials	98	3.33
Innovative Strategies		
Jurisdictional team meeting	77	3.13
Usefulness of group work	90	3.41

Reaction to the museum experience module was overwhelmingly positive. Comments included:

- “Thought provoking”
- “Amazing – informative”
- “Very informative, powerful, and well guided.”
- “The museum is wonderful, the exhibits are well thought out and meaningful.”
- “Phenomenal – very informative and educational”
- “Outstanding”
- “Excellent experience although extremely painful to see man’s capacity for evil against man as well as man’s capacity for courage in the face of extreme danger.”
- “It was an incredibly moving and powerful experience. I only wish we had more time to read all of the exhibits.”
- “Very informative and created debates among us.”
- “It was a great experience!”
- “Well done, quality exhibits, could have been interesting to split museum experience into 2 days.”
- “It was difficult to digest at the pace that was set. Try to find a way to make the groups smaller and allow more frequent intervals to reflect and discuss.”
- “Has to be more than just hate. Excellent museum! A must. Need[ed] for people in our profession.”
- “What an outstanding, interactive exhibit! The emotions, facts, and overall experience was overwhelming—positively overwhelming. Public officials and office holders should be mandated to attend!”
- “Beyond belief. One of the most awesome experiences of my life.”
- “I expected not to like it but found the exhibits very thought/emotion provoking. I think the gas chamber was in bad taste and trivialized the event.”

The experience was also quite challenging for participants, as evidenced by the following comments:

- “It was emotionally draining”
- “I never imagined the level of emotion, sadness, pain that I was going to feel today, thank you.”
- “Powerful, but almost too much at once.”
- “Rushed a bit.”

- “What an overwhelming first day! Material was so intense and day was incredibly long.”
- “Very thought provoking. I wonder if there were parts so overwhelming that we may have benefited from the question, does anyone need a time out?”

In general, participants had favorable responses to the *New Perspectives* module as they were challenged to think in new ways. Armed with new definitions and perspectives on hate crime, SBE directed the class in an exercise on developing a strategic theory and a strategic action plan. Participant comments ranged from positive (“very good – kept my attention, kept me focused”) to negative (“I’m not sure this is new thinking”). Additional positive comments included:

- “Beyond words!!”
- “Right on point”
- “I truly enjoyed the philosophical approach.”
- “This course made you think outside the box.”
- “When you ask for bigger picture you need the higher-ups of the organization to implement change to be present at the training.”
- “Very challenging.”
- “Incredible, thought provoking.”

Negative comments included the following:

- “Regarding linking hate speech to hate crimes to terrorism, one thought: if you can’t fight this conventionally because that creates martyrs, I did not hear a clear message regarding the theory about how we can fight it.”
- “It was a little unclear as to what we were to do.”
- “Fairly disappointing – too superficial and not very related to the criminal justice sector. It seems as if they would be great for businesses. Some of the notions were outright wrong. I have never met a defendant who wished to do a year in jail, ever.”

Most participants thought the *Innovative Strategies* module was a worthwhile exercise. However, most feedback on the team meetings and group work were along the lines of a comment by one participant who stated, “probably needed to be done, but didn’t add much.” Other comments included:

- “Don’t feel that we were prepared to discuss this – we had not had any internal conversation on this issue. Perhaps a “head’s up” would have allowed us to think about it prior to the short team discussion.”
- “No time – no structure.”
- “Questions too general.”
- “Already aware of our problems.”
- “No surprises. No real insights.”
- “Good start, better things to come.”
- “It was good, however did not go into much detail.”
- “Thought provoking regarding all the work yet to do.”
- “Nice to talk to people from my area. First opportunity to do so, important for “bonding” – group dynamics.”
- “Good networking within the groups.”
- “We need more time in which to carry on this conversation.”
- “Not very helpful.”
- “There could have been more explanation of what was expected of each team.”
- “I like group atmosphere.”
- “Good discussion and start to get our group going.”
- “I would have liked to hear how other jurisdictions handle bias crime.”

Reaction to Institute Speakers

A key component of the National Institutes training is to use a variety of speakers, from victims and survivors, to training moderators, to a variety of lecturers in wide-ranging areas. The National Institutes speakers included:

- SBE, Inc., consultants David Lapin and Desi Rosenfeld, who led participants through three days of workshops on defining and differentiating hate crimes from other crimes, exploring the impact of hate crime prevention on the criminal justice system, presenting a strategic planning model to assist participants in developing an Institute hate crime theory, and assisting participants in designing a new jurisdictional strategic plan in response to the challenges of hate crimes.
- Holocaust survivor Elane Geller, who spoke about her experiences as a survivor of a Nazi concentration camp.
- Little Rock Nine activist Terrence Roberts (Institute 17), who spoke about his experiences as one of the original Little Rock Nine, the group of African American teenagers who integrated Central High School in Little Rock, Arkansas, in 1957.

- Rwandan Genocide Commemoration Program (Institute 18)
- *Brown v Board's* 50th Anniversary program (Institute 19)
- Former Neo-Nazi skinhead Tim Zaal, who discussed his experiences as a former skinhead recruiter and propagandist.
- American civil liberty constitutional lawyer, Stephen Rohde discussed constitutional issues, challenges, and safeguards that affect the pursuit and subsequent prosecution of suspected hate and terror perpetrators.
- Terrorism experts Micah Halpern (Institute #17), Anthony Lukin (Institute #18), and Sabi Shabtai (Institute #19), who provided their perspectives on terrorism.
- Contemporary hate groups researcher Rick Eaton, who provided insight into hate groups worldwide. He also introduced and distributed two SWC resources, the *Digital Terrorism and Hate 2003* CD-Rom and the *New Lexicon of Hate*.

In general, the National Institutes speakers were well received (see Exhibit 8-5). The holocaust survivor was widely regarded as excellent, receiving a 3.9 out of a possible 4.0 mean score rating, while the *Brown v. Board of Education* speaker was considered marginal, receiving a 2.63 mean score rating. Most speakers, however, received ratings of 3.4 or higher.

Exhibit 8-5: Participant Reaction to the National Institutes Speakers

Speaker	n	Mean Score Scale 1 (poor) to 4 (excellent)
SBE Consultants		
Effectiveness of presenters	97	3.41
Personal Testimony		
Little Rock Nine activist	40	3.78
Holocaust survivor	99	3.90
Rwandan Genocide Commemoration	35	3.83
Brown v Board of Education speaker	24	2.63
Reformed Skinhead		
Effectiveness of presenter	97	3.51
Content of presentation	97	3.43
Quality of presentation	97	3.40
Meaningfulness	95	3.31
Contemporary Hate Groups		
Effectiveness of presenter	91	3.56
Content of presentation	92	3.43
Quality of presentation	92	3.40
Meaningfulness	90	3.39

Strategic Business Ethics, Inc., Consultants

The Strategic Business Ethics, Inc., consultants were the most prominent trainers at the training. They provided the bulk of the National Institutes training, and it was their job to help the participants through a paradigm shift in how they conceptualize acts of hate and terror. Given such an intense task as leading participants into redefining and reframing hate and terror crimes, it is understandable that reactions to these two instructors were mixed. Some of the key positive reactions included:

- “Excellent instructors and facilitators.”
- “Dave and Desi are excellent communicators, very fine presence and content.
- “Did an excellent job maintaining professional demeanor during some aggressive challenges...”

There were some negative comments. In particular, some team members felt that the National Institutes facilitators lacked expertise in dealing with GLBT issues.

Reaction to Museum of Tolerance Resources

In addition to the Museum of Tolerance resources provided in the *Experience and Environment* module (i.e., personal tour of the Museum of Tolerance and the Holocaust Museum) and *New Perspectives* module (comprehensive workbook of instructional materials and action plans), the SWC provided other resources to support the National Institutes. These included an e-learning training session that showed participants how to access a virtual network maintained by SWC staff for current and past participants, and they previewed other SWC programming, including the *Perspectives on Profiling* training. The highest ratings were reserved for the resources provided by the Museum of Tolerance, including the e-learning and the profiling exercises (see Exhibit 8-6).

Exhibit 8-6: Participant Reaction to the Museum of Tolerance Resources

M of T Resources	n	Mean Score
E-Learning Exercise		
		Scale 1 (strongly disagree) to 4 (strongly agree)
Feel familiar with website	92	3.60
Think website is useful	92	3.64
Plan to use it	90	3.61
Effectiveness of presenter	90	3.54
Perspectives on Profiling		
		Scale 1 (poor) to 4 (excellent)
Effectiveness of presenter	99	3.90
Content of presentation	99	3.80
Quality of presentation	99	3.90
Meaningful impact on profession	99	3.66

E-Learning Exercise

According to the SWC, a vital component of the National Institutes was follow-up and technical assistance to program participants. The SWC developed the *National Institute Graduate Pages* to track the progress of teams and to provide technical assistance to jurisdictions. This private website was designed to be accessed by former participants through SWC’s website. Participants were given an hour-long session on how to access the site, post information, and send messages. Participant ratings of this e-learning exercise were quite favorable—from 3.54 to 3.60 on a 4.0 scale, addressing issues of familiarity, usefulness, intent to use, and presenter effectiveness. Comments about the training session include:

- “I plan to direct police and sheriffs to the website for further education, [illegible] and profiling.”
- “Mark is great system facilitator”
- “I look forward to using it [the website].”
- “We could use something like a magnet with the website on it. I don’t know what the website is.”
- “Good resource, may have firewall/security issues.”
- “Excellent resource.”
- “It would have been most helpful to have an individual hands-on computer experience.”
- “I thought the presenters were knowledgeable and fluid which made them very effective.”

Follow-up interviews conducted months after participants had returned home from the training, however, indicated that there were a number of problems with the *National Institute Graduate Pages* website. In particular, failure to provide access to the site was a major concern for participants across all three Institutes. In the post-interviews, the majority of participants said they had not received instructions on how to access the database, along with a password. This was a problem because the database was presented as the primary tool for follow-up correspondence with participants from other jurisdictions, and for gaining assistance on jurisdictional plans from SWC staff.

Perspectives on Profiling

In addition to the National Institutes, the SWC provides an array of other training programs to many different groups, including law enforcement. For its *Tools for Tolerance for Law Enforcement* program, the SWC staff developed the *Perspectives on Profiling* training in conjunction with SBE and Will Interactive, Inc. This training, delivered on CD, provided an interactive video experience that addressed the growing need for racial profiling training for law enforcement professionals. In particular, the training addressed the differences between criminal profiling, racial profiling, and racism. It examined the idea that one racial group might be more likely to commit a given type of crime; and it showed law enforcement reactions that range from changing how a case is handled to completely disengaging, out of fear of being accused of bias and profiling.

The video presented dilemmas that law enforcement officers may encounter while performing their daily professional duties. While viewing the video, training participants make critical decisions that can alter the storyline and lead to different outcomes. This training provides participants with the unique ability to see the outcome of their choices and evaluate the consequences. In the interactive exercise, Institute classes viewed a scenario in which a fictitious police department encountered racial problems with the community and faced ethical decisions. The video was paused and participants used hand-held remotes to vote on what they thought the officer involved in the scene should do next. Once the votes were tallied, the video proceeded based on what the majority of the group had selected. In addition, training participants viewed a short film called *Lunch Date*, which features an older White woman in New York City's Grand Central station and her interaction with a Black man who works there. The film shows the emotions that come into play when prejudices cloud how people are viewed.

The *Perspectives on Profiling* interactive CD was a learning tool that was well received by the participants and generated a great deal of discussion among participants. On effectiveness, content, quality, and meaningfulness, National Institutes participants gave the *Perspectives on Profiling* training a rating of 3.66 to 3.90 on a 4.0 scale. Comments included:

- “We saw too little of it to be helpful. I understand the need to charge licensing for its use but wonder if it could be on the Institutes website for us to dig deeper?”
- “She was well organized and a very effective speaker.”
- “Loved the viewer interaction, it really made me think twice about calling someone a racist.”
- “Very good, interesting program. I would have enjoyed the whole program instead of the shorter version.”
- “Wish we could do the whole scenarios—very eye opening.”
- “As non-law enforcement, it had limited relevance.”
- “Very pertinent and makes one think about their actions.”
- “I found that it was similar to ‘preaching to the choir.’”
- “This was my favorite session. Great speaker! I learned some new things about profiling and understand better the complexity and line-drawing confusion.”
- “The video situations were over the top, but I understand their use.”
- “Important discussion.”
- “Need more time for module.”
- “Excellent presenter and presentation.”
- “Police academies need to have this training.”
- “The machine with voting was neat but all of the choices were too obvious.”
- “Dynamic interactive presentation. Some of the best training of the session.”
- “Thought provoking.”
- “The vignettes were excellent, as a prosecutor they appear to be very current and realistic to what goes on in the criminal justice system.”
- “Best I’ve seen in this field.”

Learning/Knowledge Gained

As discussed earlier, the second level of evaluating training programs involves an assessment of learning: What knowledge was gained? Findings were derived from pre-training and reaction interviews, follow-up surveys, and analysis of each team’s strategic plans.

Unfortunately, our sample sizes were small. We obtained responses from 65 individual trainees for the reaction interviews (62 percent of total attendees). We also conducted 10 completed interviews with team leaders in the pre-training phase (out of 16 total team leaders) and received 12 completed survey instruments in the follow-up phase (from a total of 16 team leaders).

When asked about their expectations of the National Institutes training, the majority of participants said they hoped to gain a better understanding of hate crimes and acts of terror. One stated that he hoped to gain “knowledge of the causes and reasons for hate crimes, how to anticipate and prepare for terrorist acts, and to learn the underlying cause for terrorism and what motivates terrorists.” Other reasons for attending the training included a desire to connect with others on their own team, as well as other jurisdictions; discover best practices; learn more about the “balance between the vigilance of a citizen and intrusion of the government;” and “find better ways to write reports and conduct investigations.”

Training participants were asked to define both hate crime and terrorism before attending the National Institutes. Most were able to define hate crimes but struggled to define terrorism. Many stated how their state defines hate crime, such as one participant’s definition of “a crime against an individual because of race, ethnicity, sex, or religious beliefs;” and most knew whether hate crime was a penalty enhancer in their state. Terrorism, on the other hand, proved more difficult to define. Some defined it simply as “an anti-American act,” while others just said “don’t know.” Some defined terrorism in terms of its goals, stating “creation of fear,” “creating panic,” and “the commission of an offense in which the goal is to achieve political results.” Others thought of terrorism as part of a continuum and defined it as “a much more violent expression of hate by a group of people than just hate crime.”

The National Institutes training aimed to change how criminal justice system personnel thought about and approached hate crimes by conveying that it is conceptually different from other crimes. In essence, the training challenged the existing criminal justice paradigm that hate crime offenders do not differ from other types of offenders. The training demonstrated that hate crime offenders vary in their belief that their crime is justified and in their strong commitment to their cause. The training sought to alter or expand participants’ views of hate crimes by demonstrating that acts of hate and terror cannot be either measured or punished like other crimes. To better understand the “values-ethics gap paradigm,” participants worked through a

series of exercises and defined hate crimes as "...a crime against a person (or their property) targeted for their cultural group that he/she *is perceived to belong* to rather than for what he/she individually *believes or does*. The perpetrator feels *morally justified* in committing the crime" [emphasis in original] (Strategic Business Ethics 2004, p 25).²¹

The key difference between hate crimes and other types of crime is that these other crimes are motivated while hate crimes are morally justified. Standard criminal justice system responses such as prison, it is argued, only serve to further the offender's cause because it allows him to cast himself as a martyr in an immoral justice system. This difference requires the criminal justice system to respond differently to crimes of hate and acts of terror than they would to perpetrators of conventional crimes. One way that jurisdictions can respond differently is by recognizing that all acts of hate and terror are serious, regardless of whether the act would be classified as a misdemeanor or felony. The suggested response is to both address the values of the individuals who commit acts of hate or sympathize with those causes, and to develop expanded strategies for dealing with acts of hate and terror.

Follow-up interviews with training participants indicated that while virtually all of them felt the training was exceptional, they were evenly divided between those who felt they hadn't learned anything new and those who felt they had discovered an entirely new way to frame hate crime and acts of terror, through the "values-ethics gap paradigm." Participants routinely pointed to the "values-ethics gap paradigm" as one of the strengths of the training because it provided them with a new perspective on hate crimes and linked hate crime with terrorism. Most attending the training had not previously connected hate crimes and acts of terror. Thus, the paradigm shift that the training sought to cause appears to have been successful for some participants. Other participants didn't feel it was something "new" but did find it a strength of the training and others disregarded it all together. As one participant stated, "I wrestled with the new theory they presented on motivation of hate crime defendants although I don't buy it... [it] works in terrorism cases but not in hate crimes cases."

In addition, when asked what they took away from the training, most participants mentioned the museum experience as one of the most memorable events. It was described as

²¹ Strategic Business Ethics created and now teaches the National Institutes Against Hate Crimes and Terrorism training curriculum. Its copyrighted materials are included in the binder distributed to all participants.

“mind-boggling,” “powerful,” and “educational.” Many interviewees stated that after the training they felt “empowered” and had a “renewed commitment” to fighting acts of hate and terror. Interviewees also said the speakers made the training more memorable. Dr. Roberts, an original member of the Little Rock Nine who spoke about his experiences and the impact of school desegregation on his life, was described as “incredible, astounding.” Both the holocaust and Rwandan genocide survivors were favorites, as was the terrorism speaker. The racial profiling exercise was a frequent favorite, particularly from those in law enforcement. The training was an intellectually stimulating experience for the participants, as indicated by these two comments:

- “Exercises with people across the country gave me a deeper understanding of how people see things... It was mind expanding.”
- “There wasn’t one day where I didn’t learn. I learned a lot about history and the effects of hate. I learned a different way of thinking.”

Overall, the museum touched a chord with people across disciplines and jurisdictions and gave them a sense of renewal, which is evident from the following comments:

- “Learned to be more sensitive to the issues”
- “New breath to come back and approach my job”
- “Even small steps can make a difference.”
- “Gave energy to want to go out and do something.”
- “Personal check on myself and my own prejudices.”
- “The power of words is tremendous.”
- “The training was inspirational, informative, and educational. I’m not as reactive now. Before I had blinders on. I learned I need to listen to each case individually before I make a decision.”

In the follow-up interviews, participants stated that they enjoyed the overall experience and liked the real-life examples they were given. For example, one officer stated that he liked the “practical application to the program.” Another respondent liked the fact that the SWC training brought “multiple voices to represent and support civil liberties.” The training energized people to want to do more and group members felt compelled to go back to their jurisdiction to share what they learned. In the follow-up interviews, the team from Madison, WI, and the team from Monmouth County, NJ, both stated that they were sending another team from their jurisdictions.

Attitude and Behavior Changes

The third level of evaluating the SWC training program was assessing behavior change: How have the participants' attitudes and behaviors changed as a result of the training? Findings for this evaluation level were derived from pre-training and reaction interviews, follow-up interviews, and analysis of each team's strategic plans.

Team leaders were asked why they wanted to attend the training and if there was a specific incident or event that prompted them to get involved in the training. The majority noted that there seemed to be an increase in hate crime incidents since September 11, 2001. Most noted racially-motivated hate crimes as a problem in their jurisdictions. The most common incidents cited included graffiti on synagogues and other places of worship; slurs, name-calling, and yelling; harassing telephone calls at Islamic organizations; mailing of anti-Semitic letters; residential property damage; and cross-burnings. One interviewee noted that in his city "there's a lot of racial tension," which breeds "a whole bunch of hate." Others noted that bias-related incidents were occurring in schools with greater frequency. Other reasons for attending included the formation of new task forces, such as anti-terrorism task forces in Wisconsin, Michigan, and California, and hate crimes task forces in Florida, Wisconsin, Michigan, and New Jersey.

After attending the training, some participants discovered that their jurisdictions were either "on-track" or were "maladaptive" in dealing with crimes of hate and terror. As one participant from West Virginia stated, "We have not been equipped to deal with these issues because they're so political. We need to deal with the hard issues a little better." A New Jersey judge stated that he is "now more knowledgeable in sentencing and in dealing with cases" after attending the training. A prosecutor from Wisconsin stated that the training forced her to look more closely at the motivation behind cases she works on. She stated, "When I get a case and read the police report, I read and see if it's bias-related. I look at everything more closely." Another prosecutor said that because of the training, he has "...more sympathy in plea bargains."

The motivation for attending the National Institutes training for the teams from New Jersey was to learn more about hate crimes and acts of terror, to find out what others around the country were doing, and to discover innovations in dealing with these crimes and incidents. The main impact the training had on these teams was to help them recognize that what they were doing was "on the right track." The lead bias crimes detective and others from the Monmouth

County Prosecutor's Office discovered they were at the forefront in dealing with bias crimes and diversity issues. The teams found that their programs were more developed than they had realized and that other teams wanted to emulate the Monmouth County programs.

Importantly, what these teams gained from the training was a renewed confidence in their strategies for combating acts of hate and terror, and a renewed enthusiasm from the top administrators on down for their current programs. When asked what the biggest change was since members attended the training, the chief of detectives replied that although the prosecutor's office takes pride in its accomplishments and knows other units will call on them for assistance, they can only get better. The team that attended the SWC training heard different views from around the country, and it reinforced their drive and commitment to their mission.

Several teams from Wisconsin felt that collecting data on hate crime was a problem in their state, and there was some fear that law enforcement may not be properly classifying these crimes. When no hate crimes were reported in 2003 in some of the major jurisdictions, there was some serious questioning as to whether hate crimes had actually never occurred that year. For these teams, the central questions were: "How come people are not reporting hate crimes? Or, are they reporting and law enforcement fails to recognize it?" The training prompted team members to research the problem. When they returned, one of the team members queried the police database for the last seven years to see how rampant an issue hate crime was. The team member discovered approximately ten hate crimes in those records.

Another Wisconsin teammate educated her staff on the issues and began collecting data on people under corrections supervision who were involved in a hate crime. Other participants attended community forums to educate the public on hate-related issues, and a number of participants got involved in their city and state hate crimes task forces. Some participants saw the National Institutes training as beneficial because it increased awareness of hate crime by the judiciary. As one assistant district attorney stated, "the training makes it easier to take these cases to court if the judges have also been trained."

When asked how the training changed their roles and duties, over 50 percent of the team leader respondents stated that it made them much more aware and sensitive to ethical and racial issues. These respondents stated that they are now more tolerant and much more likely to "speak up" when racial jokes or other biases are evident. However, nearly 50 percent of the

respondents, while praising the training, stated that it has not affected their day-to-day jobs. As one prosecutor stated, “When we get back to our jobs, it’s hard to incorporate [the training] when we’re so busy.”

The majority of team leader participants (over 60 percent) noted that they are now more involved in addressing hate crime issues in their jurisdictions and have vowed to educate both their peers and residents in the community. A hate crimes coordinator from one jurisdiction revamped his whole training curriculum after attending the training at SWC. A hate crimes trainer from another city implemented some of the Institute training into her program. Another team stated, “We think more deeply. We take more opportunities to reach out to the community and create alliances and build bridges. We may have done this if we didn’t go to the Simon Wiesenthal Center [training], but probably not.” Many participants stated that they shared their experiences with their colleagues after returning from the training. Some made reports to their superiors and shared some of the materials obtained at the training. Attendees stated that they look more closely at cases they work on to see if there is a hate crime connection.

Strategic Theory and Action Plans

Important aspects of the National Institutes training were creation of a strategic theory guiding the entire Institute and creation of strategic action plans for each jurisdiction represented. The strategic theory was decided upon by the entire group of National Institutes participants, while the strategic action plans were individualized to each team. Each is discussed below.

Strategic Theory

A portion of the training focused on developing a “strategic theory” as a way to focus and guide groups as they devise their strategic action plans. The strategic theory is similar to a mission statement in that it channels team activities in a more direct fashion. According to SBE, Inc., the strategic theory is a “blueprint for collaborative action” that is “designed to propel change in the chosen strategic direction” (SBE, Inc. 2004, p. 29). The strategic theory is touted as a “model for change” (SBE, Inc. 2004, p. 29) and uses the special talents and qualities of team members to guide activities.

Each Institute formulates a strategic theory for all of the participants in the Institute. There is one driving theory for all of the groups. In developing the strategic theory, the entire

group (i.e., all the jurisdictions) must decide on a single target on which to focus. It is expected that the strategic action plans will be a by-product of this theory. The National Institutes training facilitators give participants four target areas from which the group is to choose one as the target of their strategic theory. The four target areas are:

- **Prevention** targeting school children and parents
- **Remediation** targeting people within the criminal justice system
- **Collaboration** targeting colleagues
- **Collaboration** targeting members of the public.

The strategic theory for Institute 17 was:

We will ignite the energy of our colleagues in the criminal justice system to respond to forces of hate and terror that threaten our safety, freedom, and quality of life. Using technology and knowledge, we will build an educational coalition with vast reach that will protect our constitutional values and inspire trust in the communities we serve.

The strategic theory for Institute #18 was:

We will create effective teams as a collaborative force to teach children and their parents the value of a just society that respects cultural differences. Using our professional credibility and existing channels, programs and resources to create new insights into the causes of hate, its impact and how to prevent it, we will secure a safe tomorrow.

The strategic theory for Institute #19 stated:

In an atmosphere of increased homeland vigilance, we will use our network of government contacts and our relationship with the media to proactively educate the public into new mindsets which expose the links between bias, hate and terror as the enemies of our values of freedom and justice.

Creation of the strategic theory was challenging for many participants. In Institute 17, for instance, there was some disagreement over the focus and purpose of the strategic theory, which led to lengthy discussions between the teams and the SBE facilitators. In addition, some teams from each of the trainings complained about being “forced” into a strategic theory not of their choosing, since the facilitators encourage group consensus across all jurisdictions on the target of focus. This approach, however, may be too simplistic as different jurisdictions face different

problems. Some other participants noted that they felt they had to go along with the group even if they did not agree with the focus.

Since the strategic theory was considered a “blueprint for collaborative action,” post-training interviews asked participants if they could recall the strategic theory and what it said. Over 75 percent of participants across the three institutes could not recall the strategic theory. According to interviewees, the exercise did not address the needs of participants and took valuable time away from developing the strategic action plan, which is a major component of the training and was used by SWC to measure change and evaluate the effectiveness of the National Institutes training.

Not one participant interviewed post-training considered the strategic theory as a productive and worthwhile task. Due to the amount of time spent on developing the theory, the time spent on devising the action plan was rushed, according to many participants. Given the importance of and emphasis on going back to one’s community and implementing an action plan, most participants would have preferred more time devoted to the strategic action plan exercise.

Strategic Action Plans

As explained earlier, the National Institutes training recognizes the need to customize its training to be effective. It taught a set curriculum on hate crime but individualized the training for each attending team, allowing the teams to develop their own strategic action plans based on collaborative strategies learned in the training. The plan can focus on a variety of targets, including school children or parents, offenders, law enforcement and colleagues within the criminal justice system, and the general public.

The strategic action plans for most teams involved educating either criminal justice system personnel on the new paradigm of hate crime and terrorism learned during the National Institutes training, or focusing on prevention efforts with school children (Exhibit 8-6 shows a chart of all the strategic action plans for each jurisdiction in each Institute). Teams from Alabama, Idaho, West Virginia, Wisconsin, California, and New Jersey developed strategic action plans for training colleagues about hate crime and terrorism. Teams from Montana, Pennsylvania, California, Florida, Michigan, Ohio, and New Jersey developed strategic action plans targeting students and the larger community for education efforts on recognizing and

preventing hate crime. For example, a team from New Jersey established a blueprint for recognizing that acts of terror and hate are similarly motivated. This plan involved connecting the Bias Crime and Community Relations unit with the Homeland Security unit within the prosecutor’s office—to support community education programs and promote media outreach. A team from Florida decided to create a task force composed of school board members, elected commissioners, the state attorney, public defender, chief judge, and community leaders to create, endorse, and participate in a designated hate crimes awareness day within the schools. Likewise, teams from California, Montana, and Ohio had the primary objective of creating a “National Day of Tolerance” to coincide with September 11th remembrances.

Exhibit 8-6: Strategic Action Plans by Institute Number and Jurisdiction

TEAM	OBJECTIVES	ACTIVITES	BY WHEN
Institute 17			
Alabama	Initiate education on hate crimes and related issues in our various disciplines states conferences Initiate similar education at statewide police academies and by law enforcement coordinating committees of our three federal judicial districts Develop public service announcements to raise public awareness of hate crimes and their issues Review progress within 90 days Initiate inclusion of agenda at state conference	Individuals in each discipline will get presentations placed on conference agendas Law Enforcement Coordinator, Attorney General’s Office Law Enforcement Coordinator, Attorney General’s Office Law Enforcement Coordinator, Attorney General’s Office	End of calendar year
Idaho	Address the various disciplines within our law enforcement community with new presentations on hate crimes and terrorism Train trainers in these theories Reach out and inspire trust in communities in particular those at risk for alienation Continue meeting with team	Present to Ada County Probation Office, Boise Police Dept., Ada County SD, P&P PD office statewide training Obtain and use grant money to receive training Maintain relationships with minority groups Meet quarterly	Summer 2004 January 2005 On-going On-going
Montana	Educational meeting in our respective communities Reach and talk with students	Contact high school government class	By Fall 2004 Spring 2003-2004 school year

TEAM	OBJECTIVES	ACTIVITES	BY WHEN
	Interactive half-day conference with Little Rock Nine Rep, Aryan Nations, possibly Native American rep, Billings Memorial rep, Holocaust survivor	Liaison with MoT Liaison with Technology Coordinate e-mail and messages	September 11 or May 1
Pennsylvania	Invite more agencies to participate in the SPIRIT student dialogue program	We will make specific invitation to other agencies in strategic areas of the state to join us in this endeavor	End of March 2004
	Train facilitators from across the state	Establish dates and locations when US Justice will conduct training for facilitators	End of April 2004
	Develop criteria and then use it to develop a list of "target" schools	Gather input from US and PA Departments of Education and from PHRC database	End of March 2004
	Plan pilot program for four schools	Contact school administrators and community leaders for their agreement and participation in the four pilot programs	Hold two by the end of this school year and two more by November 2005
	Combine community seminar with SPIRIT school program	Hold community seminar to address area concerns related to changing demographics, inter-group tensions, civil rights issues, PA's hate crimes law, hate group activity, and concerns regarding terrorism	Hold two by the end of this school year and two more by November 2005
	Develop follow-up resource materials, program recommendations, and visits by contacts with participating agencies	Draft follow-up plan and resource list	End of March 2004
West Virginia	Create law enforcement training with objectives throughout state Education process Re-evaluate current process Outreach	Contact academy for support Use tools from home to train Brainstorm current producers: civic groups, churches, employers, Prosecuting Attorney Association, corrections academy, chief parole services	Update every other month Target goal 12-18 months completion
Wisconsin	Connect with past state institute participants Better education among police, DA's PD's and judges about what is a hate crime Educate specific protected groups, probation officers and community groups	Convene meeting Police: in-service and training academy PD's: regular training sessions Probation: bring in a	April 2004 December 2004 September 2004 May 2004

TEAM	OBJECTIVES	ACTIVITES	BY WHEN
	<p>Improve hate crimes data collection</p>	<p>management team and train DA's" review statute Community: forum for donors, funders, and community groups Judges: email judges about hate crimes and discuss it at felony, misdemeanor and children's court meetings</p> <p>By Judges: email judges on hate crimes and request if they've dealt with any listed as such or that they suspected By community groups: distribute internal audit which lists criminal activity for trends to other groups By police: to check with computer personnel to see if there is a special category for hate crimes and if not, could there be</p>	<p>May 2004 June 2004 April 2004</p>
Institute 18			
California	<p>Create a "National Day of Tolerance" A community-based, not criminal justice-based response</p>	<p>Develop programs & curriculum Support of elected officials, educators, community , non-gov. orgs and leverage currently existing programs No funding but: School – Policing partnership grants Utilize PTA/Local community resources already in place, business leader</p>	<p>9/11/04 On going</p>

TEAM	OBJECTIVES	ACTIVITES	BY WHEN
<p align="center">California Department of Justice</p>	<p>KIDS TOLERANCE QUIZ</p> <ul style="list-style-type: none"> - Interactive through Internet Web-site - Tools for outreach/community meetings /curriculum <ul style="list-style-type: none"> a. Define hate crime b. Cultural awareness / diversity c. Warning signs <p>Begin 5th & 6th Grade California Students?</p> <ol style="list-style-type: none"> 1. Interactive web-site <ul style="list-style-type: none"> - help identify what a hate crime is - help explain acceptance & diversity - what to do if they witness beginning of a hate crime (for parents) 2. Create a Promotional Day 3. CA Student Surveys <ul style="list-style-type: none"> - Survey 7th Grade - 5th & 6th Grades 10-11 year olds 4. AG's "Kid's Corner"—Interactive tool for kids/teachers/parents <ul style="list-style-type: none"> - Involve kids in production of materials - Tolerance/diversity – explain it - Things you can't change - What should you do 5. Safestate.org – print cards w/ web-site on them 6. Safe School Plans – sneak into their plan 7. Parent Teacher Guide. – Exploiting existing meetings 8. Community Law Enforcement Grants – Give them \$ day to think @ hate crimes 	<ol style="list-style-type: none"> 1. Regional Trainings – CD Education <p>“safe school planning”</p> <ol style="list-style-type: none"> 2. School community policing grants Support through proclamation City, elected officials <ol style="list-style-type: none"> 3. Safe School Plan - use of vehicle to promote the interactive web-site 4. Parent Teacher Guide – Civil Rights (Safe from the Start (Project 10)) <ul style="list-style-type: none"> -need ways other than interactive to get message/info out. -hate crime moot court 5. School Safety Policy Partnership grant. <p>HATE CRIME moot court – CRIM/CIVIL RIGHTS/CATIC CATIC</p> <p>AG/LA Group Cards</p> <p>Safe School Play</p> <p>Parent Teacher Guide</p> <p>School Community Policing</p> <p>Partnership Grant</p> <p>Supporting Data</p> <p>Web-site containing material</p> <p>Advertisement for materials</p>	<p>January 2005</p>
<p align="center">Florida</p>	<p>No group that has dedicated forum for dealing with hate crimes</p> <hr/> <p>Create a task force consisting of a School Board member, Elected Commissioner, State Attorney, Public Defender, Chief Judge and Community Leaders or designees to endorse and participate in creation of a day within School year to bring awareness to hate crimes</p>	<p>Approach these members to initiate idea:</p> <ul style="list-style-type: none"> - School Board member - Elected Commissioner - State Attorney - Public Defender - Chief Judge - Community Leaders <p>By letter req. resp.</p>	<p>Create Task force by 6/16/04</p>

TEAM	OBJECTIVES	ACTIVITES	BY WHEN
Michigan	<p>Problem:</p> <ul style="list-style-type: none"> - Schools addressing problem of bias motivated (hate) crimes and incidents on their own (if at all) - Existing resources go unused because schools don't know what's available - Individual schools don't believe THEY have access to gov. agencies etc. That go to "bigger" schools <p>Change:</p> <p>Reach out to teachers/educators/schools safety officers & administrators on statewide basis in order to create network and provide for education of participants, dissemination of existing programs & resources and the sharing of ideas</p>	<p>Regional Conferences (at least 2)</p> <p>Designed to:</p> <ul style="list-style-type: none"> - Educate/train participants - Identify interested schools & individuals <p>\$\$: Possible Justice Dept. grant (MIAAHC members) and/or partnership (or grant) with State Bar Access to Justice Program.</p>	<p>Conference design and detailed time line by May 1, 2004</p> <p>1st conference 2005 (though 11/2004 if possible)</p>
Ohio	<p>Establish an effective educational awareness program</p> <p>May Day incident (background hate crimes committed by African American & Hispanic youth)</p>	<p>Go into Cleveland Municipal schools "Day of Understanding"</p> <p>Judges Town Hall Meetings</p> <p>Moot Court</p> <p>Hate Crime Fact Pattern</p>	<p>April 2004</p> <p>April – June 2004</p> <p>April 2005</p>
Pennsylvania	<p>Identify agencies & organizations working in anti-hate, pro-tolerance in the public school systems</p> <p>Seek to have the work of various task forces coordinated and expanded to include relevant criminal justice organizations (i.e. courts, public defender that have been part of such organizations)</p> <p>Evaluate and expand the efforts of all groups and develop a unified program to present in all middle schools in the city</p>	<p>Information gathering from already existing groups</p> <p>Go back and talk to heads of our organizations and encourage our participation in conjunction with agencies already working on existing task force</p> <p>Get everyone at a table and develop programs we want to present</p>	<p>3 months</p> <p>3 months</p> <p>5 – 6 months (Beginning of new school year)</p>
Institute 19			
California	<p>Coordinate existing anti-bias agencies so that the community can access anti-bias training and education</p> <p>Prepare an informational flyer with an anti-bias/anti-terrorism connection and list of agencies willing to speak to community groups</p>	<ol style="list-style-type: none"> 1. Contact HRC with idea 2. Get list of existing agencies 3. Make presentation at HRC 4. Prepare solicitation letter for agency participation 5. Design flyer 6. Obtain distribution lists 7. Obtain funding 8. Media event 9. Distribution of flyer 	<p>8/1/04</p> <p>8/1/04</p> <p>8/1/04</p> <p>11/1/04</p>

TEAM	OBJECTIVES	ACTIVITES	BY WHEN
New Jersey	<p>Connect bias crimes and community relations with Homeland Security whereby connecting the bias crime with terrorism. This must be done by being sold to administration</p> <p>A. Get permission of Prosecutor</p> <p>B. Conduct programs – Bias/Homeland</p> <p>C. Create a program that will educate public</p> <p>D. Contact media – Announce Initiative</p> <p>E. Continue with program, facilitate</p> <p>F. Setup county wide task force to include representatives from: Schools, HRC, YSC, Religious leaders, Police Academy, Community Colleges, Corporate Sponsor</p>	<p>Describe new initiative, promote, educate, and sell it to the media. Open invitations, send info, buy ads, P.S.A</p> <p>1. Develop Curriculum – this can be modified</p> <p>2. Train the Trainer</p> <p>3. Implement Program</p>	<p>ASAP by July 1st.</p>
Wisconsin	<p>1. Organizational plan must be developed and/or fine tuned.</p> <p>2. Contact “principal players”; government officials, law enforcement, judiciary, Dept. of Corrections, school and the university and bring them together. <u>AND</u> convince “principle players” to commit someone, if it cannot be themselves to work on implementing our strategy.</p> <p>3. Work to ensure the ongoing funding via grants. Contact school districts. Clarify the need for a response. Distinguish between seeking funding for homeland security, and sounding the call to action about hate crimes in our community.</p>	<p>Increase contact with and communication between all individuals who have previously attended the Institute and those who may do so in the future, and individuals who have been brought on board already.</p> <p>By contacting, and recontacting the people who have already contributed to prior efforts, and focusing on bringing these officials on board through <u>our</u> agencies.</p> <p>Obtain models from other jurisdictions, that have been successful already, to guide us and help us understand how to get a core group dedicated ultimately to education about hate crimes.</p>	<p>8/1/04</p> <p>9/1/04</p>

Post-training interviews asked participants to outline the main components of their groups’ strategic action plans, key tasks, and members responsible for task completion. The follow-up surveys specifically addressed the completion status of the strategic action plans. The interviews were conducted two to six months after training, and the surveys were sent out 12 to 16 months following the training. Responses showed that most teams were not able to fully implement their plans during the evaluation’s follow-up period; however, they were still planning to fulfill their action plan mandate. For example, a team from the California Attorney General’s Office succeeded in getting approval for one aspect of its project, creation of a

children’s website on diversity. The team members were planning PowerPoint presentations on their training experiences for members of their office and were seeking to join the Attorney General’s Hate Crime Work Group. In addition, a team from West Virginia held a forum on hate crimes at an area college and was in the process of coordinating with a Little Rock Nine activist, Dr. Roberts, to be a guest speaker at another conference.

While the follow-up responses indicated that most teams had not fully implemented their SWC strategic action plans, each team leader interviewed indicated that the action items were still applicable to their plans, although the timelines had been altered greatly. Difficulty in communication among team members appears to be a key reason why some action plans had not been implemented in a timely way. Many interviewees reported that their teams had not met in the six months following the training, though most had communicated via email or telephone. Others indicated that some members of their team had met, but not the entire team. For instance, one public defender stated, “[I] haven’t seen the judge or the probation officer again.”

Time and distance are also a problem. In Montana, for example, some of the team members live eight hours from one another, making team meetings difficult to impossible. Other team members simply are consumed in completing their day-to-day tasks with little time to devote to their strategic action plan. As one public defender stated, “I’m so busy, so I’m not sure if I will make any of the face-to-face meetings.” Other reasons for slow progress included the following:

- Roles not clearly defined
- Lack of vested interest from all team members
- Correct personnel not a part of the team
- Resistance from supervisors in implementing change

Another obstacle to the successful development and subsequent implementation of the strategic action plan was that many team members did not know each other before the training. In fact, one group leader stated just four days prior to the training, “I don’t even know who is coming.” Teams have also faced obstacles in creating task forces. As one participant from West Virginia stated, “the hate crimes task force has no power, no money, and no enforcement power... law enforcement needs to get more involved.” One team member from Pennsylvania stated that “it would be nice if we actually do some things we said we were going to do. . . .The

plan we came up with was rejected [by the facilitators], and I thought that was too heavy handed. I think that's one of the reasons it's not getting off the ground." Others noted that "there's a lot of bureaucracy in my office, so they're resistant to change." While there is no doubt that the training provided by Simon Wiesenthal Center Museum of Tolerance is an excellent product, the comments by these team leaders provide an example of the sometimes poor fit between those who attend a training and those who will stay committed and work toward change once they return to their home city.

Organizational Impact

The fourth level of evaluating the SWC training program was assessing the impact of the training on the attendees' organizations. The central question for this evaluation was: What effect has the training had on strategies, policies, or programs in the participants' organizations for combating hate and terror crimes? Findings were derived from the case studies of Monmouth County, New Jersey, and Madison, Wisconsin.

The teams from New Jersey²² had been particularly active in combating hate crimes including developing the following:

- Created innovative programs including a bias unit, specialized police academy training, and a juvenile offenders program;
- Recognized that any bias crime can be both destructive and serious in its impact on the targeted community; and
- Incorporated a team approach by forming relationships across the county with other law enforcement officials, prosecutors, human relations commissioners, religious leaders, members of the gay and lesbian community, and leaders in minority communities.

In addition, the police officers on the teams from New Jersey are members of a County Bias Officers' Association. After returning from the training, the Association decided to implement a training program for all officers in the association. In August 2004, 50 officers from Monmouth County, New Jersey, received hate crimes and terrorism training that included a visit to the Holocaust Museum in Washington, D.C. In December 2004, another group of officers received training, which included lectures by the Southern Poverty Law Center.

²² See Appendix 8-B for case study reports on the New Jersey and Wisconsin teams.

As a result of the National Institutes training, a group of teams from Wisconsin²³ had:

- Formed a county hate crimes task force.
- Created and distributed 3,000 copies of a hate crimes brochure, also translated into Spanish.
- Given a presentation to the police management team asking for additional training and stressing that all hate crimes, however minor, need to be taken seriously.
- Trained 13 detectives who volunteered to be trained on hate crimes and bias incidents and now serve as resources for hate crime investigations.
- Formed a Department of Corrections hate crimes committee spearheaded by a member of the department who had attended the National Institutes training. The committee, which meets monthly, has sponsored a number of outreach efforts, including a 30-minute presentation followed by a call-in session on a new local Spanish radio station. In addition, the committee sponsored regional training for probation and parole officers and was in the process of establishing mandatory training for new agents.

Discussion

The National Institutes training provides a thorough examination of the outcomes of hatred. The training is highly regarded and energizes participants to continue their efforts in combating crimes of hate and terror in their own jurisdictions. Participant feedback about the training was overwhelmingly positive, and many participants felt they had gained new knowledge and skills at the training. Importantly, a number of participants felt the training empowered them to develop specific goals, objectives, and strategies for targeting hate crimes within their own communities.

While the training participants highlighted many positive aspects of the training, recommendations were offered on areas that could be improved. Many participants identified a number of areas, both major and minor. The major areas recommended for improvement included:

- Make a clearer connection between hate crime and terrorism.
- Implement the follow-up/access to the Simon Wiesenthal Center tracking database.
- Provide technical assistance to help jurisdictions complete their strategic action plan.

²³ This group of Wisconsin participants includes five teams that had attended the training between November 2000 and the Institutes followed in this evaluation.

The minor items recommended for change included:

- Reconsider the length of the training (discussed below).
- Rethink use of the strategic theory approach.

The follow-up interviews revealed that participants perceived a disconnect between hate crime and terrorism in the overall training. Many respondents viewed the two issues as separate topics and suggested dropping terrorism in future trainings. Participants noted that the concepts of hate crime and terrorism are “incompatible” and should have been covered separately in different training sessions. Many either failed to see the connection, or did not believe that the terrorism topics were covered in enough depth to warrant their inclusion in the training. Others believed that the focus on terrorism was impractical and distracting from the focus on hate crime. For example, one person said there was a “poor fit between terrorism and hate crime,” while another stated, “I don’t feel I could come home and do something [about terrorism].”

Team selection and screening was an area of concern for National Institutes participants. SWC aims to create teams that will work well together on their strategic action plans and be effective in making changes in their home jurisdictions. They also encourage diversity by asking jurisdictions to form teams that include a representative from the police, prosecution, public defender, probation, the judiciary, and education or human relations commission. There was some question among participants about the rationale behind the team selection. More than one participant agreed with an interviewee’s sentiment that “[teams] need to be more carefully selected.” Others argued that people in like-minded fields should be on the teams—for example, a team of police officers only. This, they believed, would ensure greater focus and better communication among team members after the training.

Another comment about team composition by one participant was that “more attention could be paid to creating teams that could do what SWC wants. Teams should not be selected by just who can come. It would be a good idea to invite people that can make changes, such as legislators.” It was argued that teams should be screened to include people who actually work on hate crimes, so that when they return to their jurisdiction, plans would be implemented with less difficulty. In the words of one participant,

“We have an opportunity to really address hate crimes and terrorism.
Need to assist each other in doing some critical analysis on how to be

more effective. Need to identify solid team members. Law enforcement people have a lot of difficulties working with non-law enforcement. Everyone should engage. There should be more effort to teach trainees on how to share information with people not there. Need to actualize it and make it practical. The selection of the team is so important. It is so important to get the right people there.”

While the training provided by SWC was intellectually stimulating and professionally delivered, it failed to provide practical applications and tools, according to some National Institutes participants. Some argued that the training was more theoretical than practical, and as such, lacked solutions. One participant stated, “We were not given practical problem solving solutions. The information was simply presented but [the] group was not told how to apply it when they got back to their jobs.” Many participants were looking for something more concrete that they could take back home with them.

At the training, keeping in touch and connected was emphasized. SWC staff designed a website specifically for Institute participants so they could discuss challenges, issues, and milestones with their colleagues across the country. The website also was designed to allow SWC to track the progress of team activities. In the follow-up interviews, participants were asked if they received access to the tracking database, including a password and instructions on how to use the website. At the time of the interviews, six months post-training, none of the participants had received information on the website, but all expressed a desire to use it. The website was viewed as a “tool” by Institute participants, and they believed the database could help with the success of their jurisdictional plans as a method to keep connected with other teams and gain assistance. Many participants felt that it would have been useful in implementing strategies against hate crime in their jurisdictions.

There were also some concerns about the overall length of the training. A lot of information was packed into the three-and-a-half day session, and many participants found it overwhelming and emotionally draining. The lengthy days tired them, making it difficult to process all the information presented. For example, as one participant stated, “You reach a level of saturation.” Another made the following comment in reference to a particularly long day, which included the tour of the museum:

“Mentally, physically . . . everything, rung out. I couldn’t debrief. I had to control my emotions. I went back to my room and the tears came. We

get overloaded. I would be willing to stay an extra day. They need to separate the tolerance and the holocaust and let people have time to get out. That was a hard day.”

Similarly, another participant stated that “they need more time, as the material was too compressed;” and in the follow-up interviews, many participants agreed that the training could be restructured with an extra day at the end to accommodate shorter days of actual training. The following comment exemplifies the sentiments expressed by many of the participants.

“There was so much information and we needed more time. I left some days feeling depressed with no outlet. The training could be a day or two longer so that we have time to recuperate and refresh. There needs to be a transition period because it’s so traumatic. Even the last day was heavy with the skinhead testimony.”

It was also suggested that less time be spent on the strategic action theory exercise vs. the strategic plan exercise. For example, one participant stated that the strategic theory exercise “dragged too long. The ideas were interesting but could have been covered in three hours. It got oppressive and did not seem to move.” In the follow-up interviews, over 75 percent of participants could not recall the strategic theory. Generally, participants did not consider the exercise to be useful, and many felt they were forced to focus on an issue that was not a concern in their jurisdiction.

In following up with participants at the three National Institutes that were part of this evaluation, it was evident that some teams remained energized and continued to work toward their goals, while others quickly faltered. A key question was: Why have some teams continued to meet and others not? The case study of Madison, Wisconsin, was instrumental in answering this question. Madison had sent five teams since the inception of the National Institutes in 1999. Some of these teams remained active and others did not. Lessons can be learned from these five Madison teams. First and foremost, teams seem to thrive or falter based on leadership. A member of the second team of Madison criminal justice personnel to attend the training (in December 2001) stated that the team met for the first year after training but is no longer meeting because of a lack of team leadership. As is common with many of the teams from across the country that have attended the National Institutes, one or two individuals shoulder the leadership responsibilities. When these individuals retire, change jobs, or become too busy to continue to bear the leadership duties, the teams tend to splinter and stop meeting.

A second lesson from the Madison teams concerns collaboration and dedication. Many team members stressed the importance of collaboration. In Madison, this collaboration has taken the form of task force meetings. Interviewees identified two previous iterations of the current hate crimes task force, each of which eventually disbanded because of a lack of leadership. The issue for many team members was not a lack of commitment to the work of the task force, but finding the time in addition to their regular job duties to carry out this work. As one detective stated, “there is no formal structure to maintain these teams as part of your formal job.” Team members expressed frustration at the lack of time they have to devote to team meetings and their strategic action plans once they return to the office. In the words of one member of the Dane County District Attorney’s Office, “it is very difficult, especially working in the D.A.’s Office, to take extra time to do the strategic action plan tasks, which are essentially done off hours and not as part of our traditional job.”

Various team members suggested solutions for dealing with leadership issues and sustaining the teams after training. These included:

- Designate a leader. Have SWC require the identification of leaders as part of the strategic action plan.
- Gain support of top management. Obtain a commitment from agency department heads to support the action plan.
- Allow training materials to be shared with colleagues who did not attend the training, or provide tools so attendees can share what they know.
- Require firm due dates for the completion of tasks on the strategic action plan by SWC.
- Improve the National Institute website and make it more user friendly.

In conclusion, the SWC National Institutes training is widely respected by participants as one of the best trainings on hate crimes and terrorism funded by BJA. Virtually all participants consider the training a very positive experience and find the combination of workshops, guest speakers, and tours of the Museums of Tolerance and Holocaust to be particularly moving. As with any training, there is always room for improvement. In particular, the training provided by the SWC could be strengthened by the following:

- Allow more time for team members to develop their strategic action plan, perhaps by beginning the process before the training starts (e.g., by requiring team members to meet).

- Alter the strategic action theory component of the training and allow more time on the strategic action plans. Do not force every team attending the training to dedicate its strategic action plan to the same strategic action topic.
- Give participants time to sit down with other participants in the same field to discuss what is and is not working in their jurisdictions. This could be done by moving the reception to a night later in the training, after participants have had a chance to get to know each other.
- Provide tools, such as materials, CDs, or websites that participants can access and use to teach the training to colleagues back home. Some participants were frustrated that they were expected to teach the training without being allowed to use the copyrighted materials.
- Employ some type of selection process when choosing participants to attend the training to ensure that team members have a commitment to address hate crimes, experience in this area, and support from their top management.

APPENDIX 8-A

SWC Training Evaluation Materials

Student Course Evaluations (reactions to training—one each day)

Pre-training Participant Interview Questions

Post-training Participant Interview Questions

Follow-up Survey

Day Two Evaluation

Your feedback is extremely important. Please rate and comment on the following:

1. Public Safety vs. Individual Rights
(Constitutional Lawyer Stephen F. Rohde)

1
poor

2

3

4
excellent

Comments:

The SBE workshop will be evaluated at the end of day three.

Thank you!

Day Three Evaluation (con't)

Your feedback is extremely important. Please rate and comment on the following:

1. Current Issues Around Terrorism (Dr. Sabi Shabtai)

Extent this presentation was helpful in raising your awareness and understanding of terrorism

1	2	3	4
poor			excellent

Comments: _____

2. Relevance of the presentation to hate crimes and terrorism:

1	2	3	4
poor			excellent

Comments: _____

3. Relevance of the presentation to your professional work:

1	2	3	4
poor			excellent

Comments: _____

4. Effectiveness of Presenter:

1	2	3	4
poor			excellent

Comments: _____

Thank you for your assistance! Simon Wiesenthal Center Museum of Tolerance Tools For Tolerance® National Institutes Against Hate Crimes And Terrorism

Day Four Evaluation
Technical Assistance

Perspectives on Profiling (Director Sunny Lee)

- | | | | | |
|--|---|---|-----------|--|
| a. Effectiveness of presenter | | | | |
| 1 | 2 | 3 | 4 | |
| poor | | | excellent | |
| | | | | |
| b. The content of the presentation | | | | |
| 1 | 2 | 3 | 4 | |
| poor | | | excellent | |
| | | | | |
| c. The quality of the presentation | | | | |
| 1 | 2 | 3 | 4 | |
| poor | | | excellent | |
| | | | | |
| d. Its capacity to have a meaningful impact on my profession | | | | |
| 1 | 2 | 3 | 4 | |
| poor | | | excellent | |

Comments: _____

Personal Testimony (Former Skinhead Tim Zaal)

- | | | | | |
|--|---|---|-----------|--|
| e. Effectiveness of presenter | | | | |
| 1 | 2 | 3 | 4 | |
| poor | | | excellent | |
| | | | | |
| f. The content of the presentation | | | | |
| 1 | 2 | 3 | 4 | |
| poor | | | excellent | |
| | | | | |
| g. The quality of the presentation | | | | |
| 1 | 2 | 3 | 4 | |
| poor | | | excellent | |
| | | | | |
| h. Its capacity to have a meaningful impact on my profession | | | | |
| 1 | 2 | 3 | 4 | |
| poor | | | excellent | |

i. Comments: _____

Simon Wiesenthal Center Tools for Tolerance
Pre-training Interview Questions

- 1) How did you hear about the training?
- 2) How did you go about selecting members for your team?
- 3) What are some tolerance and/or hate issues facing your community?
- 4) Is there a specific incident or event that prompted you to get involved in the training?
- 5) How do you define hate crime?
- 6) How do you define terrorism?
- 7) How big a problem is hate crime and terrorism in your local jurisdiction?
- 8) How has the impact of hate crimes and terrorism changed your role as a (law enforcement officer, prosecutor, etc.)?
- 9) Currently, what is your agency's approach in dealing with hate crime and terrorism?
- 10) Have you ever been to any training on hate crime and/or terrorism? If yes, what do you think is different about the National Institutes Against Hate Crime and Terrorism?
- 11) What do you hope to gain from this training?

**Simon Wiesenthal Center Hate Crimes and Terrorism
Post-Training Follow-Up Interview Questions**

1. Please list three things you took away from the training.
2. When you went back to your job, did you do anything different in your roles and duties that could be related back to the training received at SWC?
3. Has your team met since you've been back?
4. Has there been any other form of contact (phone, email, snail mail)? Plans for follow-up?
5. What is the strategic theory for Institute #18 as you recall and understand it to be?
6. Please summarize your overall strategic plan. What are the goals, objectives, timelines etc.?
 - Do you feel that your group had enough input in the development of the plan?
 - Has the plan been changed or altered in any way after further review?
 - What are the next steps in the implementation of the plan?
7. Please tell me if you were you given the tools to do the following:
{Answer choices: yes, no, don't know. If yes, ask what tools were given}
 - Strengthen ethical and bias-free decision making
 - Enhance competency in inter-cultural communication and conflict resolution
 - Develop progressive leadership practices
 - Build capacity for inclusive and equitable organization
8. Did you learn anything new?
9. How could the training be improved?
10. I noticed in the Day One evaluations that the overall consensus from the group was that the jurisdictional team meeting was not a very beneficial exercise. Do you agree with this statement? If so, what about the exercise was not beneficial?
11. Have you received access to the Simon Wiesenthal tracking database (including password and instructions for use)? If yes, have you used it? Is it easy to maneuver and use?
12. Now that you've gone through the training, do you think that any changes are needed within your organization? If so, what changes do you think are needed? How do you propose to make changes in your organization?

**Simon Wiesenthal Center's Tools for Tolerance
National Institutes Against Hate Crimes and Terrorism Training**

Follow-up Questionnaire

Institute #17

Wisconsin

The Institute for Law and Justice is conducting an independent evaluation of the Hate Crimes and Terrorism training you attended January 25-28, 2004 at the Simon Wiesenthal Center in Los Angeles, CA. The questionnaire is divided into two sections: (1) strategic action plan, and (2) cohesiveness and teamwork. Your responses will help inform us on your team's activities as well as the effects the training had on the way you do your job.

Your answers are confidential and will not be associated with you in any way.

**PLEASE RETURN COMPLETED SURVEY IN SELF-ADDRESSED STAMPED
ENVELOPE BY NOVEMBER 5, 2004**

Strategic Action Plan

The following questions are based on items taken directly from your strategic action plan developed in January 2004. Please mark responses as indicated.

I. Action Item #1: Did your group meet with members from previous institutes?

1. Yes No

If yes,

a. How many of the previous institutes were involved? _____

b. Describe the purpose of the meeting:

	Yes	No
Debrief	<input type="checkbox"/>	<input type="checkbox"/>
Set future agenda	<input type="checkbox"/>	<input type="checkbox"/>
Collaborate with other groups	<input type="checkbox"/>	<input type="checkbox"/>
Other (please specify) _____		

2. Is action item #1 still applicable to the overall plan? If not, please explain why.

II. Action Item #2: Has the group implemented hate crime education initiatives for police, prosecution, District Attorneys, and judges?

3. Yes No

If yes,

a. When will training be held? _____

b. How many people will be attending? _____

c. Briefly describe the training:

4. Is action item #2 still applicable to the overall plan? If not, please explain why.

III. Action Item #3: In the last six months, has your group engaged in any outreach activities?

5. Yes No

a. If yes, please provide a brief summary of the outreach activities.

6. Is action item #3 still applicable to the overall plan? If not, please explain why.

IV. Action Item #4: Since you've been back from the training, has the hate crimes data collection system in the State of Wisconsin been improved?

7. Yes No

a. If yes, please state whether the following tasks have been implemented:

	Yes	No
Train probation managers	<input type="checkbox"/>	<input type="checkbox"/>
Review statute in DA	<input type="checkbox"/>	<input type="checkbox"/>
Conduct community forum	<input type="checkbox"/>	<input type="checkbox"/>
Add hate crime category to police database	<input type="checkbox"/>	<input type="checkbox"/>
Hate crime discussion at Felony court mtg	<input type="checkbox"/>	<input type="checkbox"/>

8. Is action item #3 still applicable to the overall plan? If not, please explain why.

9. Have there been any challenges to implementing the strategic action plan? Please note changes or modifications.

Cohesiveness and Teamwork

10. Please indicate your level of agreement with the following statements.

		Strongly Agree	Agree	Disagree	Strongly Disagree
a.	There have been an adequate number of team meetings	1	2	3	4
b.	Team members regularly communicate with one another	1	2	3	4
c.	There are clearly defined roles and responsibilities	1	2	3	4
d.	Team members proactively engage in strategic plan tasks and activities	1	2	3	4
e.	Each team member is vested in this effort	1	2	3	4
f.	All the necessary agency personnel have been included on this team	1	2	3	4

APPENDIX 8-B

Simon Wiesenthal Center's National Institutes Against Hate Crime and Terrorism Training

Case Study of Monmouth County, New Jersey

The Simon Wiesenthal Center (SWC) in Los Angeles, California, was first established in 1977 with a mission to educate the public about the holocaust atrocities in Nazi Germany. The Center has expanded to include the Museum of Tolerance that features programs and exhibits on the civil rights movement, hate crime, terrorism, and genocide. In 1999, the SWC was awarded a grant from the U.S. Department of Justice, Bureau of Justice Assistance, to implement a hate crimes training program for criminal justice professionals throughout the country. The program, titled the National Institutes Against Hate Crimes and Terrorism (hereafter the National Institutes), is designed to assist law enforcement, prosecutors, judges, and others to formulate innovative strategies aimed at addressing problems of hate crime and terrorism in their communities. Institutes have included teams from all over the country, creating a network of criminal justice professionals whose common interest is to fight hate crime and terrorism. As of December 2004, the SWC has conducted 20 Institutes.

Methodology

The Institute for Law and Justice (ILJ)²⁴ is conducting an evaluation of criminal justice training programs under a grant from the National Institute of Justice. ILJ is responsible for developing a training evaluation model, which others in the field can use as a guide in the development of their own training programs. The National Institutes training was one of the programs selected for evaluation. A team from ILJ has been reviewing the activities of three Institutes: Institute #17 (January 2004), Institute #18 (March 2004), and Institute #19 (May 2004).

Part of the training evaluation includes case studies of selected teams that received training. The case studies allow a more in-depth analysis of activities that teams implement after

²⁴ ILJ is a criminal justice research and consulting organization in Alexandria, Virginia. (www.ilj.org)

attending the National Institutes. Of particular importance are reactions to the training, effects the training may have had on trainees' responsibilities, and policy or program changes made as a result of the training. The Monmouth County, New Jersey team was selected as a case study because others consistently lauded them as a model program in follow-up interviews.²⁵

A case study is a systematic method designed to assemble information about an individual, group, setting or event to give researchers insight into how something operates or functions (Berg 2001). The case study allows researchers to observe the "real-life" events of their subjects (Yin 2003). Case studies often involve different data collection methods, including field observation, participant observation, in-depth interviews, archival records, documentation²⁶, and physical artifacts. In this case study, ILJ researchers used in-depth interviews and field observations. The benefit of a case study is that it enhances understanding of a particular issue (Yin 2003).

Key to the evaluation design of the National Institutes training is assessing program effectiveness in changing participants' behavior in dealing with hate crime and terrorism issues within their jurisdictions. This training is a unique program that requires participants to be proactive in seeking change within their jurisdictions in order to create innovations. A key element to the evaluation of the National Institutes program is assessing how the teams are using what they learned at the training. What needs to be measured in the National Institutes program evaluation is *how* the training is being used. For this case study, the central question is what did the training participants do after they returned from the Simon Wiesenthal Center in Los Angeles?

On December 7-9, 2004, two ILJ researchers visited Monmouth County, New Jersey,²⁷ to conduct an observation visit. The visit involved observation of training at the Monmouth County Police Academy, the New Jersey Department of Corrections Training Academy at Sea Girt, and the Brookdale Community College's Juvenile Bias Program. Sixteen interviews were conducted

²⁵ The Monmouth County, New Jersey team attended the National Institutes training in May 2004.

²⁶ Documentation can include letters, memoranda, communiques, agendas, announcements, meeting minutes, and written reports. See Yin (2003) for more information.

²⁷ Monmouth County has a diverse population of approximately 632,000 people. According to the 2000 census, approximately 19 percent of the population is Black, Asian, or Latino (United States Census Bureau, 2002).

with team members and their colleagues from the Monmouth County Prosecutor's Office.²⁸ The following individuals were interviewed:

- Monmouth County First Assistant Prosecutor
- Monmouth County Chief of Investigations
- Monmouth County Deputy Chief of Investigations
- Monmouth County Captain of Special Investigations Unit
- Monmouth County Lieutenant of Bias Crimes and Community Relations Unit
- Monmouth County Sergeant of Bias Crimes and Community Relations Unit
- Two Monmouth County Detectives from the Bias Crimes and Community Relations Unit
- Monmouth County Police Academy Director
- Monmouth County Police Academy Training Coordinator
- Howell Township Police Department Sergeant of the Community Service Bureau
- Monmouth County Sheriff's Department Sergeant
- Human Relations Commission Chair
- Monmouth County Municipal Court Judge
- New Jersey Department of Corrections Training Academy Director
- State Investigator for the New Jersey Department of Law and Safety

Findings

Monmouth County takes a multi-pronged approach in combating hate crimes. In particular, they have:

- Created innovative programs including a bias unit, specialized police academy training, and a juvenile offenders program;
- Recognized that any bias crime can be both destructive and serious in its impact on the targeted community; and
- Incorporated a team approach by forming relationships across the county with other law enforcement officials, prosecutors, human relations commissioners, religious leaders, members of the gay and lesbian community, and leaders in the minority communities.

²⁸ The county prosecutor is a non-sworn attorney responsible for the investigation division within the prosecutor's office and for the 48 police agencies in the county. He sets policy for these agencies and the chiefs of police in the county report directly to him. The investigation division is composed of three sections: criminal investigations (major crimes, forensics, sex crimes), special investigations (investigation support, prosecution support, economic crimes), and narcotics.

These three principle findings will be discussed in greater detail in the following section, focusing on how the SWC training impacted Monmouth County by elaborating on their activities and key aspects of the SWC training.

Creation of Innovative Programs

The Monmouth County Prosecutor's Office is committed to combating acts of hate and terror in its county. This commitment is evident in the variety of programs the prosecutor's office has implemented, and in their staff's interest to learn new and innovative strategies from the SWC National Institutes training.

SWC National Institutes Training

The SWC National Institutes training recognizes that effective training is not one-size-fits-all. The National Institutes training simultaneously teaches a set curriculum on hate crime and individualizes the training for each attending team. The curriculum covers different types of hate crime and acts of terrorism, including an exploration of the ethical dimension of hate crime and terror threats. The training is individualized in that each team develops its own strategic action plan based on the collaborative strategies learned in the training.

The training seeks to alter or expand the participants' views of hate crimes by demonstrating that acts of hate and terror cannot be either measured or punished like other crimes. Through a series of exercises, the participants come to define hate crimes as "...a crime against a person (or their property) targeted for their cultural group that he/she *is perceived to belong* to rather than for what he/she individually *believes or does*. The perpetrator feels *morally justified* in committing the crime" (Strategic Business Ethics 2004, p 25).²⁹ The key difference between hate crimes and other types of crime is that these other crimes are motivated while hate crimes are morally justified. Standard criminal justice system responses such as prison, it is argued, only serve to further the offender's cause because it allows him to cast himself as a martyr in an immoral justice system. The suggested response is to both address the values of the individuals who commit acts of hate or sympathize with those causes, and to develop expanded

²⁹ Strategic Business Ethics created and now teaches the National Institutes Against Hate Crimes and Terrorism training curriculum. Their copyrighted materials are included in the binder distributed to all participants.

strategies for dealing with acts of hate and terror. Monmouth County employs a variety of tactics to meet this suggested response, including the creation of a bias crime unit within the prosecutor's office, specialized training at the police academy, and a program for juvenile offenders.

During the National Institutes training, the attending teams are required to develop a strategic action plan that participants would then implement in their home communities. The plan can focus on a variety of targets, including school children or parents, offenders, law enforcement and colleagues within the criminal justice system, and general members of the public. For each target, the team identifies what needs to be done, how, by when, and who is responsible. For example, the team that attended the training from Monmouth County established a blue print for recognizing that acts of terror and hate are similarly motivated. Their plan involved connecting the Bias Crime and Community Relations Unit with the Homeland Security Unit within the prosecutor's office—to support community education programs and promote media outreach.

Monmouth County Activities

Monmouth County tackles hate crime through a multi-level approach, targeting investigation and enforcement through the creation of a designated bias unit; targeting community outreach, advocacy, and education through the Human Relations Commission (HRC);³⁰ and targeting offenders through the juvenile bias crime offenders program. The Monmouth County Prosecutor's Office is the lead agency in combating acts of hate and terror. In addition, the prosecutor's office's approach to bias crime serves as a model in the state of New Jersey. Of the 21 counties in New Jersey, Monmouth County is one of only three counties that have a bias crimes unit. In addition, Monmouth County was the first county in the state to create a Human Relations Commission, and it is the only county to implement a juvenile bias crime offenders program.

³⁰ The HRC is an independent body of government workers, police, teachers, business people, victim advocates, and community leaders whose mission is to combat bias and discrimination in Monmouth County by improving community relations.

Monmouth County Prosecutor's Office Bias Crime and Community Relations Unit

In 1988 an innovative first assistant prosecutor in the Monmouth County Prosecutor's Office formed the Bias Crime and Community Relations Unit after reading a newspaper article about a new bias crimes unit and police academy training program in a sister county to Monmouth County. Since its formation, the Bias Crime and Community Relations Unit has

- Developed Standard Operating Procedures (SOPs) on how to handle bias crimes.
- Provided training to all police chiefs in Monmouth County to encourage their acceptance of the program and “get them all on board because change only happens from the top.” (*interview with Monmouth County First Assistant Prosecutor, December 7, 2004*).
- Conducted train-the-trainer sessions with supervisors across the 48 Monmouth County township police departments so that managers could go back to their own departments and initiate the training.
- Trained at least one officer from each police department in the county to be a bias officer.
- Formed the Bias Officers Association, overseen by the Monmouth County Prosecutor's Office, which meets monthly to discuss bias cases.
- Taken a leadership role in investigating and prosecuting bias crimes in the county. The bias unit, functioning as part of a county agency, interacts with all the individual jurisdictions within the county.

Monmouth County's approach to bias crimes is enforcement and education. Cases are addressed immediately due to the impact one incident can have on the entire community. Immediacy assists in keeping cases from escalating. Or as the chief of detectives stated, “a small match burning can turn into a forest fire” (*interview on December 7, 2004*). This mission is emphasized through training. There is a constant flow of information between the police administration and the community. Anti-bias policy is set by top administrators, filters down to the detectives and other staff, and is communicated to the community.

Monmouth County Police Academy

The Monmouth County Police Academy administrators and Monmouth County chief prosecutors believe there is real value in bias and diversity training for law enforcement. As the first assistant prosecutor stated, “Stereotyping...clutters an officer's mind and prevents him from seeing the real issues of a crime.” Bias and diversity training mitigates stereotyping by exposing

trainees to different cultures, religions, nationalities, and ethnicities through guest lectures and training materials.

The Monmouth County Police Academy bias and diversity training exceeds the minimum training requirements on both coverage and hours set by the New Jersey Police Training Commission. There is one full day allotted for bias training and one for diversity instruction during the 20-week recruit-training program. Terrorism is a separate block of instruction, which includes first responder and weapons of mass destruction training.

The police academy conducts a regular review of training using course evaluations and observations. Performance is measured using the instructor's performance objective sheets, student test scores, student exit evaluations, and comments from observers. Changes are made according to comments and feedback. According to the police academy's chief training officer, students consistently rate the bias crime and diversity training taught by the prosecutor's office bias detective as one of the best classes at the academy and often request that the course be extended beyond the two days.

Juvenile Bias Crime Offenders Program

In 2003, the Monmouth County Prosecutor's Office, the Center for Holocaust Studies at Brookdale Community College,³¹ the Monmouth County Probation Department, and the courts collaborated to create the Juvenile Bias Crime Offenders Program. This program, court mandated for juveniles adjudicated for bias crimes, is a 12-week course that addresses the devastating consequences hate crimes can have on individuals and communities. The curriculum includes readings on the Holocaust and victimization, writing short papers, and visits by guest speakers. Speakers represent a variety of communities including Jewish, Muslim, gay and lesbian, and the Black community. The Juvenile Bias Crime Offenders Program is aimed at helping repeat offenders understand the consequences of their actions. Organizers hope the experience prevents the participants from engaging in future bias crimes and that the interaction with different groups gives them a greater appreciation and respect for people of different races,

³¹ The Center for Holocaust Studies at Brookdale Community College first opened in 1979 to educate the public about the genocide. Over the past 25 years, the Center has grown and now offers training for teachers on "prejudice reduction" for school age children. In addition, the Center runs an art and writing contest for children in grades 9-12.

cultures, and lifestyles. Currently, the program has not been evaluated so the impact on recidivism is unknown.

Impact of SWC Training on Monmouth County

The SWC National Institutes training aims to change how criminal justice system personnel think about and approach hate crimes by understanding that it is conceptually different from other crimes. In essence, they seek to challenge the existing criminal justice paradigm that hate crime offenders do not differ from other types of offenders. Instead, the training demonstrates that hate crime offenders vary in their belief that their crime is justified and in their strong commitment to their cause. This difference requires the criminal justice system to respond differently to crimes of hate and acts of terror than they would to perpetrators of conventional crimes. One way that jurisdictions can respond differently is by recognizing that all acts of hate and terror are serious, regardless of whether the act would not be classified as a crime, classified as a misdemeanor, or classified as a felony.

In Monmouth County, acts of bias and terror are taken very seriously by the prosecutor's office. Bias crimes in Monmouth County are viewed as the "most heinous crime that can be committed" because they consider the thought process of selecting someone on the basis of who they are (i.e., color of their skin, religion, etc.) as terrorism (*interview with lead detective from the Bias Crime and Community Relations unit, December 7, 2004*). Consequently, the prosecutor's office equates bias crime to terrorism because of the far-reaching effects it can have on the community.

During the training, each attending team creates an individualized strategic action plan, providing a functional blueprint for how they are going to deal with hate crimes in their own jurisdiction. To date, one team from Monmouth County has attended the National Institutes training. Prior to attending the National Institutes training, the Monmouth County Prosecutor's Office was implementing its plan for dealing with acts of hate in their community. Monmouth County has responded to problems of bias and acts of terror through the creation of a stand alone bias crimes unit, targeted bias and diversity training at the police academy, and establishing a special program for repeat juvenile bias crime offenders.

The motivation for attending the National Institutes training for the prosecutor's office was to learn more about hate crimes and acts of terror, to find out what others around the country

were doing, and to discover any new innovations in dealing with these crimes and incidents. The main impact the training had for the prosecutor's office was a recognition that what they were doing was "on the right track." What the lead bias crimes detective and others from the prosecutor's office discovered is that Monmouth County is at the forefront in dealing with bias crimes and diversity issues. The team realized that their programs were more developed than they had realized and that other teams wanted to emulate the Monmouth County programs.

Importantly, what the team gained from the training was a renewed confidence in their strategies for combating acts of hate and terror, and a renewed enthusiasm from the top administrators on down for the Monmouth County programs. When asked what the biggest change was since members attended the training, the chief of detectives replied that although the prosecutor's office prides itself on being the best and knows other units will call on them for assistance, they can only get better. The team that attended the training at Simon Wiesenthal Center heard different views from around the country and it reinforced their drive and commitment to their mission.

Focus on Team Approach

A critical component of the SWC National Institutes training is focused on establishing collaborative relationships among agencies working to combat hate crime. This section outlines the training efforts and details Monmouth County's activities in fostering relationships with other agencies.

SWC National Institutes Training

An important focus of the National Institutes training is on building teamwork. The SWC believes that the most effective means of implementing jurisdictional and organizational change is by requiring collaboration among parties that do not often work together. To facilitate collaboration, the National Institutes training requires each jurisdiction attending the training to send a select group. The required group composition is a judge; prosecuting attorney; public defender; probation official; two law enforcement officials; and a school official, victim/witness advocate, or a human relations commissioner. Before attending the training, the team members often do not know each other. The training provides them the opportunity to establish a working relationship. This working relationship is further enhanced through their collaborating on the creation of a strategic plan. Not only are the strategic plans richer for having input from multiple

individuals with very different roles and perspectives on the criminal justice system, but each member of the team is more likely to “buy-into” the project.

Partnerships with Criminal Justice and Community Agencies

The Monmouth County Prosecutor’s Office fosters collaborative relationships with other criminal justice agencies, community organizations, and community leaders. For instance, the prosecutor’s office and the Monmouth County Human Relations Commission (HRC)³² have established a working relationship through the involvement of the lead bias crimes detective on the commission board. In addition, collaborative relationships have been formed with both state level training officials and fostered with local religious and community leaders.

The HRC is an independent body that was created in 1990 through the efforts of the prosecutor’s office and is composed of three teams: Police-Community Relations, Youth Awareness, and Economic Issues. The Police-Community Relations team examines police and community bias issues and crimes. The role of the Youth Awareness team is to build a rapport with the schools (however they are having some difficulty getting access to the schools). Lastly, the Economic Issues team assists residents in finding affordable housing. The Police-Community Relations team members help monitor the 24-hour bias hotline run by the Monmouth County Department of Human Services. The HRC members make inquiries about the incidents reported, speaking with both the reporter and the person being reported on. If the incident in question is a crime, the HRC reports its findings to the prosecutor’s office bias unit.

Equally important to establishing working relationships with community agencies such as HRC, the Monmouth County Prosecutor’s Office works with local and state criminal justice agencies on issues of bias and diversity. As detailed above, the prosecutor’s office works with bias officers from each jurisdiction in the county, and they run the bias and diversity training component for the county police academy. In addition, the prosecutor’s office has established a relationship with the state training academy. The New Jersey Department of Corrections Training Academy at Sea Girt is responsible for training all New Jersey parole, probation, and state corrections officers. When the director took his post three years ago, he realized that the curriculum was outdated and needed changing. This need for updating was substantiated by the

New Jersey Police Training Commission, which in 2003 directed state academies to provide training on domestic and international terrorism techniques and Muslim culture to prepare law enforcement in the event of another attack.

After consulting with the Association of Academy Directors, the curriculum was modified to include more training on cultural diversity. The training is evaluated every cycle and every six months and recruits have given high ratings to the diversity training taught by the Monmouth County Prosecutor's Office bias detective. The director believes that the biggest challenge in the training is in changing the emphasis from "this is what you cannot do" to "opening minds" (*interview on December 9, 2004*). The academy teaches the bias and cultural diversity class at the start of the academy to stress to recruits the importance of these issues. In addition, the training gives recruits a greater awareness of the variety of people they will encounter when they leave the academy. As such, they are constantly thinking about the concepts as they rotate through the other portions of the training. Having this class first has also helped abate clashes among recruits.

Impact of SWC Training on Monmouth County

The SWC National Institutes teams are ideally made up of police, probation, prosecution, public defenders, judges, and human relations commissioners or school officials. Involving people from the different layers of the criminal justice system presented the Monmouth County team with some unique issues. According to a number of people interviewed for this case study, including a judge, the first assistant prosecutor, and police officials, New Jersey regulations prevent professionals in the different layers of the criminal justice system from socializing or working together on joint projects such as a hate crimes task force. The rules are an attempt to restrict officials from influencing the actions of their colleagues or subordinates. For example, it is believed that if a judge and police officer work together outside of their official capacity, the police officer may be influenced to behave a certain way or make a particular decision based on contact with the judge.

³² The HRC includes representatives from the Monmouth County Freeholders (the Freeholders are another name for elected county supervisors), the Prosecutor's Office, National Conference for Communities and Justice, and the NAACP. There are approximately 60 people on the Committee and all members are self-selected.

Team collaboration among police, probation, and the judiciary is good in theory but regulations prohibit these members from working together. These restrictions prevented Monmouth County from sending a complete team as required by the National Institutes coordinators. In addition, it subverts the fostering of interagency coordination that is a valued part of the training. The Monmouth County team reflected these restrictions by being heavy on law enforcement team members from the prosecutor's office. The team consisted of the deputy chief, the captain of special investigations, two detectives, two assistant prosecutors, and one public defender. The advantage of the composition of Monmouth County team is that top level administrators formed a part of the team. Thus while Monmouth County was not able to build teamwork through the channels that the National Institutes training would prefer, they were able to both foster teamwork within the prosecutor's office and achieve a renewed commitment by the top level administrators to actively combat crimes of hate and acts of terror.

Recommendations and Conclusion

The participants involved in this case study unanimously stated that they enjoyed the learning experience and collaborative environment at the National Institutes training. When asked for recommendations to improve the training provided by Simon Wiesenthal Center in Los Angeles, they offered the following suggestions:

- Include more training on terrorism at SWC.
- Jurisdictions that send chiefs of police to the training should also send their bias officer to ensure officer buy-in.
- Invite more line officers to the training.
- Design a course specifically for top-level administrators.
- Design a course specifically for line level personnel.
- Create a follow-up training course that addresses the escalation of hate crime.
- Add a component to the training specifically for correctional staff.

The purpose of this case study is to better inform ILJ researchers about the activities teams attending the National Institutes training at Simon Wiesenthal Center were involved in. Our observations and interviews with the criminal justice professionals in Monmouth County confirmed that this county takes the issue of hate crimes very seriously and has undertaken a series of innovative strategies to combat these crimes.

References

Berg, B. (2001). *Qualitative research methods for the social sciences*. Boston, MA: Allyn and Bacon.

Strategic Business Ethics. (2004). *National Institutes Against Hate Crimes and Terrorism Strategies against hate crimes and terrorism: Consensus building and a blueprint for collaborative action*. Los Angeles: author.

United States Census. (2002). Population, housing units, area, & demographics. Retrieved February 10, 2005 from <http://www.factfinder.census.gov>.

Yin, R. (2003). *Case study research: Design and methods (3rd ed.)*. Newbury Park, CA: Sage Publications, Inc.

APPENDIX 8-C

Simon Wiesenthal Center's National Institutes Against Hate Crime and Terrorism Training

Case Study of Madison, Wisconsin

The Simon Wiesenthal Center (SWC) in Los Angeles, California, was first established in 1977 with a mission to educate the public about the holocaust atrocities in Nazi Germany. The Center has expanded to include the Museum of Tolerance that features programs and exhibits on the civil rights movement, hate crime, terrorism, and genocide. In 1999, the SWC was awarded a grant from the U.S. Department of Justice, Bureau of Justice Assistance, to implement a hate crimes training program for criminal justice professionals throughout the country. The program, titled the National Institutes Against Hate Crimes and Terrorism (hereafter the National Institutes), is designed to assist law enforcement, prosecutors, judges, and others to formulate innovative strategies aimed at addressing problems of hate crime and terrorism in their communities. Institutes have included teams from all over the country, creating a network of criminal justice professionals whose common interest is to fight hate crime and terrorism. As of January 2005, the SWC has conducted 21 Institutes.

Methodology

The Institute for Law and Justice (ILJ)³³ is conducting an evaluation of criminal justice training programs under a grant from the National Institute of Justice. ILJ is responsible for developing a training evaluation model, which others in the field can use as a guide in the development of their own training programs. The National Institutes training was one of the programs selected for evaluation. A team from ILJ has been reviewing the activities of three Institutes: Institute #17 (January 2004), Institute #18 (March 2004), and Institute #19 (May 2004).

Part of the SWC training evaluation includes case studies of selected teams that received training. The case studies allow a more in-depth analysis of activities that teams implement after attending the National Institutes. Of particular importance are reactions to the training, effects

the training may have had on trainees' responsibilities, and policy or program changes made as a result of the training. The Madison, Wisconsin, team was selected as a case study because they have been hailed by the SWC and others for using the National Institutes training as a springboard to fighting hate and bias crimes in Madison. Five teams from Madison have attended the National Institutes training. They continue to implement a host of initiatives first planned during the training.

A case study is a systematic method designed to assemble information about an individual, group, setting or event to give researchers insight into how something operates or functions (Berg 2001). The case study allows researchers to observe the "real-life" events of their subjects (Yin 2003). Case studies often involve different data collection methods, including field observation, participant observation, in-depth interviews, archival records, documentation³⁴, and physical artifacts. In this case study, ILJ researchers used in-depth interviews and field observations. The benefit of a case study is that it enhances understanding of a particular issue (Yin 2003).

Key to the evaluation design of the National Institutes training is assessing program effectiveness in changing participants' behavior in dealing with hate crime and terrorism issues within their jurisdictions. This training is a unique program that requires participants to be proactive in seeking change within their jurisdictions in order to create innovations. A key element to the evaluation of the National Institutes program is assessing how the teams are using what they learned at the training. What needs to be measured in the National Institutes program evaluation is *how* the training is being used. For this case study, the central questions are what did the training participants do after they returned to Madison, did they implement strategic action plans developed in the training, and what has sustained their commitment to their strategic action plan?

³³ ILJ is a criminal justice research and consulting organization in Alexandria, Virginia. See www.ilj.org for more information.

³⁴ Documentation can include letters, memoranda, communiqués, agendas, announcements, meeting minutes, and written reports. See Yin (2003) for more information.

On April 20-21, 2005, two ILJ researchers visited Madison, Wisconsin,³⁵ to conduct an observation visit and interviews with past National Institutes participants. The visit involved observing a hate crimes task force meeting composed of five teams of previous National Institutes attendees. The following 14 individuals were interviewed:

- City of Madison Police Department - Detective in Persons Crimes Unit
- City of Madison Police Department - Detective in Investigative Service Bureau
- University of Wisconsin at Madison Police Sergeant
- Wisconsin Department of Corrections, Division of Community Corrections - probation officers (2 interviewed)
- Dane County District Attorney's Office - Assistant District Attorneys (3 interviewed)
- Dane County District Attorney's Office - Director of Victim Services
- Dane County District Attorney's Office - victim/witness advocate
- Dane County District Attorney's Office, Critical Incident Response Team - victim/witness advocate
- Dane County Circuit Court - judge
- Madison Metropolitan School District - GLBTQ³⁶ Resource Teacher
- Madison Jewish Community Center and Jewish Social Services of Madison - Executive Director

Findings

There were two primary reasons the ILJ research team conducted a case study of Madison, Wisconsin. First, we wanted to learn personally the impact of the National Institutes training. Second, we wanted to explore why the Madison teams have been successful in both implementing their strategic action plans and creating new initiatives in their fight to combat acts of hate and terror.

Impact of National Institutes Training

In November 2000, the first group of criminal justice professionals from Madison attended a National Institutes training. Since then, Madison has sent four more teams to the

³⁵ Madison, Wisconsin is a relatively large mid-western city with a population of 208,000, not including an additional population of 40,000 undergraduate students attending the University of Wisconsin at Madison. According to the 2000 census, 16 percent of the population is African American, Asian, or Latino (United States Census Bureau, 2002).

training, with the last group attending in January 2005. Two years prior to sending the first team to the National Institutes, the Madison Police Department (MPD) began to address how hate crimes are handled within the department. It was recognized that there was a “big void” in dealing with hate crime cases because there was no formal tracking system in place. Consequently, the cases were “falling through the cracks,” according to MPD detectives. The police department began to systematically track hate crime cases and started to educate detectives on what hate crimes are and how they should be investigated.

The main types of hate crimes and bias incidents in Madison are property damage, graffiti, intimidation, and simple assault. Historically, anti-GLBTQ incidents have ranked the highest in the number of crimes committed, followed by anti-Semitic and racial incidents. In terms of reported incidents, the number of hate crimes is quite low. In 2003, the state of Wisconsin reported a total of 37 hate crimes to the Federal Bureau of Investigation’s Uniform Crime Reporting system (U.S. Department of Justice 2004). The city of Madison reported three hate crimes in 2003, two were motivated by race and one by sexual orientation. However many of those interviewed believe that the number of reported crimes does not reflect the true nature of the problem. Furthermore, they argue, bias and intolerance is pervasive in Madison. For example, Madison has recently been dealing with an increase in racial incidents between African American and Latino gangs in middle and high schools, and between different gangs and “frat boys” from the university.

Reaction

Overall, attendees to the National Institutes training were impressed with the training. Reaction to the training by Madison team members includes the following comments:

- “I thought I’d died and gone to heaven.”
- “I really admired the people doing the training, it’s their life’s mission.”
- “[The training] changed my outlook on hate and empowered me to make a difference, or try to. [It] gets people to think about more common themes of humanity.”
- “By and large, everyone who attended thought it was excellent.”
- “Racial profiling training was excellent, [I] would love to bring it here.”

³⁶ GLBTQ stands for gay, lesbian, bi-sexual, transgendered, or questioning.

- “One of the best trainings I’ve ever gone to.”
- “It really gets your mind going because we have ingrained prejudices and [we] base our decisions on our beliefs.”
- “The training really opened my eyes.”
- “The Museum [of Tolerance] is very effective.”
- “[I] have come back and asked about every case, where is my prejudice?”
- “Hands down, the best in my life.”
- “The exercises were really good and useful, [they] involved all of the group and didn’t allow you to just sit back.”
- “The training was so inspiring that we wanted to come back and reach the entire criminal justice system.”

Team members, while widely heralding the training as “excellent,” did find some aspects of the training that needed improvement. In particular, attendees were nearly unanimous in their reactions to one of the guest speakers, the National Institutes’ website, strategic action plans, and on some of the ways the facilitators faltered.

Most team members did not “buy the skinhead” because, according to one probation officer, “you don’t turn that quick.” Another National Institutes participant stated that the “skinhead was minimizing and self-important. [I] felt like I needed debriefing because he destroyed his son.”

Another issue involved the National Institutes’ website. The website is not operating as presented during the training. The general consensus among the team members was that the web-based follow-up is not widely used; mainly, they felt, because of a lack of direction from SWC. However, some team members acknowledged that they could be more proactive, but chose not to use the service as much as they could.

Another common concern among the team members was that they were not given enough time to work on their strategic action plans. As one assistant district attorney stated, “we only had one and a half hours for developing a plan, we needed more time.” One team member, whose group “could not pull together the strategic action plan,” stated that they “felt like we had failed.” Some felt that one solution, given the tight training schedule, would be to make it clearer to the attendees ahead of time that they would be developing a strategic action plan. In

that way, they could begin pondering topics and discussing ideas among themselves either prior to, or during the first day of the training.

In a related complaint, many of those interviewed expressed disappointment that the facilitators did not allow teams to freely choose the target of their strategic action plans. As one attendee stated, “we essentially had to do what [the facilitators] wanted.” Many of the teams from Madison wanted to focus their strategic action plans on education programs in schools because they had many school connections and they felt it was a needed curricular element in the Madison school system. Some felt thwarted in their efforts, and were “reprimanded” for being “too traditional.” Instead they were told “to think outside the box.”

Other team members felt that the National Institutes facilitators had a lack of expertise on dealing with GLBTQ issues. This was especially evident when one non-Madison participant at the training made very loud and negative anti-homosexual comments. According to multiple team members, the comments were loud enough that the entire class heard them, yet the facilitators ignored the comments. When this incident was brought up with the facilitators, they could provide no explanation for why they allowed the comments to stand. In a related instance at a different National Institute training, a victim advocate complained that “some participants talked under their breath and said very negative things, creating an unsafe atmosphere in the room. [The] facilitators ignored it.” Other team members, however, had the opposite experience. As one team member stated, “the staff was extraordinary, gifted in their field. They kept reminding us to leave the politics outside....”

In discussing some parts of the training content, many interviewees agreed with the statement by one assistant district attorney that the presentations and topics were “piecemeal and not tied together or discussed, and then [we were] just expected to create a plan. They didn’t need to preach to the choir as much as give tools for how I could explain hate crime charges to the community.”

Benefits and Outcomes

When asked how the training impacted the Madison participants, one team member stated, “we think more deeply, we take more opportunities to reach out to the community and create alliances and build bridges. We may have done this if we didn’t go to the Simon Wiesenthal Center [training], but probably not.” Another benefit of attending the National

Institutes training was the increase in awareness of hate crime by the judiciary. As one assistant district attorney stated, “the training makes it easier to take these cases to court, if the judges have also been trained.”

Specific outcomes based on the National Institutes training have included:

- Formation of the Dane County hate crimes task force.
- Creation and distribution of 3,000 copies of a hate crimes brochure, also translated into Spanish.
- Presentation given to the MPD management team in Fall 2003 by a team member and MPD detective asking for additional training, and stressing the importance of how all hate crimes, however, minor, need to be taken seriously.
- Training given to 13 detectives who volunteered to be trained on hate crimes and bias incidents and serve as resources for hate crime investigations because MPD does not have a specific hate crimes unit because of their decentralized command structure.
- Formation of a Department of Corrections (DOC) hate crimes committee by DOC employees based in Madison in 2001 by a member of the department who had attended the National Institutes training. The committee meets monthly. The committee has sponsored a number of outreach efforts, including giving a 30-minute presentation followed by a 90-minute call-in session on a new local Spanish radio station. In addition, the committee has sponsored regional training for probation and parole officers and is in the process of establishing mandatory training for new agents.

Challenges still facing the Madison teams include the fact that hate crimes are treated “on the side” rather than as a main focus because of the lack of an official hate crime unit within MPD. Furthermore, while the assistant district attorneys will charge hate crimes, they often agree to drop them in a plea deal.

Madison’s Success

One purpose of this case study is to determine why the Madison team members have been successful in their persistence in meeting their strategic action plan goals and creating innovative initiatives since attending the training. This is especially important given the number of teams we have followed from the National Institutes trainings 17-19 who have failed to pursue their strategic action plans or even meet upon return from the training.

Lessons Learned

A lesson that can be learned from the five Madison teams is why some of the teams continued to meet and others did not. A team member from the second team of Madison criminal justice personnel to attend the training (in December 2001) stated that their team met for the first year after attending the training. The team is no longer meeting because of a lack of team leadership. As is common with many of the teams from across the country who have attended the National Institutes, one or two individuals shoulder the leadership responsibilities. When these individuals retire, change jobs, or become too busy to continue to bear the leadership duties, the teams tend to splinter and stop meeting.

A second lesson that can be learned from the Madison teams is collaboration and dedication. Many team members stressed the importance of collaboration. In Madison, this collaboration has taken the form of task force meetings. Interviewees identified two previous iterations of the current hate crimes task force, each of which, eventually disbanded because of a lack of leadership. The issue for many team members is not the lack of commitment to the work of the task force, but finding the time in addition to their regular job duties to carry out this work. As one detective stated, “there is no formal structure to maintain these teams as part of your formal job.” Team members expressed frustration at the lack of time they have to devote to team meetings and their strategic action plans once they return to the office at the conclusion of the training. In the words of one member of the Dane County District Attorney’s Office, “it is very difficult, especially working in the D.A.’s Office, to take extra time to do the strategic action plan tasks, which are essentially done off hours and not as part of our traditional job.”

One solution for dealing with the leadership issue, as suggested by various team members, would be to have SWC require the identification of leaders as a part of the strategic action plan, and a commitment of Dane County department heads to support the plans. Solutions offered by various attendees to the problem teams dissipating upon their return from the National Institutes training, include:

- Designate a leader.
- Gain support of top management.
- Allow training materials to be shared with colleagues who did not attend the training, or provide tools so attendees can share what they know.

- Require firm due dates for the completion of tasks on the strategic action plan by SWC.
- Improve the National Institute website and make it more user friendly.

Roundtable Meeting

On April 20, 2005, the five teams of Madison National Institutes participants gathered for a roundtable meeting. This was the first time that all teams met as one large group. The purpose of the meeting was to “regroup and figure out where we go from here.” The 28 participants reported on their team activities, or lack thereof, and debated how to proceed. Four of the five teams who attended the training have continued with their strategic action plans, to some degree, upon returning from the National Institutes training. As detailed above, the teams struggled with maintaining their commitment and ability to carry out their plans without broader department support or team leadership. Some teams were limited by the lack of funds to carry out the ideas.

The roundtable participants reaffirmed their commitment to combating acts of hate and bias in Madison and asked the question, “where do we go from here?” After much debate, it was decided that the participants would abolish their original five teams and regroup into two. These two teams each chose to focus on a specific goal. The first goal is to provide outreach to the Dane County Boys and Girls Club with the aim of preventing hate crimes and bias incidents. The second goal is to upgrade the education and training of law enforcement and social service agencies on the issues of tolerance, acceptance, bias, and hate. In addition, the roundtable participants felt it was important to continue their own training through the exploration of their own biases. To that end, they agreed to attend a YWCA workshop on “Unlearning Racism” and they are currently exploring field trip options, such as a visit to the Black Holocaust Museum in Milwaukee, Wisconsin.

To sustain their focus, the participants decided to hold quarterly roundtable meetings for all the participants. The individual teams are currently negotiating their own schedule of meetings and tasks. The roundtable meeting was groundbreaking in its efforts to maintain the team members’ dedication to combating hate crime in the city of Madison and Dane County in a systematic manner. The roundtable participants realize that sustainability requires both collaboration and a system that is not dependent upon one or two leaders to shoulder the responsibilities. The roundtable is currently being driven by such a leader, but they are seeking ways to make her “replaceable.”

Recommendations and Conclusion

The Madison team members attending the National Institutes training unanimously agreed that the training was a very positive experience. Many participants were profoundly impacted by the experience. In particular, they found the Museums of Tolerance and Holocaust, coupled with the guest speakers, to be particularly moving. When asked for recommendations to improve the National Institutes training, the following suggestions were offered:

- Allow more time for team members to develop their strategic action plan, perhaps by assigning it before the training starts.
- Drop the strategic action theory component of the training and use that time on the strategic action plans.
- Do not force every team attending the training to devote their strategic action plan to the same topic.
- Give participants time to sit down with other participants in the same field to discuss what is and is not working in their jurisdictions. This could be done by moving the reception to a night later in the training, after which time participants have had the chance to get to know each other.
- Provide tools, such as creating materials, disks, or website that participants could access and use to teach the training to colleagues. Some participants were frustrated that they were expected to teach the training without being allowed to use the copyrighted materials.
- Employ some type of selection process when choosing participants to attend to ensure that team members have prior commitment, support from their top management, and have worked in the area.

The purpose of this case study is to inform ILJ researchers about the activities of the Madison, Wisconsin, team members attending the National Institutes training. Our interviews and observations with the five teams who attended the National Institutes training helped us understand the specific impact the training had on Madison. In addition, our interviews shed light on why some jurisdictions are successful in continuing the activities started during the training, and why other teams have a commitment but are not able to follow through on their plans.

References

- Berg, B. (2001). *Qualitative research methods for the social sciences*. Boston, MA: Allyn and Bacon.
- United States Census. (2002). Population, housing units, area, & demographics. Retrieved May 12, 2005 from <http://www.factfinder.census.gov>.
- United States Department of Justice. (2004). *Hate Crime Statistics 2003*. Washington, D.C.: Federal Bureau of Investigation, U.S. Department of Justice.
- Yin, R. (2003). *Case study research: Design and methods* (3rd ed.). Newbury Park, CA: Sage Publications

Chapter 9

NCLETTTC Advanced Leadership Techniques Training for First Responders, Corrections, and Security Officers

Introduction

The purpose of the evaluation described in this chapter was to apply the expanded Kirkpatrick evaluation model to the Advanced Leadership Techniques for First Responders, Corrections, and Security Officers training program offered through the National Corrections and Law Enforcement Training and Technology Center (NCLETTTC) headquartered in Moundsville, West Virginia. This research project obtained and analyzed data at all four levels. In addition, we examined data derived from both traditional and online mediums of training delivery.

Context for Evaluation

Project Initiation

A two-day kickoff meeting was arranged between all parties involved in the evaluation at NCLETTTC's location in Moundsville on November 5 and 6, 2003. At the initial meeting on November 5, project goals and responsibilities were discussed. By the end of the second day on November 6, a timeline, milestones, and a work plan chart outlining objectives, tasks, products, responsibilities, and due dates had been created (see Appendix 9-A); and a memorandum of understanding had been signed. A second follow-up meeting was held on February 3 and 4, 2004, to discuss and agree on a schedule of evaluation activities.

Overview of NCLETTTC and the Leadership Training Program

The stated mission of NCLETTTC, as a national training center, is to offer quality training and technology evaluation for criminal justice and public safety agencies in a unique facility and in a cost-effective manner. NCLETTTC is a non-profit organization located on 19 acres inside the former West Virginia Penitentiary in Moundsville, West Virginia. It became operational in May 2001 and has trained approximately 5,000 corrections, law enforcement, and public safety agency personnel. NCLETTTC has nine full-time staff with two full-time contractors. It also has approximately 40 adjunct instructors who are available to do training anywhere in the country.

The 44,000 square-foot training center includes four state-of-the-art classrooms, a 20-workstation computer lab, a firearms simulator, a technology demonstration lab, and a large bay area for training and events. Each classroom has an instructor workstation with projector and screen, VCR, and surround sound system. The computer lab contains an instructor station, VCR, DVD, and surround sound system. It also contains a 42" plotter and Smart Board. In the technology demonstration lab, approximately 50 vendors have tables set up to display their technologies with their literature.

In 1998, NCLETTTC management, recognizing the need for a facility that provides training for all public safety agencies and offers a setting for technology evaluation, secured a \$1 million appropriation from the U.S. Small Business Administration to help fund NCLETTTC. Additional funding was provided by the West Virginia Legislature to renovate certain areas to improve the quality and suitability of the physical plant for active training. Today, four different agencies financially support the facility: West Virginia Division of Corrections (owns the facility), Moundsville Economic Development Council (leases and manages the facility), Office of Law Enforcement Technology Commercialization (an NIJ program), and National Technology Transfer Center (which uses its expertise in training development and delivery to make advances at the Center).

NCLETTTC conducts training on site, online, and via a mobile training team. With respect to recruitment, the Federal Office of Management and Budget does not allow NCLETTTC to conduct direct advertising. However, NCLETTTC uses media outlets, workshops and booths at national conferences, fact sheets, websites, and word of mouth to reach potential trainees.

The center's most popular course, "Management & Leadership Skills for 1st Line Supervisors," which has been offered since the inception of NCLETTTC, was revised and updated in 2003. The resultant new course, "Advanced Leadership Techniques for First Responders, Corrections, and Security Officers" (henceforth referred to as "leadership training"), was designed as a 16-hour block of instruction conducted over two days. It was intended for personnel already in leadership positions who are responsible for the professional development, policy implementation, tactical mission execution, and formal evaluation of staff and operators within the first responder, corrections, and security officer arena. The goals of the course are to

enhance and build on a supervisor's leadership abilities and provide practical application of values-based leadership within a variety of organizations.

As a result of taking this course, trainees should be able to: conduct effective developmental counseling; identify, assess, and develop their subordinates both personally and professionally; know the importance of leading by example; and demonstrate what a leader must be, know, and do both professionally and personally. The objectives of the course include enhancing student understanding of the following topics: human dimension of leadership, mission stress, stress of change, organizational climate and culture, leadership style, intended and unintended consequences of leadership, developmental counseling, leader's responsibilities, leader as a counselor, leader counseling skill, leader's limitations, types of developmental counseling, performance and professional growth counseling, approaches to counseling, counseling techniques, counseling process, and documenting counseling. A significant portion of the course includes in-class exercises and scenario-driven counseling sessions for professional development purposes.

Rationale for Selection

Three factors contributed to the decision to evaluate the NCLETTTC leadership course. First, NCLETTTC officials were willing and eager to be involved in an evaluation. Second, all of the separate leadership classes evaluated in this project were taught by the same instructor, thereby contributing to consistency in what was delivered to trainees. The instructor, who was involved in this study since its inception, has a Ph.D. in Public Policy, Public Administration, and Politics and had a strong background in teaching leadership classes. The instructor also possessed a good understanding of evaluation research. Because of this, the instructor made a conscious effort not to revise or change course content or instructional methods in any manner across the various deliveries. Third, an online version of the course was under development at the time the decision was made to evaluate the NCLETTTC leadership course. Since the traditional and online versions of the course involved the same curriculum and instructor, the possibility existed for a quasi-experimental comparison of the two training modalities.

Overview of Literature Relevant to the Training

The terms leadership and management tend to be used rather loosely in the criminal justice literature. According to Freeman (1999), the fundamental functions of management

include planning, direction, organization, and the exercise of control. Management functions are primarily dictated by an employee's position within an organization and the responsibilities and authority attached to that position. On the other hand, leadership is not as dependent on position. Freeman notes that leadership is more about vision, values, motivation, communication, and handling change. As such, some tasks (e.g., routine aspects of recruiting, designing schedules, and ordering supplies) are clearly managerial. But the challenges associated with getting employees to work hard, pull in the same direction, and abide by organizational directives are typically addressed using both leadership and management.

Despite some exceptions (e.g., Freeman 1999; Houston 1999; Seiter 2002), most of the existing criminal justice literature on these topics relates to police. Even in the policing field, little serious research has been conducted about specific management or leadership behaviors; but the general view seems to be that police *management* has gotten much more sophisticated over the last several decades, while police *leadership* has perhaps gotten less attention and emphasis (Stamper 1992).

The focus of leadership and management might be internal or external (Cordner, Scarborough, & Sheehan 2004). The first involves internal matters in the organization and focuses on the organization itself, on such concerns as employees, tasks, rules, supervision, jobs, collective bargaining, workload, efficiency, and budget. In this role, the manager or leader is concerned with making sure that the organization or unit functions correctly.

The second focus is externally-oriented. The environments of criminal justice organizations present all kinds of demands, including both routine and unusual requests for services; inquiries about the handling of specific incidents; pressure from communities for more or different services; directives from the judiciary; and state-mandated minimum requirements for training. The manager or leader is responsible for managing the interaction between the organization and the environment so that the unit or organization is protected and successful. More ambitiously, police and other criminal justice administrators try to lead their environments by influencing community and political perspectives about crime and disorder (Reiss 1985).

The internal and external roles are both important but certainly vary by level in the organization. Top executives and, ironically, line-level officers (especially beat officers) are probably most engaged in managing and leading the external environment, while mid-level

managers and supervisors are generally most engaged in managing and leading within the organization itself.

A recent study of the Los Angeles Police Department emphasized the difficult balancing act required of most public administrators, including police chiefs (Reese 2005). On the one hand, police chiefs need to convince organization members (police officers) that the chief will support them in their difficult jobs. On the other hand, a chief must be responsive to politicians and the community—external constituencies that often complain about police misbehavior or ineffectiveness and want changes made. Failure to satisfy the “troops” can lead to low morale and interfere with a chief’s ability to get the organization to follow his or her path. Failure to satisfy external constituencies can lead to political and public relations problems and to the chief’s replacement. Few police chiefs find it easy to maintain this balancing act and to keep both internal and external audiences satisfied.

A recent book on police leadership stresses the importance of ethics, integrity, and training and considers the applicability of a wide range of general leadership theories to police organizations, including team theory, leader-member exchange theory, transformational and transactional leadership, leadership styles, situational leadership, contingency theory, path-goal theory, psychodynamics, the skills approach, trait theory, and anti-charismatic leadership (Haberfeld 2006). Likewise, leadership programs currently offered by the International Association of Chiefs of Police emphasize the applicability of leadership theories and concepts largely derived from business and the military (Prince, Halstead, & Hesser 2003). Similar approaches are taken in the field of corrections (Freeman 1999; Houston 1999; Seiter 2002). It would seem that in the current situation, there is not much unique content in the specific domain of criminal justice leadership, but rather a significant effort to apply research and theory from the more general study of public and private organizations.

The NCLETTTC leadership training course focused primarily on internal leadership practices. The training emphasized ethics and integrity and drew on various theories including theories of leadership styles, team-building, counseling, and organizational climate and culture in relation to leadership practice. In particular, the training was intended to teach participants the knowledge (evaluation Level 2) needed to improve their leadership activities (evaluation Level 3) in a manner conducive to promoting positive organizational outcomes (evaluation Level 4).

With regard to the latter, NCLETTTC officials agreed that two measurable and reasonable constructs for positive outcomes at the organizational level included employee (1) job satisfaction and (2) organizational commitment within the agencies that sent leaders to the training. That is, good leadership training should be associated with effective leadership practice that leads subordinates to be more satisfied with their jobs and more committed to the organizations for which they work. In fact, research indicates that groups of staff who have greater job satisfaction and organizational commitment tend to exhibit better job performance toward the accomplishment of organizational objectives (Lambert, Barton, & Hogan 1999).

Research Questions

The purpose of this training evaluation was to apply the expanded Kirkpatrick model to NCLETTTC's leadership training program as a means of advancing training evaluation research in the criminal justice field. Two research questions follow logically from this purpose. Though obviously interrelated, the first is more general and oriented around advancement of knowledge in the field, whereas the second is more specific to the NCLETTTC program.

- What can be learned from applying the evaluation model to the leadership program that will add to the knowledge base about model evaluations of criminal justice training activities?
- To what extent was the NCLETTTC leadership program associated with positive outcomes in the domains of: (a) trainee satisfaction with various facets of training; (b) trainee learning, as assessed in terms of pre to post-knowledge gain; (c) trainee behavior change, as assessed through both trainee self ratings and ratings by their supervisors; and (d) organizational changes, assessed in terms of employee job satisfaction and organizational commitment? How did traditional and online training modalities compare in this regard?

Method

The evaluation method for this project included the development/adoption and administration of instruments to measure: (a) trainee demographic and background characteristics, (b) participant reaction to training, (c) pre to post change in learning, (d) changes in job performance as perceived by both trainees and their supervisors, and (e) organizational changes in employee job satisfaction and organizational commitment.

Trainings and Participants

During the study, four leadership courses were taught to 76 trainees in the traditional format. (Note: 100-150 trainees were initially expected to enroll in the traditional classes.) The first traditional class, which was taught March 11-12, 2004, in Charleston, South Carolina, had 19 trainees enrolled. The second traditional class, taught June 14-15, 2004, in Roanoke, Virginia, had 29 persons enrolled. The third traditional class was taught December 9-10, 2004, in Kearneysville, West Virginia, and had 19 enrollments. The fourth traditional class, taught March 10-11, 2005, in Cincinnati, Ohio, enrolled nine trainees.

Wherever possible, this study attempted to use all available data gathered from the above-mentioned measures. In the following narrative and tables, n sizes may vary slightly, due to the following reasons: (1) enrolled training participants having to leave training early due to an emergency, thus failing to complete certain Level 1 and 2 instruments; (2) the same training participants completing follow-up Level 3 and 4 instruments; and (3) a few training participants completing some, but not all, follow-up instruments.

Two observers from the study evaluation team were present at the first course in Charleston. One study team observer from the team was present at the Roanoke and Kearneysville courses. No observers were present at the fourth course. (Because of the low number of enrollments for that course, a member of the research team arranged for the course instructor to collect data at the Cincinnati site.)

All courses enrolled law enforcement and correctional officers (i.e., sheriff's department detention officers) and included line officer, corporal, sergeant, and supervisory/managerial ranks. In the three courses attended by a study team member, training was conducted in a classroom setting. The instructor used a combination of lecture, case study, and role-play delivery strategies.

Although it was anticipated that about 150 trainees would enroll in the online version of the leadership course, by the conclusion of the evaluation period only 89 trainees had enrolled, and only 27 trainees had actually completed the course. The low number of enrollments and completions in the online course was primarily due to it beginning significantly behind schedule. Reasons for its delay are mentioned later in this report.

Immediately prior to the beginning of the courses, trainees completed the demographic form developed by the project team. This form appears in Appendix 9-B.

Summary results from the demographic form appear in Exhibit 9-1 for trainees who completed the course traditionally (n=76) as well as those who (a) enrolled online (n=89) and (b) actually completed the course online (n=27). As the table reveals, the majority of trainees in the four classes delivered traditionally were male (85.5 percent), white (86.8 percent), and married (75.0 percent). These proportions were similar for the online class, except that a higher proportion of online trainees were female (46.1 percent). In terms of age (not shown in Exhibit 9-8), traditional trainees ranged from 24 to 51 years, with a mean age of 37.3 years (SD = 6.66). Online trainees ranged in age from 24 to 61 years; and on average, online trainees were slightly older (mean age = 40.2 years, SD = 7.99). Of the 72 traditional trainees who responded to Item 4 in Exhibit 9-1, all had completed high school or the GED; fewer of these trainees responded to the remaining items dealing with education (Items 5-11). Similarly, nearly all of the online trainees had completed high school or the GED, and over half held a bachelor's degree. Over 42 percent of the traditional trainees who responded had served in the armed forces, compared to only a quarter of the online trainees.

Less than one-third (31.6 percent) of traditional trainees said they were completing the training to attain a job promotion, compared to nearly 43 percent of online trainees. At the time of training, over half (55.3 percent) of traditional trainees were employed in city police departments, and nearly half (47.4 percent) held the rank of sergeant. By contrast, the majority (88.8 percent) of online trainees were correctional employees. Fewer online than traditional trainees were managers, but more were officers. Trainees were also asked how long they had worked in their current positions. For those traditional trainees who responded (n=73), the number of months in the current position ranged from 0 to 120, with a mean of 15.1 months (SD = 25.53) and a median of 5 months. For online trainees (n=89), months in current position ranged from 0 to 24, with a mean of 7.44 months (SD = 5.74) and a median of 6.5 months. Likewise, respondents in traditional classes (n=69) varied in the total number of years they had been employed in leadership positions, ranging from 0 to 20 years, with a mean of 6.7 years (SD = 5.08) and a median of 6.0 years. For respondents in the online class (n=89), years employed in leadership positions ranged from 0 to 35, with a mean of 8.3 years (SD = 7.51) and a median of 8.0 years.

There were minimal differences between those who enrolled in the online course and those that actually completed the online course. The only apparent differences were that a higher proportion of those who completed the online course had masters and professional degrees, as well as less experience in the armed forces.

Exhibit 9-1: Trainee Demographics and Backgrounds

Variable	All Trad. Classes n (%)	Trad. Class 1 n (%)	Trad. Class 2 n (%)	Trad. Class 3 n (%)	Trad. Class 4 n (%)	Online Class Enrolled n (%)	Online Class Completed n (%)	All Classes n (%)
1. Gender								
N	76	19	29	19	9	89	27 ³⁷	165
Male	65 (85.5)	16 (84.2)	25 (86.2)	16 (84.2)	8 (88.9)	48 (53.9)	15(55.6)	113 (68.5)
Female	11 (14.5)	3 (15.8)	4 (13.8)	3 (15.8)	1 (11.1)	41 (46.1)	12(44.4)	52 (31.5)
2. Ethnicity								
N	76	19	29	19	9	89	27	165
White/Caucasian	66 (86.8)	13 (68.4)	26 (89.7)	18 (94.7)	9 (100)	69 (77.5)	22(81.5)	135 (81.9)
African American	8 (10.5)	5 (26.3)	3 (10.3)	0 (0.0)	0 (0.0)	16 (18.0)	3(11.1)	24 (14.5)
Hispanic	2 (2.6)	1 (5.3)	0 (0.0)	1 (5.3)	0 (0.0)	1 (1.1)	1(3.7)	3 (1.8)
Asian	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	2 (2.2)	1(3.7)	2 (1.2)
Other	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (1.1)	0(0.0)	1 (0.6)
3. Marital Status								
N	76	19	29	19	9	89	27	165
Single	9 (11.8)	2 (10.5)	4 (13.8)	3 (15.8)	0 (0.0)	13 (14.6)	4(14.8)	22 (13.3)
Married	57 (75.0)	17 (89.5)	20 (69.0)	12 (63.2)	8 (88.9)	66 (74.2)	19(70.4)	123 (74.5)
Divorced	10 (13.2)	0 (0.0)	5 (17.2)	4 (21.1)	1 (11.1)	10 (11.2)	4(14.8)	20 (12.1)
4. High School Grad/GED								
N	72	18	28	17	9	89	26	161
Yes	72 (100)	18 (100)	28 (100)	17 (100)	9 (100)	87 (97.8)	26(100)	159 (98.8)
No	0 (0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	2 (2.2)	0(0.0)	2 (1.2)
5. Associate's Degree								
N	22	4	7	6	5	89	26	111
Yes	11 (50)	4 (100.0)	5 (71.4)	1 (16.7)	1 (20.0)	28 (31.5)	7(26.9)	39 (35.1)
No	11 (50)	0 (0.0)	2 (28.6)	5 (83.3)	4 (80.0)	61 (68.5)	19(73.1)	72 (64.9)
6. Bachelor's Degree								
N	31	7	9	9	6	89	26	120
Yes	16 (51.6)	4 (57.1)	8 (88.9)	3 (33.3)	1(16.7)	49 (55.1)	15(57.7)	65 (54.2)
No	15 (48.4)	3 (42.9)	1 (11.1)	6 (66.7)	5 (83.3)	40 (44.9)	11(42.3)	55 (45.8)
7. Master's Degree								
N	19	3	5	6	5	89	26	108
Yes	1 (5.3)	0 (0.0)	1 (20.0)	0 (0.0)	0 (0.0)	9 (10.1)	5(19.2)	10 (9.3)
No	18 (94.7)	3 (100.0)	4 (80.0)	6 (100.0)	5 (100)	80 (89.9)	21(80.8)	98 (90.7)

³⁷ Demographic data from one of the 27 trainees who completed the online course was missing after variable 3 (i.e., marital status).

Exhibit 9-1: Trainee Demographics and Backgrounds (continued)

Variable	All Trad. Classes n (%)	Trad. Class 1 n (%)	Trad. Class 2 n (%)	Trad. Class 3 n (%)	Trad. Class 4 n (%)	Online Class n (%)	Online Class Completed n (%)	All Classes n (%)
8. Doctorate Degree								
N	16	3	3	5	5	89	26	---
Yes	0 (0)	0 (0.0)	0 (0.0)	0 (0.0)	0(0.0)	0(0.0)	0(0.0)	
No	16 (100)	3 (100.0)	3 (100.0)	5 (100.0)	5 (100.0)	89(100.0)	26(100.0)	
9. Vocational School								
N	32	4	10	12	6	89	26	---
Yes	22 (68.8)	1 (25.0)	7 (70.0)	9 (75.0)	5 (83.3)	16 (18.0)	6(23.1)	
No	10 (31.3)	3 (75.0)	3 (30.0)	3 (25.0)	1 (16.7)	73 (82.0)	20(76.9)	
10. Professional Degree								
N	17	3	3	6	5	89	26	---
Yes	3 (17.6)	0 (0.0)	1 (33.3)	2 (33.0)	0 (0.0)	6 (6.7)	3(11.5)	
No	14 (82.4)	3 (100.0)	2 (66.7)	4 (66.7)	5 (100.0)	83 (93.3)	23(88.5)	
11. Military Schooling								
N	41	10	12	12	7	Missing	Missing	---
Yes	28 (68.3)	7 (70.0)	9 (75.0)	9 (75.0)	3 (42.9)			
No	13 (31.7)	3 (30.0)	3 (25.0)	3 (25.0)	4 (57.1)			
12. Service in Military Active or Reserve Status								
N	71	19	26	18	8	89	26	160
Yes	30 (42.3)	9 (47.4)	10 (38.5)	8 (44.4)	3 (37.5)	22 (24.7)	9(34.6)	52 (32.5)
No	41 (57.7)	10 (52.6)	16 (61.5)	10 (55.6)	5 (62.5)	67 (75.3)	17(65.4)	108 (67.5)
13. Branch of Military								
N	30	9	10	8	3	22	26	---
Army	16 (53.3)	3 (33.3)	7 (70.0)	4 (50.0)	2 (66.7)	13 (59.1)	5(18.5)	
Air Force	6 (20.0)	3 (33.3)	1 (10.0)	2 (25.0)	0 (0.0)	4 (18.2)	2(7.4)	
Marines	6 (20.0)	3 (33.3)	2 (20.0)	1 (12.5)	0 (0.0)	1 (4.5)	1(3.7)	
Other	1 (3.3)	0 (0.0)	0 (0.0)	0 (0.0)	1 (33.3)	4 (18.2)	0(0.0)	
Multiple Branches	1 (3.3)	0 (0.0)	0 (0.0)	1 (12.5)	0 (0.0)	0 (0.0)	0(0.0)	
14. Are you seeking completion of this course to attain a promotion?								
N	76	19	29	19	9	89	27	165
Yes	24 (31.6)	9 (47.4)	12 (41.4)	1 (5.3)	2 (22.2)	38 (42.7)	9(33.3)	62 (37.6)
No	52 (68.4)	10 (52.6)	17 (58.6)	18 (94.7)	7 (77.8)	51 (57.3)	18(66.7)	103 (62.4)

Exhibit 9-1: Trainee Demographics and Backgrounds (continued)

Variable	All Trad. Classes n (%)	Trad. Class 1 n (%)	Trad. Class 2 n (%)	Trad. Class 3 n (%)	Trad. Class 4 n (%)	Online Class n (%)	Online Class Completed n (%)	All Classes n (%)
15. Type of Employer								
N	76	19	29	19	9	89	26	165
City Police	42 (55.3)	13 (68.4)	11 (37.9)	9 (47.4)	9 (100.0)	0 (0.0)	0(0.0)	42 (25.5)
Sheriff's Office	22 (28.9)	6 (31.6)	12 (41.4)	4 (21.1)	0 (0.0)	3 (3.3)	0(0.0)	25 (15.2)
University Police	6 (7.9)	0 (0.0)	6 (20.7)	0 (0.0)	0 (0.0)	(0.0)	0(0.0)	6 (3.6)
Federal Agency	6 (7.9)	0 (0.0)	0(0.0)	6 (31.6)	0 (0.0)	2 (2.2)	0(0.0)	8 (4.8)
Juvenile Agency	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	3 (3.4)	2(7.7)	3 (1.8)
Dept. Corrections	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	79 (88.8)	23(88.5)	79 (47.9)
Other	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	2 (2.2)	1(3.8)	2 (1.2)
16. Rank								
N	76	19	29	19	9	83	26	159
Officer	17 (22.4)	1 (5.3)	7 (24.1)	5 (26.3)	4 (44.4)	41 (49.4)	13(50.0)	58 (36.5)
Corporal	3 (3.9)	1 (5.3)	0 (0.0)	2 (10.5)	0 (0.0)	0 (0.0)	0(0.0)	3 (1.9)
Sergeant	36 (47.4)	8 (42.1)	17 (58.6)	6 (31.6)	5 (55.6)	35 (42.2)	8(30.8)	71 (44.7)
Management	20 (26.3)	9 (47.4)	5 (17.2)	6 (31.6)	0 (0.0)	7 (8.4)	5(19.2)	27 (17.0)

Design, Instrumentation, and Data Collection

Though optimal for permitting causal inferences, use of an experimental design was impractical in this study. NCLETTTC officials did not believe it was appropriate or feasible to randomly assign trainees to conditions (e.g., traditional versus online instruction, a training group vs. no training control group, etc.) because training participants voluntarily signed up for the course and were then placed in the course on a first-come-first-served basis. Training participants were selected as a result of recruiting from sites based upon contacts in the field. As an alternative, quasi-experimental design (Shadish, Cook, & Campbell 2002) was implemented where possible across the four levels of evaluation. The following sub-sections describe design, measurement, and data gathering at each level.

Level 1

Two methods were used to study students' reactions to the training, including the perceptions of both training participants and researchers. One method was the participant rating instrument appearing in Appendix 9-B and developed by the research team. This instrument asked participants to rate 14 items on a five-point Likert scale ranging from Strongly Agree (5) to Strongly Disagree (1). Items 1-11 addressed participant satisfaction with some important specific components of training, including objectives, organization, instructional methods/presentation, content, relevance, and assessment. Items 12-14 were measures of overall satisfaction and sought to determine whether trainees (a) perceived the positive aspects of training as outweighing the negative ones, (b) would take another similar course in the future, and (c) thought the training would help them perform their jobs more effectively. In addition, the instrument included three open-ended questions asking trainees to describe the most and least beneficial features of the training, as well as what changes they would suggest. This form was administered to participants by the observer and trainer at the conclusion of training. The instrument exhibited very good internal reliability as measured by Cronbach's alpha (.97).³⁸

³⁸ While factor analysis can be used to estimate the validity of a Level 1 rating instrument of the type used in this research, Tabachnick & Fidell (1996) recommend a minimum of 300 cases for factor analysis to produce stable estimates. We had just 99 cases, and factor loadings ranged from .79 to .93. The solution had one factor that explained 75.73 percent of the variance in rating scores. Altering factor extraction techniques did not change this pattern.

The second method involved field observations by the research team—present at three of the four traditional trainings (two researchers attended one of the trainings and two trainings were attended by one researcher). Trainers were rated on a five-category scale (Unacceptable – Poor – Fair – Good – Excellent) in the following areas: organizational skills, delivery methods, teaching strategies, and questioning skills. The percent of agreement averaged across all four trainings rated was 66.67 percent.

Level 2

Knowledge was assessed using pretest – posttest methods with the group undergoing traditional training. The testing instrument appearing in Appendix 9-B was administered before the onset of the two-day training. Immediately after the conclusion of training, trainees were administered the instrument in Appendix 9-B (Post-Test 1). This instrument was designed to serve as a parallel forms test to the pretest instrument because re-administration of the identical pretest two days subsequent to the first administration would have created a testing threat to validity (Cook & Campbell 1979). Similarly, to guard against possible testing effects, a third version of the test was created for administration as a second posttest six months after the end of training (see Appendix 9-B—Post-Test 2). Another reason for development of a second posttest was that many online training courses emphasize “mastery learning” (see Minor et al. 2005) and therefore allow repeated taking of tests. The pretest and first posttest were administered by the trainer and observer. Training participants were allowed 30 minutes to complete the tests. The second posttest was administered via U.S. mail with a cover letter and additional survey.

The process used in constructing the three forms of the test conformed to known principles of good item construction, such as using multiple choice format questions, attempting to construct distracters of equal length and difficulty, etc. (see Minor et al. 2005, pp. 17-26 for detailed coverage of such principles). The test construction process proceeded as follows:

1. The curriculum was reviewed, and the number of objectives and objective categories were identified.
2. A pool of 75 test items was produced from the curriculum’s objectives. (The numbers of items produced for each objective category were commensurate with the length of that objective category).
3. Each item was assessed in relationship to Bloom’s taxonomy (Bloom et al. 1956). (Refer to the discussion in Chapter 3)

4. Two raters then assessed Bloom's taxonomy for each item independently. (Inter-rater agreement was 61 percent between Bloom's six levels and 71 percent between Bloom's taxonomy collapsed into three levels.)
5. Twenty-five items from the pool were randomly assigned to each of the three tests, proportionate to each objective category.
6. The three tests were then checked to ensure that they were approximately equal with respect to Bloom's taxonomy (approximately 40 percent knowledge, 24 percent comprehension, 24 percent application, 8 percent analysis, 4 percent synthesis, and 0 percent evaluation per test).
7. Each test was then checked to make sure there were an approximately equal number of a, b, c, and d response choices.
8. Items for each test were put in random order.

Although Cronbach's alpha is usually used for scores that fall along a continuum, it will produce the same results as KR-20 with dichotomous data (0 or 1). Cronbach's alpha for the pretest, posttest 1 and posttest 2 were, respectively: .314, .359 (inclusive of the online class data; the posttest 1 alpha without the online data was .236), and .578. We deemed it inappropriate to calculate parallel forms reliability between the pretest and either posttest due to the possibility of posttest scores being affected by training. Parallel forms reliability between the two forms of the posttest was .423. This coefficient was statistically significant ($p < .01$) – meaning that the probability is no greater than one percent that the true correlation is zero – but of only moderate strength. The most likely explanation for Cronbach alpha being lower than parallel forms reliability is that Cronbach alpha is often deflated when tests measure several attributes/dimensions rather than one.

The initial study design called for comparing the pre-post knowledge data from the traditional classes with pre-post data from online participants. However, the scores of the pretest taken by online trainees were not recorded because of miscommunication between the online course developer and the research team; the online developer thought the test was merely a “quiz” and did not arrange for scores to be recorded. The miscommunication was clarified before the posttest. But of the 26 persons who had completed the course online by the end of the study, 13 had to be excluded from the analysis of posttest data because they were permitted to take the posttest up to 10 times each, thus creating testing threats to validity.

Level 3

We developed two instruments to study the degree to which trainee behavioral change was associated with the leadership training. One instrument (known as the Leadership Performance Survey or LPS) asked trainees to rate their own behavioral changes, whereas the other one (the Supervisor's Appraisal of Employee Leadership Performance) asked trainees' immediate supervisors to rate changes that they as supervisors perceived in trainee behavior. Both instruments are included in Appendix 9-B.

Items on both instruments were designed to mirror the specific curriculum objectives. Hence, the purpose of the surveys was to determine the extent to which training objectives were associated with behavioral change and to examine this question in a manner that allowed trainee self-reports to be corroborated by reports from their supervisors. However, certain survey items (e.g., "being aware of my own values and biases") called for considerable trainee introspection and, though appropriate for trainees, were judged inappropriate for supervisors. As such, items 12-13, 17-18, 20-25, 30-32, and 34-35 from the LPS were not included on the supervisor's rating instrument.

We were also concerned that some trainees might have undergone changes in supervisors, and the supervisor rating instrument was designed to take this possibility into account. Part A of the instrument was to be completed by those individuals who had supervised the trainee both before training and at least six months thereafter; Part B was to be completed by those who had not supervised trainees within this time frame.

Groups of trainees to form a class were not identified far enough in advance of the start of training to make pre-testing possible. Alternatively, the surveys were administered six months after completion of training, and raters (whether supervisors or trainees themselves) were asked to rate post-training performance in relationship to pre-training performance in terms of post performance being more effective (ratings of 4 or 5), quite similar (ratings of 2 or 3), or less effective (ratings of 0 or 1). In addition, due to their late start dates in relationship to the study's time frame, Level 3 surveys were not administered to the fourth traditional class taught in Cincinnati or to the online class.

Members of traditional classes 1-3 and their supervisors were administered the Level 3 surveys via mail following the total design method developed by Dillman (1978). After sending the proper reminders and follow-up surveys, the following response rates were achieved.

	<i>Trainees (percent)</i>	<i>Supervisors (percent)</i>
<i>Class 1</i>	63.16	83.33
<i>Class 2</i>	65.50	100.0
<i>Class 3</i>	57.90	88.24

Although the response rates for trainees from classes 1-3 are less than desired, all rates shown above are acceptable by social science standards.

The rating instruments in Appendix 9-B demonstrated high internal reliability. Cronbach's alpha for the LPS was .98. The alpha for Part A of the supervisor rating instrument was also .98, and that for Part B was .95.

Level 4

As noted earlier, one of our research questions pertained to whether the NCLETTTC leadership program was associated with positive organizational outcomes in the areas of employee job satisfaction and organizational commitment. Given the curriculum of the leadership training program, it is reasonable to believe that more effective leadership practice should be associated with greater job satisfaction and organizational commitment among employees in the trainee's agency.

Therefore, organizational impacts at Level 4 were measured using two previously validated instruments, the Job Descriptive Index (JDI) as a measure of job satisfaction and the Organizational Commitment Questionnaire (OCQ) as a measure of organizational commitment. The JDI was developed and standardized by Blazer et al. (1997). The JDI is designed to provide separate measures of employee satisfaction with the following aspects of a job: (a) the work itself, (b) pay, (c) promotion opportunities, (d) supervision, (e) co-workers, and (f) the job in general. The possible point range on each of these six measures is 0-54, with higher scores indicative of greater job satisfaction. Blazer et al. (1997) recommend that the various scales be scored separately, rather than pooled into an overall measure of job satisfaction. Cronbach alpha

reliability coefficients for the work (.87), pay (.86), promotion (.87), supervision (.91), coworker (.88), and job in general (.91) scales were all high in the present study.

The OCQ was developed and standardized by Mowday, Steers, & Porter (1979). This 15-item questionnaire requires the respondent to answer each item on a 7-point rating scale ranging from strongly disagree to strongly agree. For each of the statements, commitment is rated on a scale of 1 to 7, with ratings of 1 representing the lowest commitment rating and those of 7 representing the highest commitment level. Cronbach's alpha for the OCQ in this research was .93.

The JDI and OCQ were administered as pretest and posttest instruments to employees from one agency represented by trainees from each of classes 1-3. As in the case of the Level 3 surveys described above, no Level 4 data were obtained from class 4 or the online class due to the late start dates in relationship to the time frame of the research.

A concern about measuring job satisfaction and organizational commitment as Level 4 impacts was that a sufficient number of employees from a particular agency needed to undergo the leadership training to make such organizational change tenable. This concern was addressed for each class as follows.

Class 1

Immediately before the start of class 1, the research team learned that a small 25-officer police department would be sending four of its officers to the training, including all of its emerging leaders (new sergeants). One of the 25 officers was a school safety resource officer. Permission was obtained from the department's chief for participation in the Level 4 evaluation.

Class 2

Immediately before the training for class 2, it was learned that a police department planned to send six of its 39 police officers to the leadership training. These six officers were sergeants, acting sergeants, or officers who would likely become sergeants in the near future. The agency had an additional ten support staff (i.e., two administrative specialists and eight full-time dispatchers). Permission was again obtained from the chief to allow all 39 employees to participate in the research.

Class 3

A week prior to the training for class 3, the research team learned that a city police department would be sending eight of its 51 officers to the training. These eight officers were sergeants, lieutenants, or shift managers. The agency had an additional 13 support staff (i.e., seven clerical or secretarial staff and six full-time dispatchers). The chief was contacted, and permission was obtained to collect Level 4 data.

The pretest versions of the JDI and OCQ were administered on-site at the training locations by the research staff immediately prior to the start of training. The research staff visited the agencies at shift changes and administered the JDI and OCQ to small groups of all available employees over a two-day period. This process was repeated six months after training to collect posttest data. Since the variable of interest was organizational pre to post change, rather than individual pre to post change, it was not necessary to ensure that everyone who completed a posttest had also completed a pretest, or that everyone who had initially taken the pretest also took a posttest. That is, personnel changes in an agency did not threaten validity because the variable of interest was the organization rather than the specific individuals comprising it at any given time.

To examine the implications on organizational outcomes of anyone who trained and was no longer employed six months later, the research team verified the employment status of all trainees who did not complete and return the six-month follow-up posttest. The research team was able to confirm that all but one of the trainees who took the pretest was still employed by their agencies at posttest. Given that only one trainee was no longer employed by the same agency at the time of this study, the research team viewed any potential impact on organizational outcome to be negligible.

Results

Level 1 Reaction Results

It can be seen from Exhibit 9-2 that the means for the traditional class were quite high for each item, with no item having a mean of less than 4.5 across the four classes. The average rating across all items and across all classes (i.e., the mean of means) was 4.6 (SD=0.48). There were no appreciable differences between the four traditional classes; the average ratings across

all items for classes 1, 2, 3, and 4 respectively were: 4.7 (SD=0.38), 4.5 (SD=0.60), 4.5 (SD=0.43), and 4.8 (SD=0.28).

By contrast, the items means for the online participants were lower in magnitude. The average rating across all items for the online class (i.e., the mean of means) was 3.5 (SD=0.87), which was significantly lower than the overall item mean (4.6) of the traditional class, $t(df=32.29) = 6.26, p = .00$.

Exhibit 9-2: Trainee Reaction to Training

Variable	All Trad. Classes	Trad. Class 1	Trad. Class 2	Trad. Class 3	Trad. Class 4	Online Class	All Classes
1. Objectives were clearly set at the beginning of training.							
N	72	19	27	17	9	27	99
Mean	4.5	4.6	4.5	4.3	4.7	3.7	4.3
Standard Deviation	0.53	0.51	0.58	0.47	0.50	1.00	0.78
% Strongly Agree or Agree	98.6	100.0	96.3	100.0	100.0	77.8	92.9
2. Objectives were met by the end of training.							
N	72	19	27	17	9	27	99
Mean	4.6	4.7	4.6	4.5	4.8	3.7	4.3
Standard Deviation	0.55	0.48	0.75	0.51	0.44	0.78	0.74
% Strongly Agree or Agree	97.2	100.0	92.6	100.0	100.0	77.8	91.9
3. The training was well organized.							
N	72	19	27	17	9	27	99
Mean	4.7	4.7	4.6	4.5	4.9	3.6	4.4
Standard Deviation	0.59	0.45	0.75	0.51	0.33	0.93	0.84
% Strongly Agree or Agree	97.2	100.0	92.6	100.0	100.0	77.8	91.9
4. Instructional methods were used effectively.							
N	72	19	27	17	9	27	99
Mean	4.7	4.7	4.6	4.6	4.9	3.6	4.4
Standard Deviation	0.58	0.45	0.75	0.49	0.33	1.12	0.91
% Strongly Agree or Agree	97.2	100.0	92.6	100.0	100.0	74.1	90.9
5. Instructional aids (e.g., graphics, etc.) were beneficial.							
N	71	19	26	17	9	27	98
Mean	4.5	4.5	4.5	4.4	4.6	3.2	4.1
Standard Deviation	0.79	0.77	0.90	0.70	0.73	1.78	1.07
% Strongly Agree or Agree	87.3	86.3	88.5	100.0	88.9	59.3	79.6

Exhibit 9-2: Trainee Reaction to Training (continued)

Variable	All Trad. Classes	Trad. Class 1	Trad. Class 2	Trad. Class 3	Trad. Class 4	Online Class	All Classes
6. Content was presented at a level appropriate to my background and experience.							
N	72	19	27	17	9	27	99
Mean	4.5	4.7	4.4	4.4	4.7	3.4	4.2
Standard Deviation	0.67	0.58	0.84	0.49	0.50	1.15	0.95
% Strongly Agree or Agree	95.8	94.8	92.6	100.0	100.00	66.7	87.9
7. Content was presented in a clear and understandable manner.							
N	72	19	27	17	9	27	99
Mean	4.7	4.8	4.7	4.6	4.9	3.6	4.4
Standard Deviation	0.50	0.37	0.54	0.62	0.33	0.88	0.80
% Strongly Agree or Agree	97.2	100.0	96.3	94.1	100.0	77.8	91.9
8. Topics covered were relevant to actual leadership skills called for in my organization.							
N	72	19	27	17	9	27	99
Mean	4.5	4.6	4.4	4.5	4.8	3.3	4.2
Standard Deviation	0.60	0.61	0.70	0.51	0.44	1.27	0.99
% Strongly Agree or Agree	97.2	94.8	96.3	100.0	100.0	59.3	86.9
9. Exercises meant to practice or demonstrate knowledge and/or skills were used effectively.							
N	71	19	27	16	9	27	98
Mean	4.7	4.8	4.6	4.5	4.9	3.4	4.3
Standard Deviation	0.55	0.42	.056	0.73	0.33	1.08	0.94
% Strongly Agree or Agree	95.8	100.0	96.3	87.5	100.0	63.0	86.7

Exhibit 9-2: Trainee Reaction to Training (continued)

Variable	All Trad. Classes	Trad. Class 1	Trad. Class 2	Trad. Class 3	Trad. Class 4	Online Class	All Classes
10. Content of assessment (e.g., exercises, case study, test, etc.) was aligned with the content of training topics.							
N	72	19	27	17	9	27	99
Mean	4.6	4.6	4.6	4.5	4.8	3.6	4.3
Standard Deviation	0.60	0.50	0.64	0.51	0.44	0.93	0.82
% Strongly Agree or Agree	94.4	100.0	92.3	88.3	100.0	74.1	88.9
11. Difficulty level of methods of assessment (e.g., exercises, case study, test) was aligned with difficulty level of the training topics.							
N	72	19	27	17	9	27	99
Mean	4.5	4.5	4.5	4.5	4.8	3.3	4.2
Standard Deviation	0.58	0.61	0.64	0.51	0.44	1.10	0.94
% Strongly Agree or Agree	95.8	94.7	92.6	100.0	100.0	59.3	85.9
12. Positive aspects of course outweigh its negative aspects.							
N	72	19	27	17	9	27	99
Mean	4.7	4.8	4.5	4.6	5.0	3.6	4.4
Standard Deviation	0.55	0.37	0.64	0.62	0.00	1.09	0.89
% Strongly Agree or Agree	95.8	100.0	92.6	94.1	100.0	70.4	88.9
13. I would take another similar type of course if it were offered (and needed).							
N	72	19	27	17	9	27	99
Mean	4.7	4.8	4.6	4.6	5.0	3.5	4.4
Standard Deviation	0.61	0.37	0.84	0.51	0.0	1.42	1.06
% Strongly Agree or Agree	97.2	100.0	92.6	100.0	100.0	70.4	89.9

Exhibit 9-2: Trainee Reaction to Training (continued)

Variable	All Trad. Classes	Trad. Class 1	Trad. Class 2	Trad. Class 3	Trad. Class 4	Online Class	All Classes
14. This training will help me perform my job more effectively.							
N	72	19	27	17	9	27	99
Mean	4.7	4.8	4.6	4.6	4.9	3.6	4.4
Standard Deviation	0.578	0.42	0.75	0.49	0.33	1.12	0.91
% Strongly Agree or Agree	97.2	100.0	92.6	100.0	100.0	74.1	90.9

The participant evaluation instrument also contained three open-ended items that allowed participants to list what they thought were the most beneficial and least beneficial aspects of the course, as well as changes they would recommend to the course. Slightly over half of the traditional class participants (57.6 percent) as well as the online class participants (50.6 percent) responded to these three questions. With respect to the most beneficial aspects of the traditionally taught course, the most prevalent comments were, in hierarchical order: counseling sessions/scenarios, hands-on activities, class interaction, presentation style of the instructor, and discussion of leadership styles. The most prevalent comments about the least beneficial aspects of the traditional course were, in hierarchical order: none, need for more scenarios, and need for more practical examples. Recommended changes for the traditional course included, in hierarchical order: none, need for more hands-on activities, need to add more scenarios, need to extend the training time, and need to reduce the military flair.

The most beneficial aspects of the online course were, in hierarchical order: usefulness of leadership knowledge gained in course toward career, explanation and video of counseling sessions/scenarios, convenience of the online class, and realism of the examples used. The most prevalent comments about the least beneficial aspects of the online course were, in hierarchical order: none, inconvenience of the online course (e.g., not able to ask instructor questions, downloading of documents very time consuming, etc.), course needs to be less choppy and more fluid, and training oriented too much toward military leadership. Recommended changes for the online course included, in hierarchical order: none, reduce the time needed to download

documents, need to add more police/correctional scenarios and examples, and reduce the military flair.

The second method of studying training delivery involved field observations by the research team. In this way, we tested the instructor rating tool and method. Although we did not apply rigorous methodology to this test, it still served as an observation assessment of training quality. Members of the research team who were present during the first three traditional classes recorded their impressions of training quality. The data in the cells of Exhibit 9-3 refer to the number of courses receiving the various ratings (unacceptable, poor, fair, good, and excellent). There are four ratings for many of the items because one of the three trainings was attended by two members of the research team and rated by both. With few exceptions, the items were rated as either “good” or “excellent,” and no items received “unacceptable” or “poor” ratings. While this table is limited to just four (and in some instances three) observations, the data here generally corroborate the pattern of findings shown in Exhibit 9-2 for the first three traditional courses.

Exhibit 9-3: Observer Ratings

Rating Category/Item	Unacceptable	Poor	Fair	Good	Excellent	% Agreement
<i>Organizational Skills</i>						
1. Stated objectives clearly	0	0	0	1	3	75%
2. Defined new/unusual terms	0	0	0	1	3	75%
3. Explained directions*	0	0	0	0	3	75%
4. Adjusted pace for complexity & understanding*	0	0	0	2	1	50%
5. Stressed important content	0	0	0	3	1	75%
6. Managed time well	0	0	1	2	1	50%
7. Handled materials easily	0	0	0	2	2	50%
8. Linked to previous training*	0	0	0	2	1	50%
<i>Delivery Methods</i>						
1. Anecdotes/examples given	0	0	0	1	3	75%
2. Case study/scenario*	0	0	0	0	3	75%
3. Demonstration*	0	0	0	3	0	75%
4. Facilitation*	0	0	0	2	1	50%
5. Learning activity*	0	0	0	2	1	50%
6. Lecture	0	0	0	3	1	75%
7. Panel discussion*	0	0	0	0	0	NA
8. Power point presentation as outline only*	0	0	0	0	3	75%
9. Interactive power point	0	0	1	1	2	50%
10. Questions and answers	0	0	0	1	3	75%
11. Role playing	0	0	1	0	3	75%
12. Video	0	0	2	1	1	50%
13. Workbook*	0	0	0	1	2	50%
14. Problem based *	0	0	2	1	0	50%
<i>Teaching Strategies</i>						
1. Provided overview of lesson	0	0	0	2	2	50%
2. Provided feedback to learners	0	0	0	2	2	50%
3. Created transition from phase to phase	0	0	0	3	1	75%
4. Conducted periodic review/summary of material	0	0	2	2	0	50%

5. Included examples of references to use on the job	0	0	0	0	4	100%
6. Was able to use hardware and software easily	0	0	0	1	3	75%
<i>Questioning Skills</i>						
1. Answered questions thoroughly and clearly	0	0	0	1	3	75%
2. Questions asked were clear	0	0	0	0	4	100%
3. Put appropriate challenge in questions	0	0	0	3	1	75%
4. Invited questions	0	0	0	0	4	100%
5. Handled incorrect answers supportively*	0	0	0	2	1	50%
6. Used open-ended questions where possible	0	0	0	3	1	75%

*Missing case(s)

Level 2 Knowledge Results

Each question on the knowledge pretest and the two posttests was counted as one point, so that the possible point range for each test was 0-25. Pre-post data are provided in Exhibits 9-4 and 9-5. Panel A of Exhibit 9-4 shows that on average, across all traditional classes, trainees got just over 34 percent of the pretest items correct; no trainee scored higher than 56 percent.

Although the Posttest 1 mean (53.89 percent) is significantly higher than the pretest mean (see Panel B of Exhibit 29), this mean is still low in magnitude. In addition, there was a significant decline between Posttest 1 and Posttest 2 from a mean of 53.89 percent to a mean of 41.30 percent. Dissipation of knowledge gain from Posttest 1 to Posttest 2 is apparent in Exhibit 9-5 and was evident in each class for which data were available. It is also interesting to note that the posttest mean for the online class (64 percent) was somewhat higher than the posttest means of the traditional classes, although the online class mean was still relatively low in absolute terms.

Exhibit 9-4: Pretest and Posttest Data

Panel A: Descriptive Results

Data Source	All Trad. Classes	Trad. Class 1	Trad. Class 2	Trad. Class 3	Trad. Class 4	Online Class	All Classes
Pretest							
N	74	17	29	19	9	Not Applicable	Not Applicable
Mean	34.27%	35.29%	37.93%	35.58%	17.78%		
SD	0.11	0.10	0.08	0.10	0.05		
Range	8 – 56%	12 – 52%	16 – 52%	16 – 56%	8 – 24%		
Posttest 1							
N	74	20	28	17	9	20	94
Mean	53.89%	52.20%	54.14%	54.59%	55.56%	64.00%	57.74%
SD	0.09	0.10	0.10	0.06	0.11	0.12	0.11
Range	32 – 76%	32 – 76%	40 – 72%	48 – 68%	32 – 72%	44 – 88%	32-88%
Posttest 2							
N	43	13	19	11	Not Applicable	Not Applicable	Not Applicable
Mean	41.30%	39.69%	42.11%	41.82%			
SD	0.13	0.15	0.14	0.11			
Range	8 – 68%	8 – 68%	16 – 64%	32 – 64%			

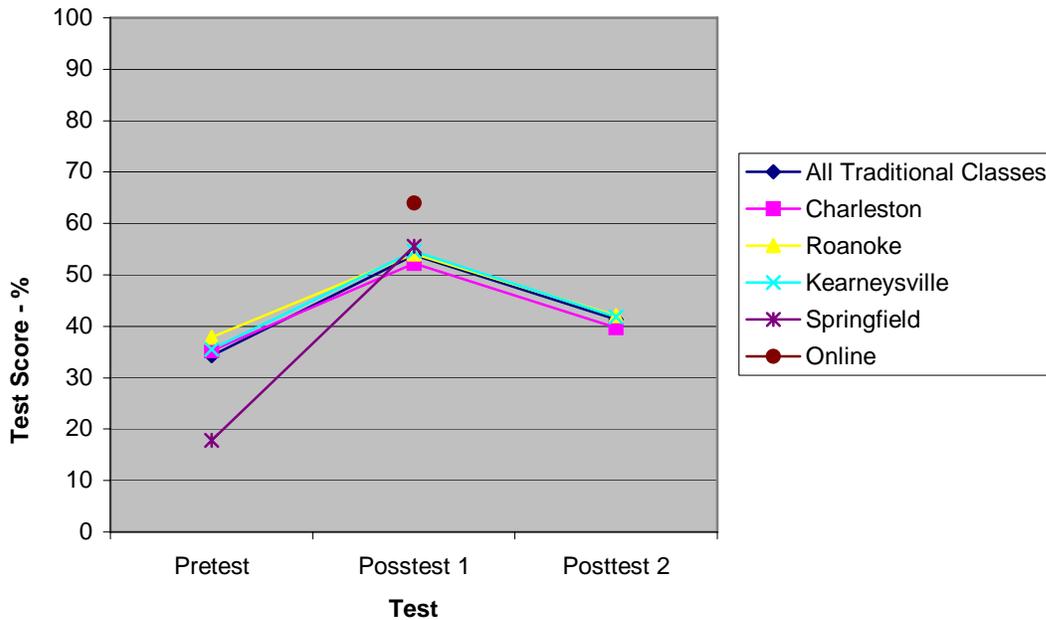
Panel B: Inferential Results

Data Source	All Trad. Classes	Trad. Class 1	Trad. Class 2	Trad. Class 3	Trad. Class 4	Online Class
Difference between Pretest & Posttest 1	t=11.86 df=146 p < .0001	t=5.19 df=35 p < .0001	t=6.50 df=51.77 p < .0001 ^b	t=6.93 df=34 p < .0001	t=9.01 df=16 p < .0001	Not Applicable
Difference between Posttest 1 & Posttest 2	t=3.129 df=115 p < .01	t=.959 df=28 p=.346	t=1.20 df=26.85 p=.242 ^b	t=1.59 df=28 p=.124	Not Applicable	Not Applicable

^a In order to use all available data and to avoid limiting analyses to only those training participants who completed all three tests, independent sample t-tests rather than matched samples t-tests were employed.

^b Equal variances not assumed in calculation of t-test

Exhibit 9-5: Pre to Post Differences by Class



Level 3 Behavior Change Results

Leadership Performance Survey (LPS) – Trainees’ Follow-up

This survey asked trainees to rate their own behavioral changes following training, based on a comparison to what their performance had been like prior to training. Ratings were made on a 0-5 scale, with higher scores indicating positive behavioral change ratings.

Trainee responses to the LPS items are summarized in Exhibit 9-6 for the three traditional classes from which data were obtained. When examining the table, recall that ratings of 0-1 indicated less effective post-training leadership performance, ratings of 2-3 suggested similar pre and post performance, while those of 4-5 indicated more effective post-training performance. As can be seen, the table contains no means that exceed 4.0 and only four means of 4.0, all of which were for class 1. By the same token, close to two-thirds of all the item means are 3.5 or higher, indicating at least some progress.

The overall means for the LPS data set (i.e., the means of means in Exhibit 9-6) were as shown below with the grand mean rating across classes of 3.51.

	<i>Mean</i>	<i>SD</i>
<i>Class 1</i>	3.63	0.30
<i>Class 2</i>	3.60	0.59
<i>Class 3</i>	3.30	0.87

These mean scores indicate that, on average when compared with pre-training performance, trainees rated their post-training performance as about midway between “quite similar” and “more effective.” This suggests that though not as much progress was evident as might have been desired, some behavior change did transpire, as rated by trainees in those areas targeted by the training curriculum.

Exhibit 9-6: Trainees' Leadership Performance Survey Ratings

Survey Item	Class 1 N=11	Class 2 N=19	Class 3 N=11
1. Addressing the human dimension of leadership (discipline, morale, taking care of people)			
Mean	3.5	3.7	3.2
SD	0.71	0.73	0.87
2. Preparing for and dealing with mission stress (hostile/violent situations)			
Mean	3.7	3.3	3.1
SD	0.82	0.65	1.04
3. Preparing for and dealing with the stress of change (new laws, technology, etc.)			
Mean	3.3	3.4	3.1
SD	0.82	0.69	0.83
4. Building positive organizational climate (short-term environment)			
Mean	3.5	3.6	3.5
SD	0.97	0.84	1.04
5. Building positive agency culture (long-term environment)			
Mean	3.5	3.7	3.5
SD	0.85	0.73	1.13
6. Mixing leadership styles successfully (participating, delegating, transformational, & transactional)			
Mean	3.6	3.6	3.4
SD	0.52	0.76	1.03
7. Anticipating both the intended and unintended consequences of my decisions and actions			
Mean	3.9	3.6	3.3
SD	0.74	0.76	0.90
8. Practicing developmental counseling with subordinates			
Mean	3.8	3.8	3.1
SD	0.42	0.71	0.94
9. Developing leadership capabilities among subordinates			
Mean	4.0	3.8	3.6
SD	0.47	0.71	1.12
10. Conducting performance reviews with subordinates (strengths, weaknesses, and plans of action)			
Mean	4.0	3.8	3.5
SD	0.82	0.79	1.13
11. Mentoring subordinates			
Mean	3.6	3.6	3.5
SD	0.70	0.77	1.04

Exhibit 9-6: Trainees' Leadership Performance Survey Ratings (continued)

Survey Item	Class 1 N=11	Class 2 N=19	Class 3 N=11
12. Clearly defining the purpose of counseling sessions			
Mean	4.0	3.7	3.1
SD	0.67	0.81	0.83
13. Fitting counseling style to the subordinate/situation in question			
Mean	3.6	3.7	3.3
SD	0.70	0.75	1.01
14. Demonstrating respect for subordinates			
Mean	3.6	3.7	3.3
SD	0.52	0.75	0.90
15. Establishing/maintaining communication with subordinates			
Mean	3.6	3.6	3.4
SD	0.52	0.76	1.03
16. Supporting subordinates			
Mean	3.6	3.5	3.4
SD	0.70	0.77	1.12
17. Being aware of my own values and biases			
Mean	3.6	3.7	3.5
SD	0.70	0.75	1.13
18. Displaying empathy in dealings with subordinates			
Mean	3.1	3.4	3.0
SD	0.74	0.76	0.77
19. Establishing my credibility by being honest and consistent			
Mean	3.7	3.5	3.4
SD	0.95	0.77	1.03
20. Engaging in active listening			
Mean	3.7	3.7	3.4
SD	0.67	0.75	0.92
21. Paying attention to the gestures of subordinates			
Mean	3.8	3.8	3.2
SD	0.63	0.79	0.87
22. Responding verbally and non-verbally to subordinates during counseling			
Mean	3.9	3.5	3.1
SD	0.57	0.77	0.94

Exhibit 9-6: Trainees' Leadership Performance Survey Ratings (continued)

Survey Item	Class 1 N=11	Class 2 N=19	Class 3 N=11
23. Asking subordinates the proper number and type of questions			
Mean	3.4	3.5	3.1
SD	0.70	0.61	0.83
24. Avoiding counseling errors (rash judgments, inflexibility, etc.)			
Mean	4.0	3.8	3.5
SD	0.67	0.85	1.04
25. Being aware of my limitations as a leader			
Mean	3.6	3.7	3.4
SD	1.17	0.75	1.03
26. Making referrals as needed			
Mean	3.7	3.5	3.4
SD	0.67	0.70	1.03
27. Counseling in relation to a specific event or situation (substandard performance, new employee, crisis, etc.)			
Mean	3.8	3.6	3.3
SD	0.42	0.83	1.01
28. Counseling to improve subordinate performance			
Mean	3.7	3.8	3.3
SD	0.48	0.71	1.01
29. Counseling to promote professional growth of subordinates			
Mean	3.7	3.7	3.4
SD	0.67	0.82	1.03
30. Practicing nondirective counseling			
Mean	3.3	3.2	3.4
SD	0.67	0.79	1.03
31. Practicing directive counseling			
Mean	3.3	3.3	3.1
SD	0.67	0.67	0.83
32. Being able to combine nondirective and directive counseling			
Mean	3.2	3.5	3.5
SD	0.42	0.77	1.04
33. Identifying the need for counseling			
Mean	3.6	3.6	3.3
SD	0.52	0.84	1.01

Exhibit 9-6: Trainees' Leadership Performance Survey Ratings (continued)

Survey Item	Class 1 N=11	Class 2 N=19	Class 3 N=11
34. Preparing for counseling (scheduling, notifying, organizing)			
Mean	3.8	3.5	3.2
SD	0.63	0.84	0.98
35. Conducting counseling sessions (opening, developing a plan, closing)			
Mean	3.5	3.5	3.4
SD	0.85	0.84	1.03
36. Following-up after counseling sessions			
Mean	3.4	3.4	3.3
SD	0.70	1.12	1.01

An open-ended item was included on the LPS to ascertain whether trainees perceived barriers or factors that may have interfered with the transfer of skills learned from the leadership training to their workplace. Of the two participants (one from class 2 and one from class 3) who indicated there was such a barrier, only one (from class 3) described the nature of that barrier (i.e., “limited authority, support”).

An additional open-ended question asked trainees to indicate any other factors besides the leadership training that might have affected their leadership skills since completion of training. Ten of the 42 participants indicated there were other factors. Descriptions of those factors included: experience (two participants from class 2, one participant each from classes 1 and 3), other training (one participant from class 1 and two from class 2), increased workload due to elimination of positions (class 1), an incident of disciplining subordinates (class 1), and subordinates being in the same training as the participant (class 2).

Supervisors' Follow up Ratings of Trainees' Leadership Performance

As discussed earlier, while the survey instrument for trainees' immediate supervisors was similar to the one administered to trainees themselves, the former instrument excluded certain items that called for trainee introspection and left little or no basis for an objective supervisor appraisal. Both surveys were administered six months after training completion. The

supervisors who provided rating data varied in rank and title (e.g., sergeant, lieutenant, captain, deputy chief, chief, and mayor).

Recall further that the supervisor survey was designed to allow for changes that may have occurred with respect to the trainee's supervisor. Forty-four respondents to the supervisor survey indicated that they had supervised trainees an average of 24.77 months (SD = 26.72) prior to completion of the training as well as an average of 5.91 months (SD = 1.14) after the training; these supervisors are henceforth referred to as "Supervisors A." Ten supervisors indicated that they had not supervised the trainee before training and only an average of 4.15 months (SD = 1.80) since completion of the training; these supervisors are henceforth referred to as "Supervisors B." As in the case of the survey administered to trainees, post-training performance was rated by supervisors on a scale of 1 to 5, with ratings of 0 - 1 representing "less effective," 2 - 3 "quite similar," and 4 - 5 "more effective" in comparison to pre-training performance.

Item responses from both types of supervisors are summarized in Exhibit 9-7. Excluding the far right column (where n=1), Exhibit 9-7 contains relatively few item means of 4.0 or greater. However, as with Exhibit 9-6, the majority of means are 3.5 or higher, indicating some progress as perceived by supervisors.

The overall item means are shown below. Exclusive of the Supervisor B data from Class 3 (based on n=1), there was little difference between the Supervisor A and B data. These overall mean scores indicate that, on average when compared with pre-training performance, supervisors generally rated the post-training performance of trainees as roughly midway between "quite similar" and "more effective." This is consistent with the pattern of findings from the trainee survey and implies that though not as much progress was evident as might have been desired, some behavior change did transpire, as rated by trainees' supervisors in those areas targeted by the training curriculum.

	<i>Mean</i>	<i>SD</i>
<i>Class 1</i>		
<i>Sup. A</i>	3.7	0.63
<i>Sup. B</i>	3.8	0.30
<i>Class 2</i>		
<i>Sup. A</i>	3.5	0.64
<i>Sup. B</i>	3.5	0.56
<i>Class 3</i>		
<i>Sup. A</i>	3.2	0.44
<i>Sup. B</i>	4.3	-----

Exhibit 9-7: Supervisors' Leadership Performance Survey Ratings

Survey Item	Class 1	Class 1	Class 2	Class 2	Class 3	Class 3
	Sup. A N=10	Sup. B N=6	Sup. A N=18	Sup. B N=5	Sup. A N=14	Sup. B N=1
1. Addressing the human dimension of leadership (discipline, morale, taking care of people) Mean SD	3.9 0.57	3.8 0.41	3.5 0.92	3.4 1.14	3.4 0.63	5.0 ---
2. Preparing for and dealing with mission stress (hostile/violent situations) Mean SD	3.7 0.67	4.2 0.41	3.6 0.70	3.8 0.45	3.4 0.50	5.0 ---
3. Preparing for and dealing with the stress of change (new laws, technology, etc.) Mean SD	4.0 0.67	3.8 0.41	3.6 0.78	3.4 0.89	3.0 0.68	4.0 ---
4. Building positive organizational climate (short term environment) Mean SD	3.7 0.82	3.8 0.41	3.5 1.15	3.2 0.84	3.2 0.89	4.0 ---
5. Building positive agency culture (long term environment) Mean SD	3.8 0.63	3.8 0.41	3.4 0.98	3.2 0.84	3.1 0.86	4.0 ---
6. Mixing leadership styles successfully (participating, delegating, transformational, and transactional leadership) Mean SD	3.5 1.27	4.0 0.00	3.7 0.77	3.2 0.84	3.1 0.86	4.0 ---
7. Anticipating both the intended and unintended consequences of my decisions and actions Mean SD	3.8 0.79	3.7 0.52	3.4 0.61	3.8 0.84	3.4 0.51	4.0 ---
8. Practicing developmental counseling with subordinates Mean SD	3.6 0.84	3.5 0.84	3.6 0.92	3.0 0.71	3.3 0.73	3.0 ---

Exhibit 9-7: Supervisors' Leadership Performance Survey Ratings (continued)

Survey Item	Class 1	Class 1	Class 2	Class 2	Class 3	Class 3
	Sup. A N=10	Sup. B N=6	Sup. A N=18	Sup. B N=5	Sup. A N=14	Sup. B N=1
9. Developing leadership capabilities among subordinates						
Mean	3.5	3.7	3.4	3.0	3.3	4.0
SD	0.71	0.82	0.85	0.71	0.61	---
10. Conducting performance reviews with subordinates (strengths, weaknesses, and plans of action)						
Mean	3.5	3.8	3.6	3.4	3.3	4.0
SD	0.85	0.41	0.92	0.89	0.73	---
11. Mentoring subordinates						
Mean	3.6	3.7	3.6	3.6	3.1	5.0
SD	0.84	0.52	0.78	1.14	0.53	---
12. Demonstrating respect for subordinates						
Mean	3.8	3.8	3.3	3.8	3.3	5.0
SD	1.03	0.41	0.91	0.84	0.47	---
13. Establishing/maintaining communication with subordinates						
Mean	4.0	3.8	3.6	3.8	3.3	5.0
SD	0.82	0.41	0.92	0.84	0.47	---
14. Supporting subordinates						
Mean	3.9	4.0	3.7	4.0	3.5	5.0
SD	0.88	0.00	1.08	0.00	0.52	---
15. Establishing credibility by being honest and consistent						
Mean	3.6	4.2	3.7	4.2	3.4	5.0
SD	0.84	0.41	1.08	0.45	0.76	---
16. Making referrals as needed						
Mean	3.9	4.0	3.6	3.6	3.0	4.0
SD	0.57	0.63	0.85	0.55	0.55	---
17. Counseling in relation to a specific event or situation (substandard performance, new employee, crisis, etc.)						
Mean	3.7	3.8	3.7	3.2	3.2	4.0
SD	0.82	0.41	0.84	0.84	0.80	---

Exhibit 9-7: Supervisors' Leadership Performance Survey Ratings (continued)

Survey Item	Class 1 Sup. A N=10	Class 1 Sup. B N=6	Class 2 Sup. A N=18	Class 2 Sup. B N=5	Class 3 Sup. A N=14	Class 3 Sup. B N=1
18. Counseling to improve subordinate performance						
Mean	4.1	3.8	3.5	3.2	3.3	4.0
SD	0.57	0.41	0.71	0.84	0.73	---
19. Counseling to promote professional growth of subordinates						
Mean	4.1	3.7	3.6	3.4	3.3	4.0
SD	0.57	0.52	0.62	0.89	0.83	---
20. Identifying the need for counseling						
Mean	3.4	4.0	3.4	3.2	3.1	4.0
SD	0.84	0.63	0.09	0.84	0.66	---
21. Following-up after counseling sessions						
Mean	3.4	3.7	3.3	3.2	2.9	4.0
SD	0.84	0.52	0.69	0.84	0.66	---

Supervisors were also asked open-ended questions. Only three of the supervisors indicated that they were aware of any barriers that may have interfered with the transfer of skills learned in the NCLETTTC leadership training to the trainee's workplace. One supervisor, from class 1, indicated that the subordinate had been intentionally run over by a vehicle and was out of work for several months due to injuries. One supervisor from class 2 mentioned that trainee currently did not have any counseling responsibilities. One other supervisor from class 2 mentioned "personal strife" as a barrier.

Seven supervisors (one from class 1 and 6 from class 2) indicated that they were aware of other factors besides the leadership training that may have affected the trainee's leadership skills since completion of training. Explanations were provided by six supervisors (all from class 2). Explanations included: additional supervisory training, the trainee took some time off (4-5

weeks), other more experienced supervisors had counseled the trainee, and the trainee required additional counseling and supervision.³⁹

Level 4 Organizational Impact Results

At Level 4, the research question was whether NCLETTTC leadership training was associated with increases in employee job satisfaction and organizational commitment at the level of the organization where trainees worked. To this end, the Job Descriptive Index (JDI) and Organizational Commitment Questionnaire (OCQ) were administered as pretest and posttest instruments to employees from one agency represented by trainees from each of traditional classes 1-3. The results are presented below according to class.

The possible point range on the JDI is 0-54, with higher scores indicating greater satisfaction. The authors of the JDI (Blazer et al. 1997, p. 26) observe that “scores well above 27 (i.e., 32 or above) indicate satisfaction, while those well below 27 (i.e., 22 or below) indicate dissatisfaction.” Scores above 22 and below 32 fall in the neutral range. Scale items left blank were omitted from overall scale analyses and results presented in the tables below. The OCQ is scored by calculating mean ratings for each item on a 1-7 scale (with higher ratings indicating greater commitment) as well as by calculating an overall item mean or mean of means.

Class 1

JDI data for the agency represented in class 1 are summarized in Exhibit 9-8. The table reveals that at pre-testing, employees from the police department studied were satisfied with their work, supervision, co-workers, and job in general. They were dissatisfied with pay and neutral on the promotion measure. At post-testing, employees also exhibited satisfaction on all scales except pay (where they were neutral) and promotion (where they were dissatisfied). Although

³⁹ As a means of estimating the predictive validity of Level 2 test scores, we analyzed the relationship between these scores and supervisor ratings. Pearson correlations were negligible. Correlations between Supervisor A ratings and pretest, Posttest 1, and Posttest 2 scores were -.052 (n=35), -.322 (n=36), and -.160 (n=26) respectively. Correlations between Supervisor B ratings and pretest, Posttest 1, and Posttest 2 scores were -.089 (n=12), -.528 (n=11), and -.279 (n=6). None of these correlations were statistically significant. However, neither pretest and posttest distributions nor supervisor rating scores exhibited much variability (as evidenced by the standard deviation data), and the numbers of cases for these analyses were rather low. The result of this situation is unstable estimates of predictive validity.

pre to post declines were evident on all but one scale (pay), the declines were not large. None of the mean differences were statistically significant, as measured by t-tests.

Exhibit 9-8: Pre – Post Job Satisfaction Data – Traditional Delivery (Class 1)

Instrument	JDI Work	JDI Pay	JDI Promotion	JDI Supervision	JDI Co-workers	JDI Job in General
Pretest						
N	26	27	26	26	26	24
Mean	45.31	21.78	22.38	38.38	39.23	47.04
SD	7.33	15.54	14.87	14.11	11.29	6.00
Posttest						
N	31	31	31	24	23	30
Mean	43.29	23.03	19.74	33.71	38.61	46.13
SD	10.97	16.09	12.04	14.15	11.48	7.36

Data from the OCQ for the agency represented in traditional class 1 are presented in Exhibit 9-9. Negatively worded items were reverse scored so that higher means are always indicative of greater commitment. The overall pretest item mean (i.e., the mean of pretest means) was 5.71 (SD = .83) indicating moderate to high organizational commitment. By comparison, the overall posttest item mean (i.e., the mean of posttest means) was 5.44 (SD = .78), also indicating moderate to high organizational commitment. Based on data provided by the two survey administrations, there was a very slight, statistically insignificant decrease in overall organizational commitment.

Exhibit 9-9: Pre – Post Organizational Commitment Data, Traditional Delivery (Class 1)

OCQ Item	Pretest	Pretest	Posttest	Posttest
	Mean N=27	SD	Mean N=30	SD
1. I am willing to put in a great deal of effort beyond that normally expected in order to help this organization be successful.	6.26	0.94	6.20	1.03
2. I talk up this organization to my friends as a great organization to work for.	6.11	1.01	5.79	1.03
3. I feel very little loyalty to this organization.	5.67	2.13	5.67	2.07
4. I would accept almost any type of job assignment in order to keep working for this organization.	4.81	1.47	4.63	1.65
5. I find that my values and the organization's values are very similar.	5.33	1.52	5.33	1.37
6. I am proud to tell others that I am part of this organization.	6.04	1.32	6.13	0.97
7. I could just as well be working for a different organization as long as the type of work is similar.	4.52	1.78	3.87	1.74
8. This organization really inspires the very best in me in the way of job performance.	5.37	1.21	4.87	1.28
9. It would take very little change in my present circumstances to cause me to leave this organization.	5.48	1.60	5.47	1.52
10. I am extremely glad that I chose this organization to work for over others I was considering at the time joined.	6.44	0.85	6.07	1.10
11. There's not much to be gained by sticking with this organization indefinitely.	5.56	1.78	4.43	2.14
12. Often, I find it difficult to stick with this organization's policies on important matters relating to its employees.	5.30	1.73	4.30	1.82
13. I really care about the fate of this organization.	6.22	1.37	6.10	1.03
14. To me this the best of all possible organizations for which to work.	5.74	1.43	5.24	1.35
15. Deciding to work for this organization was a definite mistake on my part.	6.74	0.59	6.60	0.86

Class 2

Descriptive JDI data are presented in Exhibit 9-10 for the department represented in class 2. It can be seen that at pre-testing, trainees expressed satisfaction on all scales except pay and promotion (where they were neutral). At post-testing, all scores fell into the satisfaction range, except for the pay score, which remained in the neutral range. Slight pre to post declines are evident on all measures except the promotion measure. None of the mean differences were statistically significant, as measured by t-tests.

Exhibit 9-10: Pre – Post Job Satisfaction Data, Traditional Delivery (Class 2)

Instrument	JDI Work	JDI Pay	JDI Promotion	JDI Supervision	JDI Co-Workers	JDI Job in General
Pretest						
N	42	43	43	42	43	41
Mean	42.93	24.09	29.77	43.40	43.79	45.10
SD	10.34	16.72	19.46	12.68	10.69	8.54
Posttest						
N	38	38	37	29*	15*	38
Mean	41.87	23.74	32.92	41.10	40.2	44.55
SD	14.40	17.46	17.23	14.37	11.67	13.88

*Lower numbers of cases on these scales were due to a photocopying error that cut off the bottom portion of one of the survey pages.

OCQ data from the agency represented in class 2 are presented in Exhibit 9-11. The overall pretest item mean (i.e., the mean of pretest means) was 5.32 (SD = 1.11) indicating moderate to high organizational commitment. By comparison, the overall posttest item mean (i.e., the mean of posttest means) was 5.01 (SD = 1.30), also indicating moderate to high organizational commitment. Based on data provided by the two survey administrations, there was a very slight, statistically insignificant decrease in overall organizational commitment.

Exhibit 9-11: Pre – Post Organizational Commitment Data, Traditional Delivery (Class 2)

OCQ Item	Pretest Mean N=43	Pretest SD	Posttest Mean N=39	Posttest SD
1. I am willing to put in a great deal of effort beyond that normally expected in order to help this organization be successful.	6.23	1.19	6.15	1.04
2. I talk up this organization to my friends as a great organization to work for.	5.65	1.54	5.28	1.88
3. I feel very little loyalty to this organization.	5.60	2.13	4.85	2.41
4. I would accept almost any type of job assignment in order to keep working for this organization.	4.58	1.89	4.46	2.11
5. I find that my values and the organization's values are very similar.	5.19	1.67	4.92	1.67
6. I am proud to tell others that I am part of this organization.	5.74	1.66	5.62	1.84
7. I could just as well be working for a different organization as long as the type of work is similar.	4.07	1.80	3.92	1.76
8. This organization really inspires the very best in me in the way of job performance.	4.79	1.90	4.97	1.65
9. It would take very little change in my present circumstances to cause me to leave this organization.	4.63	2.13	4.46	2.33
10. I am extremely glad that I chose this organization to work for over others I was considering at the time I joined.	5.77	1.27	5.36	1.71
11. There's not much to be gained by sticking with this organization indefinitely.	4.74	2.03	4.74	2.20
12. Often, I find it difficult to stick with this organization's policies on important matters relating to its employees.	4.76	1.86	4.23	1.86
13. I really care about the fate of this organization.	6.24	1.21	6.00	1.69
14. To me this is the best of all possible organizations for which to work.	5.17	1.74	4.95	2.08
15. Deciding to work for this organization was a definite mistake on my part.	6.52	1.09	5.82	1.82

Class 3

Job satisfaction and organizational commitment data analogous to those just presented are shown for the agency represented in class 3 in Exhibits 9-12 and 9-13. (Recall that job satisfaction and organizational commitment data were not obtained for lass 4 or for the online class due to the late start of these classes in relation to the time frame of the study.) Exhibit 9-12 reveals that at pre-testing, employees expressed satisfaction on all measures except pay (where they were neutral) and promotion (were they were dissatisfied). At post-testing, slight declines were present on the pay, promotion, and co-workers measures, while slight gains were found on the others. None of the mean differences were statistically significant, as measured by t-tests.

Exhibit 9-12: Pre – Post Job Satisfaction Data, Traditional Delivery (Class 3)

Instrument	JDI Work	JDI Pay	JDI Promotion	JDI Supervision	JDI Co-Workers	JDI Job in General
Pretest						
N	29	29	29	29	28	28
Mean	36.97	24.21	21.31	32.14	39.57	35.82
SD	12.36	15.54	16.85	13.84	12.24	15.55
Posttest						
N	36	35	36	36	35	36
Mean	39.14	24.06	19.89	33.50	36.40	38.5
SD	9.90	16.28	14.55	16.00	14.64	14.12

On the OCQ (Exhibit 9-13), the overall pretest item mean (i.e., the mean of pretest means) was 4.35 (SD = 1.40) indicating moderate organizational commitment. The overall posttest item mean (i.e., the mean of posttest means) was just slightly lower at 4.34 (SD = 1.36), also indicating moderate organizational commitment. Based on data provided by the two survey administrations, there was a very slight, statistically insignificant decline in overall organizational commitment.

Exhibit 9-13: Pre – Post Organizational Commitment Data – Traditional Delivery (Class 3)

OCQ Item	Pretest	Pretest	Posttest	Posttest
	Mean N=29	SD	Mean N=36	SD
1. I am willing to put in a great deal of effort beyond that normally expected in order to help this organization be successful.	5.17	1.44	5.53	1.68
2. I talk up this organization to my friends as a great organization to work for.	4.31	2.04	4.36	1.97
3. I feel very little loyalty to this organization.	4.86	1.90	4.50	2.14
4. I would accept almost any type of job assignment in order to keep working for this organization.	4.10	1.95	3.52	1.92
5. I find that my values and the organization's values are very similar.	3.93	2.00	4.03	1.83
6. I am proud to tell others that I am part of this organization.	4.83	1.95	4.44	1.99
7. I could just as well be working for a different organization as long as the type of work is similar.	3.93	1.71	3.67	1.94
8. This organization really inspires the very best in me in the way of job performance.	3.66	2.01	3.75	1.71
9. It would take very little change in my present circumstances to cause me to leave this organization.	4.21	2.01	4.53	1.89
10. I am extremely glad that I chose this organization to work for over others I was considering at the time I joined.	4.59	1.97	4.50	1.88
11. There's not much to be gained by sticking with this organization indefinitely.	3.66	2.06	4.06	1.94
12. Often, I find it difficult to stick with this organization's policies on important matters relating to its employees.	3.45	1.79	3.69	1.67
13. I really care about the fate of this organization.	5.10	1.95	5.06	2.10
14. To me this the best of all possible organizations for which to work.	4.17	2.09	3.97	1.83
15. Deciding to work for this organization was a definite mistake on my part.	5.31	1.95	5.47	1.67

Discussion

The purpose of this study was to address a void in criminal justice evaluation research, namely the paucity of attention that has been devoted to applying evaluation methodology to training programs. To this end, we applied our expanded Kirkpatrick evaluation model to study the NCLETTTC leadership training program. Specifically, we examined the extent to which the program was associated with positive outcomes in the domains of participant satisfaction with training, knowledge gain, and behavioral change. At the level of organizational impacts, we studied the degree to which the program was associated with enhancements in job satisfaction and organizational commitment among employees from select participants' agencies. Finally, we attempted to compare outcomes across traditional and online methods of training delivery, but our capacity for doing this was limited by the time frame of the study in relationship to the later than expected start date and lower than anticipated enrollments for the online program.

Main Findings

A total of 76 trainees completed the four training classes delivered via traditional format during the course of this study, compared with the 100 to 150 who were initially expected to enroll. Furthermore, the online course that was expected to enroll 150 trainees actually enrolled 89, and of these persons, only 26 had actually completed the course by the conclusion of the study's data collection phase.

These data suggest that recruitment was less effective than anticipated for both instructional formats. Although there may be unknown explanations for the lower than anticipated enrollments, the project team observed that it was often difficult for NCLETTTC to obtain firm commitments for training from those agencies expressing interest. Trainings were usually not verified until a few weeks prior to the start date of training. In addition, some trainings were cancelled at very late notice.

Given that the traditional version of the course ran only two days and was completed by most participants who enrolled, our findings also suggest delays, enrollment issues, and problems in course completion among online participants. The online version of the class got off to a far later start than planned. As expected, its rate of enrollment was much faster than the traditional course, but only a small percentage of participants had completed the online course by

the time of this report. The delay in starting the online course was primarily due to the following: (a) the complex nature of the automated evaluation components of the online course (e.g., pretest, posttest, evaluation form, etc.) slowed its design and development and (b) its primary developer (an employee with NCLETTTC who was inexperienced with developing online courses with significant evaluation components) had multiple additional responsibilities. Shortly after the start of the evaluation, an administrator with NCLETTTC was terminated. The person who was scheduled to be the primary developer of the online course was placed in the administrative vacancy. A Boston-based corporation that specializes in multimedia course design and development volunteered to assist with the evaluation by contributing free labor and resources. Despite the best efforts of this corporation, the online enrollments were delayed and much lower than expected by the time of this report.

Demographically, participants who enrolled in the traditional version of the course differed in important ways from those who enrolled in the online version. Greater proportions of online enrollees were women, ethnic minorities, and persons who had not served in the armed forces. Moreover, greater proportions of online than traditional enrollees: (a) said they were taking the course to attain a job promotion, (b) were employed in correctional rather than police agencies, and (c) held the rank of officer vs. manager. At the same time, online enrollees reported having worked in leadership positions for a somewhat longer period (i.e., median of eight years) than traditional enrollees (i.e., median of six years).

One must be careful about making too much of such differences. For example, the finding that the online class enrolled more correctional than police employees primarily reflects the focus of recruitment. At the same time, though, these findings suggest that some groups, particularly those that have traditionally been underrepresented in criminal justice leadership positions (i.e., women and ethnic minorities) may find online training more accessible; the same is likely true of those employees in non-managerial positions who have less flexible schedules. Additionally, the data suggest the possibility that motives for online training may differ, with online trainees being more likely to be motivated by a desire for promotion.

Data such as these on the profile of training participants can also be used to help an agency compare the group(s) actually being reached by training with the group(s) the agency intends to target. Over 83 percent of the people who enrolled in NCLETTTC leadership training

held line or middle management positions. In addition, almost 55 percent held bachelor's degrees. Thus, despite the fact that over 62 percent of participants said they were not completing the course specifically to attain promotion, the data suggest that the training generally targeted a group having prospects for promotion and occupational advancement.

Interestingly, a greater proportion of training participants who completed the online course had some post baccalaureate educational work. Though there are too few cases (n=27) to draw any conclusions, one possible explanation for their ability to complete the course in an expedient manner may be somehow related to their higher level of education.

Level 1

Findings at this level of evaluation revealed that all four traditional format classes displayed very positive reactions to training. By contrast, the trainees completing the course online displayed reactions that, though moderate in magnitude, were significantly more negative than the perceptions of those who completed the course traditionally. Inspection of the mean ratings suggests that the lowest online class ratings were for items addressing instructional aids and exercises, relevance of training to the trainee's organization, alignment of content with trainee background and experience, and alignment between assessment method difficulty and content difficulty. Somewhat higher ratings were achieved on items addressing the clarity and accomplishment of objectives, organization and understandability of content, and instructional methods. Unfortunately, because so few enrollees completed the online training, there is no way to know whether this more negative perception of online training generalizes to all online enrollees. Assuming that it might generalize, the research team suspects one explanation for the lower online ratings was the lack of an interactive, dynamic instructor who could identify one-on-one with the trainees.

Level 2

The overall knowledge pretest mean of 34.27 percent across the four traditional classes suggests that these participants entered training with minimal knowledge of training content. This finding is important since many of the trainees entered training with considerable prior job experience generally, and considerable prior experience at leadership in particular; trainees with such experience potentially held some knowledge of the training topic before enrolling. As in the case of the demographic and background data discussed above, the low knowledge pretest

mean is consistent with the proposition that the training targeted an appropriate audience. In addition, the fact that a small percentage of trainees reported on the demographic form that they had attended any type of leadership training either before (16.4 percent) or during their current line of work (33.9 percent), further supports the proposition that the training targeted the appropriate audience.

All four traditional classes demonstrated significant pre to post gains in knowledge scores, and the overall pre-post gain for the combined classes was also significant. Still, the posttest means were low in absolute terms, with no posttest mean from the traditional classes exceeding 56 percent. This suggests that while some knowledge gain was associated with training, the magnitude of that gain was insufficient. Moreover, across the three traditional classes for which data were available, there was evidence of significant dissipation of knowledge from the first to second posttest, which was administered six months subsequent to training. This suggests that while the training may have produced measurable short-term gains in knowledge of leadership, the training content, method, and/or duration may have been such that gains were not sustained over time. A related possibility is that trainees lost some of the knowledge initially gained because they were not regularly applying it to their jobs after training. Unfortunately, as part of Level 3 measurement, we did not include questions about how frequently trainees had applied the content since training, so this possibility cannot be addressed directly with Level 3 data. Indirectly, however, we know from Level 1 qualitative data that some trainees believed the training could be improved by making it more practical and oriented toward applications on the job.

Level 3

Leadership Performance Survey (LPS) data were obtained from three traditional classes. Though the magnitude of self-rated pre to post performance change was not optimal, participants reported at least some improvement in leadership practices. Mean scores showed that when comparing their pre- and post-training performance, trainees rated their post-training performance about halfway between “quite similar” and “more effective.” This suggests that although not as much progress was evident as might have been desired, some behavior change did transpire, as rated by trainees in those areas targeted by the training curriculum. Across all classes, the least progress was reported on items dealing with showing empathy toward

subordinates as well as practicing nondirective and directive types of counseling. By contrast, the consistently highest progress ratings were found on items pertaining to improving agency/organizational climate and culture, conducting subordinate performance reviews, mentoring subordinates and helping them develop into leaders, and being aware of one's own values and biases and avoiding errors during counseling.

Analogous ratings of trainee leadership performance by their immediate supervisors corroborated the findings from the LPS. Though not as much progress was evident as might have been desired, trainees' supervisors did perceive some trainee behavior change in those areas targeted by the curriculum. This pattern of findings is similar to those made at Level 2, where pretest to posttest 1 change transpired but was not of the desired magnitude (i.e., posttest 1 scores were still relatively low). These findings also support the interpretation that the content, delivery, and/or length of training, though adequate to produce minimal knowledge and behavioral gains, were insufficient to yield the desired or optimal levels of gain.

Level 4

The Level 4 results are also consistent with the interpretation of content, delivery, and/or length insufficiency. Although outcomes were largely favorable at Level 1, and some gains associated with training were demonstrated at both Levels 2 and 3, this was not as true at the level of organizational outcomes.

Across select agencies from each of the three traditional classes studied, staff indicated satisfaction on both pretest and posttest measures of the work, supervisory, and co-worker facets of their jobs; they also exhibited both pre and post satisfaction with their jobs in general. By contrast, the staff members surveyed were less satisfied on both pre and post measures of the pay and promotion facets of their jobs. At the pretest, classes 2 and 3 were, on the whole, neutral toward pay, while class 1 was dissatisfied with pay. Classes 1 and 2 were both neutral toward promotion at the pretest, while class 3 was dissatisfied. By the posttest, only one class (class 2) had moved into the satisfaction range on one measure (promotion).

One interpretation of this pattern of findings is that it was not necessary for the training to emphasize leadership practices specifically intended to improve satisfaction with work, supervision, co-workers, or the job generally because employees in the agencies studied were content on these dimensions of job satisfaction prior to the classes. Although it is clear that

personnel in the agencies studied were far less positive toward pay and promotion prospects before the training classes, it is less clear what trainees could have done as a result of their training experiences to alter actual pay levels and promotion structures, especially within a period of six months following training. Arguably, improved leadership could have improved organizational morale and thereby promoted more positive perceptions of pay and promotion, but the posttest data generally did not support this contention.

Virtually identical interpretations can be applied to the other Level 4 measure, the Organizational Commitment Questionnaire (OCQ). As with four of the six JDI scales, the OCQ indicated reasonably good pre-training organizational commitment in two of the agencies studied. In a third agency, the pretest OCD data suggested moderate organizational commitment. Post-training levels of organizational commitment were very similar in each agency to pre levels. Although there was some room for OCQ scores to improve in each agency, the training did not have the effect of promoting such change.

Strengths, Limitations, and Recommendations

The principal strength of this study is that we were able to demonstrate application of the expanded Kirkpatrick evaluation model to a criminal justice training program. Data were obtained at all four levels of evaluation on measures logically aligned with the objectives and desired outcomes of training. While evidence of desired outcomes was not obtained at Level 4, we demonstrated at least some evidence of positive outcomes in the cumulative fashion implied by the model across Levels 1, 2, and 3. Our research indicates that the expanded Kirkpatrick evaluation model represents a promising tool for structuring and guiding training evaluations in the criminal justice field. In particular, the model can be employed to help determine the degree to which training is sufficient to produce desired outcomes across theoretically cumulative levels. If, as in our study, the data reveal positive, or somewhat positive, outcomes at lower levels (e.g., 1-3) but not at higher ones (e.g., 4), there is empirical foundation for questioning the sufficiency of training along the dimensions of content, delivery, and/or length.

Although there are serious limitations to our online dataset, a second strength of the study is that we were at least able to begin the complex undertaking of comparing traditional with online delivery methods. Our comparison was based on a non-experimental design and restricted to participant demographic and Level 1 data, but we did demonstrate some rather interesting

differences between traditional and online participants on both background and reaction measures. At a minimum, these observed differences warrant comparison of traditional and online modalities in future studies. It should not be assumed that participants completing training traditionally and those completing it online have similar backgrounds or reactions.

Many of the limitations and problems encountered in our research can be expected to be more or less typical of efforts to evaluate training programs of this nature. Consequently, in the remainder of this chapter, we address the major limitations and challenges we encountered and provide recommendations for dealing with these, recommendations that we believe could improve future training and evaluation efforts in criminal justice.

Less Than Anticipated Enrollments

Probably the single most important problem in the NCLETTTC evaluation was the less than expected number of enrollments in both the traditional and online courses. It was assumed that during the course of the study, at least 100 trainees would enroll in the traditional class, which is the minimal number that was agreed upon with the training agency at the outset of the study. The traditional course actually had 76 participants. And while the online course was expected to enroll 150 participants, it actually enrolled under 90 at the time of the study, and only about a third of the online enrollees had completed the course in time to provide evaluation data.

Low online course completion, coupled with delays in starting the online course in a time frame commensurate with the time frame of the study and the fact that Level 2 pretest data were not obtained for the online class, essentially nullified the quasi-experimental design we planned to employ. Since there was neither a sufficient number of online participants nor sufficient time to gather Level 2 (Posttest 2), Level 3, and Level 4 data from online trainees, there was no group against which to compare the pre to post data generated from the traditional classes. To permit valid causal inferences to be drawn about the effects of training, the research design would have necessitated collection of pre and post data from at least two comparison groups exposed to controlled differentials in training (Cook & Campbell 1979; Shadish et al. 2002). This ideal could not be achieved at any level of evaluation in this study.

As already mentioned, NCLETTTC experienced difficulty getting firm commitments from agencies to send their personnel to the leadership training. This is understandable in a field

where it is not uncommon for agencies to experience shifting demands for staffing and periodic staff shortages; administrators may be reluctant to commit staff to training programs too far in advance of training. One recommendation for alleviating this situation is to over-recruit and over-book trainings as a precaution to ensure adequate numbers.

Where low enrollments exist, there will be negative implications for evaluation efforts. Statistical estimates of instrument reliability and validity (e.g., factor analysis results) are rendered unstable by low sample sizes; sometimes it is not even possible to calculate such estimates when sample size is low. Estimates of pre to post group changes and between group differences can be unstable as well. The power of inferential statistical tests to detect differences is reduced by low sample size. These problems are potentially serious from an evaluation standpoint. For instance, our data showed evidence of declines in knowledge from the first posttest to the second one. Clearly, it is important to determine whether any gains made in knowledge (and other constructs) accompanying training are sustained over time.

One recommendation for dealing with this issue is for training program evaluators to perform power analyses at the time studies are being designed and well before trainings are to be scheduled. Such analyses help inform researchers about the sample size that is going to be needed in order to accurately detect pre to post changes or differences between groups, and trainings can be planned accordingly. References to consult regarding power analyses include Cohen (1988) and Kraemer (1987).

Further Challenges Associated with Online Training

One major recommendation for criminal justice agencies that may be thinking about developing and evaluating online training programs is to seek, from the outset, the services of an expert consultant team with knowledge and experience in online course development and evaluation. The complexities of online courses, particularly those that are being evaluated, not only require the expertise of a skilled and experienced course designer and computer programmer, but also the expertise of an experienced program evaluator. Evaluation of training conducted traditionally is complex in its own right. In addition, online evaluations often do not have the benefit of someone present to explain questionnaire items or instructions to a test. The amount of preparation, pilot-testing, and refinement that goes into an online course and subsequent evaluation can easily be underestimated. Furthermore, not only should the content of

online courses be pilot-tested thoroughly, but there also needs to be verification efforts to ensure that online software is recording data as intended. Our experience teaches that failure to anticipate and allocate sufficient time for these things can result in failure to obtain data within the timetable of the research, especially the type of longitudinal data necessitated by Level 3 and 4 evaluations.

Our results also imply that a program which experiences minimal if any difficulty with trainee retention when delivered traditionally over a two-day period may be associated with considerable retention problems when delivered online in a participant-paced manner. When, as in this study, evidence begins to emerge that large numbers of online enrollees are not completing training in a timely fashion, the reason(s) should be ascertained and corrective actions undertaken to the extent possible. In the training program evaluated in this research, there was no mechanism in place to address course non-completion issues. Moreover, the research design did not give adequate attention to systematically gathering data on the reasons for non-completion of online work.

Communication Issues

Other challenges pertained to lack of communication between the different parties involved in the project. As one illustration, it was sometimes difficult to get personnel from NCLETTTC to return telephone calls or emails in an expeditious manner due to their schedules. This is understandable in that two nationally advertised events that NCLETTTC hosts (i.e., Mock Prison Riot and Most Disaster Event) require months of preparation and consume much of staff time as each event draws closer.

Another illustration involved miscommunication between the research team and the corporation that volunteered to assist with course design and development. Despite an in-person meeting at the company's headquarters, conference calls, and numerous telephone and email exchanges, some miscommunications still occurred. Although the content of the online course was reviewed and pilot-tested several times, its record keeping capabilities were not tested. Had the record keeping components been ready for review at pilot testing, it would have become apparent that the pretest scores were not being recorded.

Although there are often valid reasons for delays and misunderstandings of this nature, they can have deleterious effects for evaluations. One recommendation is that the work plans

and timelines agreed to by the various stakeholders in an evaluation must be as detailed and precise as possible in specifying who is responsible for what and when. Although changes and refinements are inevitable, there must be mechanisms in place for clearly communicating these to all parties. Furthermore, it is essential that online course developers work very closely with all relevant parties and pilot test and monitor on a regular basis all of the components of the course and its evaluation.

Course Revision and Research Design

Given that the study involved a single training program that was to be delivered multiple times to different classes, we urged training officials not to alter the course content, duration, or delivery (aside from desired varying of the traditional versus online formats) once the evaluation had begun. This was done to help ensure consistency of what, in the language of research, was the independent variable. On the other hand, the person who served as the instructor of all four traditional classes and the online class pointed out that not being able to make what were identified as needed refinements and revisions to the course turned out to be a major problem. In fact, the instructor's agreement to undertake no efforts to improve the course across deliveries may be one of the reasons why greater pre to post changes were not observed at Levels 2, 3, and 4. That is, the lack of revision could have contributed to insufficiency of course content, delivery, and/or or duration to produce optimal outcomes.

When a training program is to be delivered to multiple classes over time, any variation in the program between classes must be controlled for in the research design if study validity is to be preserved. Therefore, prior to evaluation, course developers should be reasonably confident that during the course of the evaluation, very little or no change in curricular content, delivery, or length will be required. Revisions that are made need to be carefully documented and then considered during the analysis of outcome data.

Data Loss

Another limitation in this study pertained to the six-month follow-up response rate for Level 2 and 3 data. Some non-response is unavoidable in a study of this nature (i.e., employees moving, retiring, etc.). However, we found that some of the trainees continually declined to provide follow-up information, even though incentives (i.e., free online coursework, eligible to enter drawing for \$400 worth of free courses) were used and Dillman's (1978) widely used (and

typically effective) procedure was employed. We believe that the six-month follow-up knowledge posttest was one reason why some trainees refused to participate. The tests may have been perceived as more time/effort consuming and anxiety producing than a simple rating instrument that can be completed more easily.

At the same time, we would not suggest the omission of follow-up knowledge posttests in a training evaluation, especially in view of our finding that initial pre to post 1 knowledge gains showed dissipation effects by the second posttest. Instead, we would recommend the use of better incentives for trainees to complete follow-up knowledge posttests or substituting in-person test administrations for mail survey methods.

In addition, as with any survey, respondents sometimes did not provide information for every item on a given instrument. Similarly, based on some of the Level 3 supervisor surveys received, it is suspected that some supervisors were confused with which part of their six-month follow-up survey to complete (i.e., whether they supervised the trainee before or after the training). Data loss also resulted from the timetable of this study vis-à-vis the pace of training delivery. As mentioned previously, there was not time to collect Level 3 and Level 4 data from the online class, and less data than desired were collected from the fourth traditional class.

Higher Level Evaluation Issues

Due to the nature of the questions involved (i.e., changes in individual and organizational functioning), it will often prove more challenging to obtain valid and comprehensive Levels 3 and 4 data compared to Levels 1 and 2 data. A major limitation of the instrument used in this research to study trainee behavior change at Level 3 is that the instrument did not contain items directly related to trainee application of training content on the job. Both training participants and their supervisors were asked to compare the effectiveness of participant post-training performance with performance prior to training in the areas that were the focus of the curricular objectives. This method is useful but indirect and limited as a means of studying application. The instrument could be improved considerably by adding a series of items measuring the frequency or degree with which knowledge and skills acquired during training were actually used on the job. Ideally, such application data could be collected both before and after training from trainees as well as others (e.g., supervisors, subordinates, independent observers). In

addition, the data collection method need not be limited to surveys, as was the case in this study; interviews and focus groups would also provide potentially valuable information.

An unavoidable difficulty encountered at the Level 4 evaluation stage was not knowing until a few days before the start of training what agency (of the various ones represented at the training) might be suitable for an organizational evaluation. This was primarily due to people having the opportunity to register until the very start of training. A closely related issue involved being sure that a sufficient number of staff from a given agency were attending the training to make organizational change feasible. For example, it is debatable whether having a single mid-level unit manager from an agency of 200 staff complete leadership training is sufficient, in and of itself, to promote measurable organizational change. The trainee to agency staff ratio must be considered in relation to training objectives before decisions about Level 4 evaluation can be made.

Similarly, when conducting Level 4 evaluation, decisions must be made about the appropriate unit of analysis for study. The organizational measures we used (i.e., the JDI and OCQ) did not involve all the same persons completing both pre and post instruments due to various staffing changes. A staff member present for pre-testing might not have been available for post-testing, or vice versa. We did not see this as a serious threat to the validity of our research because the organization itself was being treated as the unit of analysis; and based on discussions with NCLETTTC officials, the objectives of leadership training clearly encompassed creating agencies with more satisfied staff and improved levels of organizational commitment. But as a matter of principle, the objectives of a given training program will dictate whether a pre to post measurement strategy of this nature is valid.

On a related note, our pretest data showed that, in the agencies studied, job satisfaction and organizational commitment were generally not low prior to training. Of course, this raises the question of whether improvements in job satisfaction and organizational commitment were viable or reasonable training objectives in the first place. When Level 4 evaluations (and indeed training curricula themselves) are being planned, it will often be useful to conduct a needs assessment during the pre-training period to determine whether any presumed Level 4 goal of training should in fact be a goal (i.e., whether much improvement can be expected given both the pre-training level of the construct and the objectives of training). Just as researchers have to be

sure that their Level 4 measures are valid indicators of the constructs the training is designed to alter, those responsible for curriculum design and delivery must ensure that the constructs being targeted by training in fact need targeting.

Another consideration is that Level 4 evaluation will often be far less convenient and far more consuming of time and other resources than measurement at other levels. In the present study, for instance, sometimes the agency that agreed to participate in a Level 4 evaluation was not located close to the training site. On one occasion, this required a member of the project team to drive approximately 80 miles to and from the training site in order to collect Level 4 data. Level 4 evaluations should be planned only when time and resources are sufficient.

In addition, it was sometimes difficult to collect Level 4 data from all of an agency's personnel. Most personnel were able to be contacted by having a member of the project team at each roll call over the two-day period when all shifts were present. Administrative staff were contacted during the course of the day. However, personnel on vacation, sick leave, reserve duty, etc., were more difficult to contact. The project team attempted to collect data from these personnel by leaving survey forms with an agency administrator, who would make arrangements to get the surveys completed and mailed back to the project team member. However, this method of data collection produced mixed results.

Trainee and Trainer Detachment from Evaluation

In this study, the research team made decisions about what measures to use at all levels of evaluation, with some input from training officials and none from trainees. While it may not be possible or desirable for trainees to have input on Level 1 or 2 measures (other than during pilot tests), the same does not necessarily hold true for Level 3 and 4 data. If it is the trainee's behavior and his or her organization that training is designed to alter, trainees will often be uniquely positioned to provide insights about what change is realistic and measurable. One approach we have used in other evaluations is to have trainees explicate certain key performance indicators (KPIs) at Levels 3 and 4 that they think might be affected by training and to help them decide how these KPIs can be measured. Under this approach, the role of an external evaluator becomes to collaborate with agency staff (and, when possible, internal evaluators or research staff) to collect the relevant data and accurately interpret the results of analyses. Such an approach has the important advantage of giving the trainee and his or her agency a more direct

investment and sense of ownership in the evaluation project. The evaluation becomes something that is collaborative, rather than something imposed. Consequently, there is a greater probability that evaluation results will be put to use to promote improvements at the agency.

Similarly, it is easy to see how collaboration can be established with training officials and how these officials can develop an investment in the evaluation. Training officials can provide valuable insight and assistance with (a) the development of measures at all levels, particularly at Level 2; (b) data collection; and (c) interpretation of findings. Astute training officials will have reviewed the literature on the constructs being targeted by training, and this places them in a strong position to consult with evaluators. Here again, the ultimate benefit is the potential for evaluation results to be put to use to improve training curricula and delivery.

A worthwhile tool in such a collaborative evaluation effort (particularly when those involved in the collaboration are geographically dispersed) is a project website consisting of links to all the main project materials and products. This would include links to project design documents (such as work plans and timelines), relevant bodies of literature, measurement instruments, data files, output files, reports and drafts thereof, and dialogue boards for questions and commentary.

APPENDIX 9-A

NCLETTTC Study Timeline, Milestone & Workplan Chart

Objectives	Date	NCLETTTC	EKU	ILJ
<ul style="list-style-type: none"> • Tasks ⇒ Products (Starting date November 2003) (Anticipated Completion date December 2004)				
1. Make arrangements to obtain adequate number of participants to participate in evaluation. <ul style="list-style-type: none"> • Prepare draft explanation of evaluation to incorporate in registration materials • Review explanation of evaluation • Finalize explanation • Incorporate explanation in online and traditional course registration procedures • Advertise discounted course at openhouse and in magazines, corrections.com & at conferences ⇒ Product – Pool of adequate number of participants	Nov 21 Nov 28 Dec 1 Dec 7 Dec 15 plus	Yes Yes Yes Yes Yes	Yes (lead) Yes Yes	Yes

Objectives • Tasks ⇒ Products (Starting date November 2003) (Anticipated Completion date December 2004)	Date	NCLETTTC	EKU	ILJ
2. Conduct Level 1 Evaluation (measures reaction of participants to training). <ul style="list-style-type: none"> • Review existing training evaluation instruments that record demographic data and measure reactions to training • Draft new instruments • Review new instruments • Finalize new instruments • Incorporate new instruments in both online and traditional courses • Observe traditional course & interview trainer and students to gain feedback on course implementation • Observe online course (i.e., review online class) • Collect demographic and reaction data from all classes • Code data • Analyze data ⇒ Product – Results of Level 1 evaluation	Nov 15 Nov 21 Nov 28 Dec 1 Dec 7 Jan (TBA) Jan (TBA) Start Jan - Ongoing Ongoing Ongoing	Yes Yes Yes Yes Yes Yes Yes Yes	Yes Yes (lead) Yes Yes Yes Yes Yes Yes	Yes Yes Yes Yes Yes Yes Yes

Objectives • Tasks ⇒ Products (Starting date November 2003) (Completion date December 2004)	Date	NCLETTTC	EKU	ILJ
<p>4. Conduct Level 3 Evaluation (measures behavior change, such as what information gained from training is being used on job).</p> <ul style="list-style-type: none"> • Draft, review and finalize letter to agency • Draft supervisor instrument and trainee graduate performance/activity logs • Review instruments • Finalize instruments • Incorporate new instruments in both online and traditional courses • Send instruments to supervisors and trainee graduates • Collect data from supervisors and trainee graduates • Code data • Analyze data <p>⇒ Product – Results of Level 3 evaluation</p>	<p>Nov 21</p> <p>Nov 21</p> <p>Nov 28</p> <p>Dec 7</p> <p>Dec 15</p> <p>6 mos. Post Grad - Ongoing</p> <p>Ongoing</p> <p>Ongoing</p> <p>Ongoing</p>	<p>Yes</p> <p>Yes</p> <p>Yes</p> <p></p> <p>Yes</p> <p>Yes</p> <p></p> <p>Yes</p>	<p>Yes</p> <p>Yes</p> <p>Yes</p> <p>Yes</p> <p></p> <p>Yes</p> <p></p> <p>Yes</p> <p>Yes</p>	<p></p> <p></p> <p>Yes</p> <p></p> <p></p> <p></p> <p></p> <p></p>

Objectives • Tasks ⇒ Products (Starting date November 2003) (Completion date December 2004)	Months	NCLETTTC	EKU	ILJ
<p>5. Conduct Level 4 Evaluation (measures impact of training on organization, cost/benefit, return on investment, morale etc.)</p> <p>Tasks - to be determined</p> <p>⇒ Product – Results of Level 4 evaluation</p>				

APPENDIX 9-B

NCLETTTC Training Evaluation Materials

Participant Demographic Information

Training Reaction Survey

Pre and Post Knowledge Tests

Performance Survey

Supervisor's Appraisal Form

**ADVANCED LEADERSHIP TECHNIQUES FOR FIRST RESPONDERS,
CORRECTIONS, AND SECURITY OFFICER TRAINING COURSE**

Participant Demographic Information Form

Date of Training: Month: ____ Days: ____ Year: ____

Personal Information:

1. Name: _____

2. Age: _____

3. Please indicate your gender (circle one): Male Female

4. Please indicate your race/ethnicity (circle one):

White/Caucasian African-American American Indian

Hispanic Asian Other (please specify)_____

5. Please indicate your marital status (circle one): Single Widowed Married Divorced

6. Please indicate your current home mailing address:

Street: _____ City: _____ State: _____

Zip code: _____

7. Educational Information:

Level Of Education:	Major Area of Study:	Number Years Completed:	Received Degree/Certificate (circle one)	
High School/GED			Yes	No
College Education:			Yes	No
Associate Degree:			Yes	No
Bachelor Degree:			Yes	No
Masters Degree:			Yes	No
Doctorate Degree:			Yes	No
Professional Degree			Yes	No
Specialized Schooling:			Yes	No
Vocational:			Yes	No
Military:			Yes	No

8. Military Experience:

Have you ever served in the Armed Forces, either in active duty or reserve status (circle one)?
 Yes No

If yes, which branch did you (or do you) serve in (circle one)?

Army Air Force Marines Navy Other (please specify): _____

What was the highest rank you achieved while in the armed forces? _____

Briefly explain the responsibilities associated with your highest rank: _____

What was (or is) the total number of years you have served in the armed forces? _____

9. Current Employment:

Current Employer (name of agency): _____

Current Position: _____

How long have you worked in this current position? _____

Immediate Supervisor's Name: _____

Immediate Supervisor's Mailing Address:

Street: _____ City: _____ State: _____

Zip code: _____ Email: _____

10. Leadership Training History

Are you seeking completion of this leadership course as a means to attain a promotion (circle one)?

Yes No

Have you completed another leadership course(s) while employed in your current line of work (circle one)?

Yes No

If yes, when did you complete the leadership course(s)? Month: ____ Year: ____

Please briefly describe the course(s):

Have you completed other leadership courses while employed in any other line of work (circle one)?

Yes No

If yes, when did you complete the leadership course(s)? Month: ____ Year: ____

Please briefly describe the course(s): _____

What is the combined total number of years you have you been employed in a leadership position? _____

ADVANCED LEADERSHIP TECHNIQUES FOR FIRST RESPONDERS, CORRECTIONS, AND SECURITY OFFICER TRAINING COURSE EVALUATION FORM

Name of Officer: _____ Date of Training: _____

Please check the box that best describes your level of satisfaction with each statement listed below.

Ratings = Strongly Agree (5) Agree (4) Neutral (3) Disagree (2) Strongly Disagree (1)

5 4 3 2 1

1. Objectives were clearly set at the beginning of training.
2. Objectives were met by the end of training.
3. The training was well organized.
4. Instructional methods were used effectively.
5. Instructional aids (e.g., graphics, etc.) were beneficial.
6. Content was presented at a level appropriate to my background and experience.
7. Content was presented in a clear and understandable manner.
8. Topics covered were relevant to actual leadership skills called for in my organization.
9. Exercises meant to practice or demonstrate knowledge and/or skills were used effectively.
10. Content of assessment methods (e.g., exercises, case study, test) was aligned with content of training topics.
11. Difficulty level of methods of assessment (e.g., exercises, case study, test) was aligned with difficulty level of the training topics.
12. Positive aspects of course outweigh its negative aspects.
13. I would take another similar type of course if it were offered (and needed).
14. This training will help me perform my job more effectively.

COMMENTS

What was most beneficial about the training?

What was least beneficial about the training?

What changes (if any) would you recommend?

ADVANCED LEADERSHIP TECHNIQUES FOR FIRST RESPONDERS, CORRECTIONS, AND SECURITY OFFICER TRAINING----- PRE-TEST

Answer the following multiple choice questions by marking the letter below that corresponds with the BEST answer on your answer sheet. If you do not know the answer, please select the choice that you most agree with.

1. The best leadership style a Field Training Officer should adopt with new recruits is:

- a. directing
- b. participating
- c. delegating
- d. transformational

2. A person wearing a t-shirt with a New York City Fire Department logo is an example of _____, because it denotes that person wanting to be part of something bigger than them.

- a. culture
- b. climate
- c. admiration
- d. identification

3. One of the most important characteristics of an effective leader is:

- a. flexibility
- b. strict adherence to rules
- c. compassion
- d. fairness

4. While counseling Sergeant Wright about his deficient performance the past several months, Captain Johnson discovers that Sergeant Wright has been spending a lot of his time dealing with his invalid parents and their financial matters. Captain Johnson should:

- a. give Sergeant Wright information on how to get power of attorney
- b. refer the matter to his immediate supervisor for guidance
- c. refer Sergeant Wright to a nearby social work office
- d. refer Sergeant Wright to a nearby legal assistance office

5. Training director Donovan wants to inject stress into her department's in-service training so her staff can be better prepared to handle stress in emergency situations. The best method in which to add stress to her department's training environment is to:

- a. require high standards of performance
- b. conduct realistic training with limited resources
- c. conduct realistic training with detailed constraints
- d. add unanticipated conditions

6. One caveat of using the Transactional Leadership Style is that:

- a. in the wrong situations, it can make the leader seem self-serving
- b. it limits the leader's ability to influence individuals
- c. it reduces individual commitment to the task
- d. it allows for no input from the subordinates

7. Which type of counseling communicates standards and is an opportunity for leaders to establish and clarify expected values, attributes, skills, and actions?.
- performance counseling
 - professional growth counseling
 - developmental counseling
 - evaluation counseling
8. The transformational leadership style can best be characterized by:
- giving subordinates the authority to make decisions and holding them accountable
 - motivating subordinates to use their skills and knowledge and to exercise initiative
 - motivating subordinates to work by offering rewards and punishment
 - asking subordinates for input and recommendations prior to making a final decision
9. Good leaders should avoid counseling atmospheres that:
- promote one way communication
 - are formal in nature
 - are informal in nature
 - reinforce the leader's authority
10. During a hostage situation involving inmates at a medium security institution, Warden Springhill decides to give commands from outside the institution rather than inside the institution where he/she can be closer to the situation. This action of the warden has the potential of influencing the most the:
- discipline of the staff
 - morale of the staff
 - stress of the staff
 - commitment of staff
11. Public Safety Commissioner Thornton is concerned about how her staff feel about their organization. That is, she wants to know more about their shared perceptions and attitudes of the organization, as well as what they believe about the day-to-day functioning of their organization. Commissioner Thornton is concerned about the organization's:
- culture
 - ethos
 - morale
 - climate
12. The anticipated results of a leader's decisions and actions are referred to as:
- planned actions
 - anticipated findings
 - intended products
 - intended consequences
13. Leaders who want to emphasize that some acts are not tolerated and some acts are rewarded would use which style of leadership:
- transformational leadership style
 - participating leadership style
 - transactional leadership style
 - directing leadership style

14. Active listening involves:

- a. listening and letting the subordinate do all of the talking
- b. Speaking only to reinforce importance of what is being said
- c. Not asking questions so that the subordinate can feel relaxed
- d. Communicating nonverbally and listening for common themes

15. In order for subordinates to be successful in a developmental counseling atmosphere, they must be:

- a. willing to share the role of leader
- b. honest in their own personal assessment of their progress and goals
- c. prepared to critique co-workers
- d. able to communicate effectively with their leader

16. To use this style, you must have the courage to communicate your intent and then step back and let your subordinates work. You must also be aware that immediate benefits are often delayed until the mission is accomplished.

- a. transactional
- b. participating
- c. transformational
- d. direct

17. The purpose of developmental counseling is to:

- a. cause subordinates to do things to improve their performance
- b. develop subordinates to achieve personal, professional, and organizational goals
- c. correct substandard performance and reward good performance
- d. help the subordinate understand the expectations of their position

18. Leaders with empathy:

- a. are less likely to be biased toward subordinates
- b. are more aware of differences between individuals of different backgrounds
- c. are more likely to act consistently with their own values
- d. have more credibility and influence with subordinates

19. Which of the following is the most difficult to measure”?

- a. morale
- b. discipline
- c. commitment
- d. teamwork

20. While counseling Corporal McGuire about his substandard performance, you, his supervisor, learn that he has been having a difficult time dealing with his wife, who suffers from severe depression. Which of the following two types of counseling, and in what order, should be utilized when dealing with this situation?

- a. performance counseling; followed by crisis counseling
- b. performance counseling; followed by referral counseling
- c. referral counseling; followed by performance counseling
- d. crisis counseling; followed by performance counseling

21. An effective counseling style includes the following characteristics:
- a. fairness, honesty, support
 - b. sympathy, fairness, respect
 - c. flexibility, respect, communication
 - d. identification, fairness, motivation
22. Upon arrival at an accident site involving multiple injuries, Captain Chapman immediately begins to give orders to those emergency personnel present on what to do. What type of leadership style is he/she exhibiting?
- a. transactional
 - b. delegating
 - c. direct
 - d. transformational
23. In addition to changing the speed of response to crime, accidents and other emergency situations, technology is:
- a. decreasing mental stress
 - b. decreasing physical stress
 - c. decreasing both mental and physical stress
 - d. increasing both mental and physical stress
24. The directive counseling approach:
- a. takes too much time in many situations
 - b. allows the subordinate to develop the solution
 - c. teaches the subordinate how to attain standards
 - d. advises the subordinate that a given course of action is best
25. Which type of counseling is preferred for most counseling sessions?
- a. directive
 - b. transactional
 - c. nondirective
 - d. combined

ADVANCED LEADERSHIP TECHNIQUES FOR FIRST RESPONDERS, CORRECTIONS, AND SECURITY OFFICER TRAINING -----POST-TEST 1

Answer the following multiple choice questions by marking the letter below that corresponds with the BEST answer on your answer sheet. If you do not know the answer, please select the choice that you most agree with.

1. _____ is a long lasting, complex set of shared expectations consisting of shared attitudes, values, goals, and practices characterizing the larger institution.

- a. Climate
- b. Morale
- c. Culture
- d. Discipline

2. An effective leader:

- a. will have all the necessary skills to help his or her subordinates
- b. will refer family-related matters pertaining to subordinates to a superior
- c. should refer a troubled subordinate to a person or agency more qualified
- d. will recognize his/her limitations in helping subordinates

3. Your new chief informs the department that one of her main priorities will be to improve the organizational climate. The chief will be primarily concerned with changing:

- a. long held beliefs, customs and practices of the officers
- b. shared attitudes, values, goals and practices of the officers
- c. what officers believe about the department's day to day functioning
- d. the environment of the department in which the officers work

4. Major Jones is considered by many to be a very good leader because he/she has a proven track record of counseling subordinates and developing new leaders out of them. He/she allows subordinates to take responsibility, does not project his/her biases onto subordinates, helps each subordinate develop an action plan that fits them, and is honest and consistent with subordinates. Which of the following best describes Major Jones' qualities as an effective developmental counselor for subordinates?

- a. firmness, fairness, credibility, discipline, cultural awareness
- b. respect, self awareness and cultural awareness, empathy, credibility
- c. respect, cultural awareness, motivation, discipline
- d. self awareness, empathy, credibility, motivation

5. Sergeant Johnson's squad typically performs well at equipment inspections. However, at the most recent inspection, his squad failed to meet several equipment standards. His supervisor should conduct which of the following types of developmental counseling:

- a. event oriented counseling
- b. performance counseling
- c. professional growth counseling
- d. crisis counseling

6. Supervisor's must mentor their subordinates through:

- a. discipline, guidance, and empathy
- b. respect, support, and counseling
- c. teaching, discipline, and counseling
- d. teaching, coaching, and counseling

7. Counseling errors include:

- a. keeping all information from the subordinate confidential
- b. not confirming subordinate prejudices or beliefs
- c. drawing conclusions based on more than what the subordinate tells you
- d. encouraging the subordinate to take the initiative

8. Subordinates are most likely to take care of each other, to be more committed and innovative, and to be more likely to take calculated risks when their leaders combine techniques from both:

- a. directing and participating leadership styles
- b. transactional and transformational leadership styles
- c. participating and delegating leadership styles
- d. directing and delegating styles

9. As a result of constant and intensive training, combined with rewards for good performance ratings and punishments for bad performance ratings, Sergeant Dennington's unit has more confidence in its ability to respond to an emergency situation. Which of the following best describes the reason for the unit's new confidence?

- a. discipline
- b. morale
- c. ethos
- d. spirit

10. When carefully writing tasks and outlining the unit's orders, Lt. Barnes goes over the potential benefits of success as well as consequences of failure. This leadership style typically:

- a. teaches subordinates to exercise initiative and to take calculated risks when facing a crisis
- b. is a good team building approach that teaches subordinates how to create a plan
- c. evokes only short-term commitment from subordinates and discourages risk taking and innovation
- d. is considered to be demeaning and abusive to subordinates

11. What new change is forcing first responders to change the way they operate?

- a. downsizing of first responder groups due to financial constraints
- b. connection of local criminals with more organized criminal organizations
- c. the fast rate of technological change
- d. retirement of trained officers

12. First responders, who are the first to see the casualties of strangers and possibly their friends, have the greater burden of:

- a. being afraid of the unknown
- b. waiting on additional help to arrive
- c. realizing their own mortality
- d. letting their leaders and co-workers down

13. Effective leaders:

- a. cross train by training lower ranking personnel the advance skills of senior personnel
- b. adjust leadership styles and techniques according to the type of people they lead
- c. cross train by training senior personnel the basic skills of lower ranking personnel
- d. are fair and try to treat everyone consistently by adopting one leadership style

14. In order to lessen the effect of staff vacancies, Superintendent Marshall changed the institution's shifts from straight to rotating. Shortly thereafter, several staff members complained about how attending college classes became much more difficult because of their rotating work schedule. This action resulted in:
- a. unintended consequences
 - b. intended consequences
 - c. changing the organization's climate
 - d. changing the organization's culture

ADVANCED LEADERSHIP TECHNIQUES FOR FIRST RESPONDERS, CORRECTIONS, AND SECURITY OFFICER TRAINING-----POST-TEST 2

Answer the following multiple choice questions by marking the letter below that corresponds with the BEST answer on your answer sheet. If you do not know the answer, please select the choice that you most agree with.

1. Chief Thomas is a police chief who values training and teamwork. She often participates in the in-service training that may of her officers attend. She also makes it a point to show up at each different shift, including weekends, and talk to all of the department's staff, including officers, dispatchers, and office personnel. These actions of the police chief have the potential of influencing the most the:
 - a. discipline of the staff
 - b. morale of the staff
 - c. stress of the staff
 - d. commitment of the staff

2. Which of the following purposes would best be served by training that is rigorous, realistic, and that prepares officers to be obedient, to expect the unexpected and to take the correct action even when the leader is not around?
 - a. improvement of morale
 - b. improvement of discipline
 - c. improvement of commitment
 - d. improvement of teamwork

3. First responders, such as those in law enforcement, need to be prepared for missions that might be stressful. To prepare first responders to act under stressful conditions, leaders need to make sure the training environment appears real by adding:
 - a. constraints
 - b. high standards
 - c. unanticipated conditions
 - d. limited resources

4. The volume of information that current technology makes available renders which of the below more important than in the past?
 - a. centralizing decision-making
 - b. operating with other organizations
 - c. determining mission-critical information
 - d. retirement of trained officers

5. A new leader is helping to positively influence an organization's climate when he/she:
 - a. attempts to change the way things were done in the past by pointing out deficiencies
 - b. lives up to the same expectations he/she sets for his/her followers
 - c. offers rewards for getting the job done, no matter how it was accomplished
 - d. accepts the responsibility for faults on behalf of everyone under his/her command

6. Police officers throughout a state who wear black tape over their badges at the funeral of a slain officer represent an example of _____, because it is a practice that consists of the shared values, attitudes, goals and practices that characterize the larger institution.
- climate
 - morale
 - ethos
 - morale
7. The best leadership style a police chief can adopt in a hostage situation at a bank robbery is:
- directing
 - participating
 - delegating
 - informational
8. When asking for her unit's input and recommendations in creating an evacuation plan for school staff and students in the event of a bomb threat, Captain Saylor:
- is exhibiting signs of weakness and lack of self-confidence
 - is obligating herself in having to follow her subordinate's advice
 - is creating a sense of subordinate ownership of the plan
 - is challenging her subordinates to rise above their immediate needs
9. By practicing developmental leadership, wherein both the professional and personal growth of subordinates are emphasized, a leader:
- is teaching mature and experienced subordinates to be accountable for their actions
 - is evoking only short term commitment from his subordinates and is discouraging risk taking
 - must be aware that the benefits are not immediate
 - still insists on outlining all of the conditions of task completion
10. This leadership style is used by leaders who are not involved in fast paced operations and have the time to consult with subordinates:
- transactional
 - transformational
 - participating
 - direct
11. During a natural disaster the state governor requests that police cadets who have not yet graduated from the state police academy perform some limited law enforcement duties in the areas affected. The cadets will patrol affected areas in an effort to prevent widespread looting. What type of leadership style should their supervisor, who is the National Guard commander, exhibit?
- direct leadership style
 - transformational leadership style
 - participating leadership style
 - delegating leadership style
12. What leadership style motivates subordinates to perform through a system of rewards and punishments?
- direct leadership style
 - transformational leadership style
 - participating leadership style
 - transactional

13. Leaders combine what two styles by doing the following with their subordinates: outline all of the conditions of task completion, describe benefits of success and the consequences of failure, and inspire and show individualized concern?

- a. directing and participating leadership styles
- b. transactional and transformational leadership styles
- c. participating and delegating leadership styles
- d. directing and delegating styles

14. In order to increase the firepower of police officers so that fewer officers would be killed or injured in shootouts, most of the nation's police departments switched from revolvers to semi-automatic pistols. However, the percentage of officers involved in accidental shootings where their firearm accidentally discharged has increased. This action resulted in:

- a. unintended consequences
- b. intended consequences
- c. changing the organization's climate
- d. changing the organization's culture

15. In a _____ review, a leader works with a subordinate to reach agreement on a plan of action that will build on the subordinate's strengths as well as improve areas of weakness:

- a. leadership development
- b. counseling
- c. after action
- d. performance

16. The technique a leader uses to counsel effectively must:

- a. be based on a very close observation of the subordinate's behavior
- b. be the same technique that is used for all of the leader's subordinates
- c. include passive listening, responding, questioning, and a plan of action
- d. fit the situation, the leader's capabilities, and the subordinate's expectations

17. Lieutenant Harlan is respected by the entire department because the values that are important to him, such as being respectful, honest, and consistent in thought and deed, are reflected in the quality of his work. He is appointed by the chief to be district commander for the area of the city that is largely inhabited by Hispanics because he has made it a point to become familiar with the values, perspectives, and concerns of this community. Which of the following best describes the qualities of Lieutenant Harlan that make him an effective developmental counselor for subordinates?

- a. fairness, discipline, cultural awareness
- b. respect, self awareness and cultural awareness, credibility
- c. respect, cultural awareness, motivation
- d. self awareness, empathy, motivation

18. Counseling errors include:

- a. keeping all information from the subordinate confidential
- b. not confirming subordinate prejudices or beliefs
- c. drawing conclusions based on more than what the subordinate tells you
- d. encouraging the subordinate to take the initiative

19. An effective leader should:
- be trained and/or educated in having all of the necessary skills to help his or her subordinates
 - refer all of a subordinate's family-related matters, including marital problems, to a superior
 - always refer a troubled subordinate to a person or agency more qualified to handle the problem
 - recognize his/her limitations and when needed, refer a subordinate to another person or agency
20. One of the subordinate officers under your command, who normally has very high performance reviews, has recently been witnessed by other officers to be very disrespectful of detainees in your facility's lockup. You should conduct which of the following types of developmental counseling:
- event oriented counseling
 - performance counseling
 - professional growth counseling
 - crisis counseling
21. A leader should not conduct any form of developmental counseling when a subordinate officer:
- has been using a lot of sick leave to take care of a family problem
 - has been recognized by the mayor and city council for heroism
 - has just graduated from the academy and has been assigned to your unit
 - has taken drug contraband from the evidence room without approval
22. The directive counseling approach:
- encourages subordinates to talk freely and openly
 - is the least time consuming counseling approach
 - allows the leader to decide which solution is best
 - stresses the subordinate's decision-making responsibilities
23. The major differences between the nondirective, directive and combined counseling approach pertains to:
- how much training and experience in counseling the leader has
 - the amount of time it takes to complete the session
 - the degree to which the subordinate participates in the session
 - the degree the counselor participates in the session
24. To reduce your subordinate's perception that you may be unnecessarily biased or judgmental:
- you should get other staff members to give examples of your subordinate's performance
 - you should require your subordinate to give examples of his/her own performance
 - you should provide examples or specific observations of your subordinate's performance
 - both you and your subordinate should provide examples of your subordinate's performance
25. End the counseling process by:
- summarizing its key points and scheduling future meetings
 - implementing a plan of action and evaluating results
 - asking the subordinate if he/she understands the plan of action
 - following through on referrals and scheduling future meetings

ADVANCED LEADERSHIP TECHNIQUES FOR FIRST RESPONDERS, CORRECTIONS, AND SECURITY OFFICER TRAINING

LEADERSHIP PERFORMANCE SURVEY

Name of officer (provided by consultant): _____

Please rate your job performance in the following areas (1 – 36) since completion of the Advanced Leadership Techniques for First Responders, Corrections, and Security Officer Training Course at the National Corrections & Law Enforcement Training & Technology Center. Notice that the questions are meant to compare your performance since this training with what your performance was like before you completed the training. Please circle one rating per question on the rating scale provided.

Compared to my pre-training performance, my post-training performance has been (circle one):

	More Effective		Quite Similar		Less Effective	
1. Addressing the human dimension of leadership (discipline, morale, taking care of people)	5	4	3	2	1	0
2. Preparing for and dealing with mission stress (hostile/violent situations)	5	4	3	2	1	0
3. Preparing for and dealing with the stress of change (new laws, technology, etc.)	5	4	3	2	1	0
4. Building positive organizational climate (short term environment)	5	4	3	2	1	0
5. Building positive agency culture (long term environment)	5	4	3	2	1	0
6. Mixing leadership styles successfully (participating, delegating, transformational, and transactional leadership)	5	4	3	2	1	0
7. Anticipating both the intended and unintended consequences of my decisions and actions	5	4	3	2	1	0

Because the topic of counseling was a central part of the training, the following questions pertain to counseling. Please respond according to the same instructions as above.

8. Practicing developmental counseling with subordinates	5	4	3	2	1	0
9. Developing leadership capabilities among subordinates	5	4	3	2	1	0

Compared to my pre-training performance, my post-training performance has been (circle one):

	More Effective		Quite Similar		Less Effective	
10. Conducting performance reviews with subordinates (strengths, weaknesses, and plans of action)	5	4	3	2	1	0
11. Mentoring subordinates	5	4	3	2	1	0
12. Clearly defining the purpose of counseling sessions	5	4	3	2	1	0
13. Fitting counseling style to the subordinate/situation in question	5	4	3	2	1	0
14. Demonstrating respect for subordinates	5	4	3	2	1	0
15. Establishing/maintaining communication with subordinates	5	4	3	2	1	0
16. Supporting subordinates	5	4	3	2	1	0
17. Being aware of my own values and biases	5	4	3	2	1	0
18. Displaying empathy in dealings with subordinates	5	4	3	2	1	0
19. Establishing my credibility by being honest and consistent	5	4	3	2	1	0
20. Engaging in active listening	5	4	3	2	1	0
21. Paying attention to the gestures of subordinates	5	4	3	2	1	0
22. Responding verbally and non-verbally to subordinates during counseling	5	4	3	2	1	0
23. Asking subordinates the proper number and type of questions	5	4	3	2	1	0
24. Avoiding counseling errors (rash judgments, inflexibility, etc.)	5	4	3	2	1	0
25. Being aware of my limitations as a leader	5	4	3	2	1	0
26. Making referrals as needed	5	4	3	2	1	0

27. Counseling in relation to a specific event or situation (substandard performance, new employee, crisis, etc.)	5	4	3	2	1	0
28. Counseling to improve subordinate performance	5	4	3	2	1	0

Compared to my pre-training performance, my post-training performance has been (circle one):

	More Effective		Quite Similar		Less Effective	
29. Counseling to promote professional growth of subordinates	5	4	3	2	1	0
30. Practicing nondirective counseling	5	4	3	2	1	0
31. Practicing directive counseling	5	4	3	2	1	0
32. Being able to combine nondirective and directive counseling	5	4	3	2	1	0
33. Identifying the need for counseling	5	4	3	2	1	0
34. Preparing for counseling (scheduling, notifying, organizing)	5	4	3	2	1	0
35. Conducting counseling sessions (opening, developing a plan, closing)	5	4	3	2	1	0
36. Following-up after counseling sessions	5	4	3	2	1	0

37. How many times have you referenced or used the developmental counseling form since completion of training? _____ times

38. Have there been any barriers that have interfered with the transfer of skills learned in the NCLETTTC leadership training to your workplace (circle one)? Yes No

If yes, please describe the nature of those barriers:

39. Have there been any other factors besides the leadership training at NCLETTTC that may have affected your leadership skills since completion of your training (circle one)? Yes No

If yes, please describe the nature of those factors:

Thank you very much for your cooperation. Please mail the completed questionnaire in the pre-paid, self-addressed envelope that has been provided.

**ADVANCED LEADERSHIP TECHNIQUES FOR FIRST RESPONDERS,
CORRECTIONS, AND SECURITY OFFICER TRAINING**

SUPERVISOR'S APPRASIAL OF EMPLOYEE LEADERSHIP PERFORMANCE

Name of Supervisor (provided by consultant): _____

_____, an employee with your agency who our records indicate you are responsible for directly supervising, completed the Advanced Leadership Techniques for First Responders, Corrections, and Security Officer Training on _____ (provide date) at the National Corrections & Law Enforcement Training & Technology Center. This survey is concerned with this employee's job performance since completion of the training. Notice that the questions are meant to compare the employee's performance since training with what performance was like before training. Under Part A below, please rate job performance in the areas indicated (1 – 21), making one rating per question on the rating scale provided.

If you did not supervise this employee prior to completion of the training and at least six months following the training, please skip Part A below and complete Part B. If you complete Part A, please leave Part B blank. Please make sure to complete Parts C and D before returning the survey.

PART A (complete only if you supervised this employee before training and at least six months thereafter)

How many months did you supervise the above-mentioned employee prior to completion of training? __ months

How many months did you supervise the above-mentioned employee after the training? __ months

Compared to this employee's pre-training performance in the below-mentioned areas, post-training performance has been:

	More Effective		Quite Similar		Less Effective
1. Addressing the human dimension of leadership (discipline, morale, taking care of people)	5	4	3	2	1 0

2. Preparing for and dealing with mission stress (hostile/violent situations)	5	4	3	2	1	0
3. Preparing for and dealing with the stress of change (new laws, technology, etc.)	5	4	3	2	1	0
4. Building positive organizational climate (short term environment)	5	4	3	2	1	0
5. Building positive agency culture (long term environment)	5	4	3	2	1	0

Compared to this employee's pre-training performance in the below-mentioned areas, post-training performance has been:

	More Effective		Quite Similar		Less Effective	
6. Mixing leadership styles successfully (participating, delegating, transformational, and transactional leadership)	5	4	3	2	1	0
7. Anticipating both the intended and unintended consequences of decisions and actions	5	4	3	2	1	0
8. Practicing developmental counseling with subordinates	5	4	3	2	1	0

Because the topic of counseling was a central part of the training, the following questions pertain to counseling. Please respond according to the same instructions as above.

9. Developing leadership capabilities among subordinates	5	4	3	2	1	0
10. Conducting performance reviews with subordinates (strengths, weaknesses, and plans of action)	5	4	3	2	1	0
11. Mentoring subordinates	5	4	3	2	1	0
12. Demonstrating respect for subordinates	5	4	3	2	1	0
13. Establishing/maintaining communication with subordinates	5	4	3	2	1	0
14. Supporting subordinates	5	4	3	2	1	0

15. Establishing credibility by being honest and consistent	5	4	3	2	1	0
16. Making referrals as needed	5	4	3	2	1	0
17. Counseling in relation to a specific event or situation (substandard performance, new employee, crisis, etc.)	5	4	3	2	1	0
18. Counseling to improve subordinate performance	5	4	3	2	1	0
19. Counseling to promote professional growth of subordinates	5	4	3	2	1	0
20. Identifying the need for counseling	5	4	3	2	1	0
21. Following-up after counseling sessions	5	4	3	2	1	0

PART B (complete if you did not supervise this employee before training and at least six months thereafter)

How long have you supervised the above-mentioned employee since completion of the training?
 _____ months

During the time I have supervised this employee, his/her performance in the below-mentioned areas has been:

	More Effective		Quite Similar		Less Effective	
1. Addressing the human dimension of leadership (discipline, morale, taking care of people)	5	4	3	2	1	0
2. Preparing for and dealing with mission stress (hostile/violent situations)	5	4	3	2	1	0
3. Preparing for and dealing with the stress of change (new laws, technology, etc.)	5	4	3	2	1	0
4. Building positive organizational climate (short term environment)	5	4	3	2	1	0
5. Building positive agency culture (long term environment)	5	4	3	2	1	0
6. Mixing leadership styles successfully (participating, delegating, transformational, and transactional leadership)	5	4	3	2	1	0
7. Anticipating both the intended and unintended consequences of decisions and actions	5	4	3	2	1	0
8. Practicing developmental counseling with subordinates	5	4	3	2	1	0

Because the topic of counseling was a central part of the training, the following questions pertain to counseling. Please respond according to the same instructions as above.

9. Developing leadership capabilities among Subordinates	5	4	3	2	1	0	
10. Conducting performance reviews with subordinates (strengths, weaknesses, and plans of action)	5	4		3	2	1	0
11. Mentoring subordinates	5	4	3	2	1	0	
12. Demonstrating respect for subordinates	5	4	3	2	1	0	

During the time I have supervised this employee, his/her performance in the below-mentioned areas has been:

	More Effective		Quite Similar		Less Effective	
13. Establishing/maintaining communication with subordinates	5	4	3	2	1	0
14. Supporting subordinates	5	4	3	2	1	0
15. Establishing credibility by being honest and consistent	5	4	3	2	1	0
16. Making referrals as needed	5	4	3	2	1	0
17. Counseling in relation to a specific event or situation (substandard performance, new employee, crisis, etc.)	5	4	3	2	1	0
18. Counseling to improve subordinate performance	5	4	3	2	1	0
19. Counseling to promote professional growth of subordinates	5	4	3	2	1	0
20. Identifying the need for counseling	5	4	3	2	1	0
21. Following-up after counseling sessions	5	4	3	2	1	0

PART C (Please complete this section)

22. Are you aware of any barriers that have interfered with the transfer of skills learned in the NCLETTTC leadership training to this employee's workplace (circle one)? Yes No

If yes, please describe the nature of those barriers:

23. Are you aware of any other factors besides the leadership training at NCLETTTC that may have affected the above-mentioned employee's leadership skills since completion of training (circle one)? Yes No

If yes, please describe the nature of those factors:

PART D (please complete this section)

What is your current job title? _____

Date you completed this survey _____

Thank you very much for your cooperation. Please mail the completed questionnaire in the pre-paid, self-addressed envelope that has been provided.

Chapter 10

National Judicial College's Civil Mediation Training

National Judicial College

The National Judicial College (NJC), located at the University of Nevada, Reno, has served as the country's leading state judicial educational institution for more than 40 years. The NJC mission is to provide "leadership in achieving justice through quality judicial education and collegial dialogue." The main objective of the NJC is to improve judicial proficiency, competency, skills, and productivity through education and training programs. The NJC educates judges from all areas of the state judicial system, including general jurisdiction, appellate, administrative law, tribal, and military. The national perspective of the NJC allows judges to learn firsthand from national experts and fellow participants about new and innovative techniques, as well as alternative methods of performing the work of the courts.

The NJC was formed in the early 1960s through the efforts of the American Bar Association's Joint Committee for the Effective Administration of Justice to develop a national training organization as a means for improving the administration of justice. Over time, the NJC changed its location, title, and mission. Originally known as the National College for State Trial Judges, the NJC moved from Boulder, Colorado, to Reno, Nevada, with funding from the Max C. Fleischman Foundation. The NJC mission expanded in the 1980s to include a master's degree program; in the 1990s with the addition of the National Tribal Judicial Center; and in 2001 with the addition of a Ph.D. program in judicial studies. The master's degree program, developed jointly with the National Council of Juvenile and Family Court Judges, was one of the nation's first advanced degree programs for trial court judges. In addition to the advanced degree programs, the NJC provides ongoing training through courses, seminars, and web-based courses.

Since its creation over 40 years ago, the NJC has awarded more than 70,000 professional education certificates. The NJC offers courses on site, online, across the nation, and internationally. Approximately 95 courses are offered annually to more than 2,700 judges from all 50 states, U.S. territories, and more than 150 countries. In an effort to make its services available to the widest possible audience, the NJC provides scholarship assistance to some

participants. In addition, it continually develops new courses in an effort to keep learning current and also holds faculty development workshops.

Literature Review

In the United States, people rely upon the courts to resolve differences. As caseloads have increased, wait times have gotten longer, and associated costs have increased, alternatives to the justice system have been sought by court personnel, judges, lawyers, litigants, and society at large (Hensler 2003). Interest in alternative dispute resolution (ADR) is high, as reflected in its ever growing domains, including:

- Family realm
 - divorce
 - custody
 - eldercare
 - family businesses
 - estate disputes
- Workplace
 - wrongful termination
 - discrimination
 - harassment
 - labor management
- Public realm
 - environmental
 - land use
- Youth
 - school conflicts
 - peer mediation
- Criminal justice
 - victim-offender restorative justice
 - domestic violence
- Contractual
 - landlord-tenant disputes
 - homeowners' associations
 - builders/contractors/realtors/homeowners
 - medical malpractice
 - legal malpractice

Alternative dispute resolution processes include negotiation, mediation, arbitration, and hybrid methods such as mediation-arbitration. Mediation is a process in which disputants are assisted by a neutral third party to come to a mutually agreeable, and often self-determined, solution. Mediation and arbitration both rely on a neutral third party. In arbitration, however, it is the role of the third party to render a decision, whereas in mediation, the third party helps the disputants to reach a voluntary settlement (Alaska Judicial Council 1999). Mediation can be court-connected, meaning that a court may refer the parties to mediation, or the parties can decide on their own to seek mediation. Mediation is most effective when it is used in cases in which the parties have an ongoing relationship, such as parents, business partners, employees, students, or neighbors (National Center for State Courts 2005).

The key qualities of the mediation process are that it is voluntary, collaborative, controlled, confidential, informed, neutral, and empowering (Melamed 2004; Yeend 2003). Mediation is voluntary—parties can leave at any time for any reason. While the mediation process can be mandated, the parties cannot be forced to agree and settle. Mediation is collaborative and controlled, in that parties work together to reach a solution, and they have complete decision making power in reaching the solution. Mediation is informed, meaning that the process offers the parties an opportunity to obtain legal and expert opinion if desired. Mediation is neutral in that the mediator is impartial and balanced; and it is empowering because the participants make the decisions.

While there are no compulsory, formal steps to the mediation process, there are five general steps that usually occur. These are (Yeend 2003):

- Initiation—educating the parties about mediation, gathering information, and disclosing relevant facts, such as potential conflicts of interest and costs.
- Opening remarks—a welcome, discussion of the process, encouragement to participate, and explanation of procedures.
- Problem solving—clarification of issues and values, identification of options, and discussion of various solutions.
- Resolution—clarifying options, evaluating terms, and formalizing an agreement or deciding next steps if there is an impasse.
- Closure—debriefing by the mediator on the skills and techniques they used and what could be done in similar situations.

Research on mediation results has been mixed (Benjamin & Irving 1995; National Center for State Courts 1994; RAND 1996). However, one consistent finding is that people are more satisfied with the mediation process than with traditional court processes (National Center for State Courts 2005). People who have mediated rather than adjudicated their cases have considered mediation to be “fair and responsive” to their needs and have had lower rates of re-litigation compared to those who adjudicated their cases (Alaska Judicial Council 1999).

Civil Mediation Training

The purpose of the NJC Civil Mediation course is to familiarize student participants with the civil mediation process and qualify them for certification in states that require it. After course completion, participants should be able to conduct a mediation session, outline common standards of conduct and ethical considerations, handle special problems and avoid classic errors, and summarize the interpersonal dynamics of mediation.

Program Overview

NJC’s Civil Mediation is a week-long, 40-hour course that the instructors describe as “...designed to be interactive, hands-on, fast-paced and practical” (Yeend 2003, p. ii). The instructors, Nancy Neal Yeend and John Paul Jones, designed the course for the NJC and developed the curriculum materials. At the conclusion of the course, it is expected that participants would be proficient in conducting civil mediations. Skill development, however, must continue because mediation is an art that requires constant practice and refinement.

The Civil Mediation course begins on Sunday afternoon with lecture, discussion, and exercises on the fundamentals of mediation. Instruction during the week starts at 8:00 a.m. each day and ends between 5:00 and 6:00 p.m. The course concludes with a graduation on Friday at 4:00 p.m. The course curriculum covers the following modules:

- Fundamentals
 - Expectations
 - Defining Mediation
 - Models
 - Dispute Resolution Process
 - Process Selection Considerations
 - Development of ADR

- Role of the Mediator
- Mediator Requirements
- Process Challenges
- Court-connected ADR
- Process
 - Case Initiation
 - Causes of Conflict
- Opening Remarks
 - Mediator’s Opening Remarks
 - Critical Components
 - Participants’ Opening Remarks
- Strategies
 - Co-mediation
 - First Impressions
 - Confidentiality
 - Behavioral Guidelines
 - Separate Meetings
- Problem Solving
 - Problem Articulation & Resolution
 - Issues Identification
 - Trust
 - Negotiation
 - Options
 - Reframe
- Communication
 - Communication Styles
 - Listening & Listening Filters
 - Metaphors
 - Gathering Information
- Ethics and Bias
 - Ethical Standards & Challenges
 - Bias Development
 - Biases in Mediation
- Role of Advocates
 - Representation
 - Multiple Parties
 - Unrepresented Participants

- Experts
- Insurance Issues
- Agreements
 - Fundamentals
 - Closure
 - Traps for Mediators
 - Successful Mediators
- Creating a Court Program
 - Fundamental Elements
 - Sample Program

One of the key methods in this course is role playing. The instructors are committed to providing participants every opportunity to engage in a variety of mediation sessions by creating mediation simulations that let participants practice their newly emerging knowledge and skills. Throughout the week, there are 10 role plays in which all participants take part as mediator, plaintiff, defendant, and attorneys for plaintiff and defendant. The role plays focus on different strategies and techniques the participants need to learn to be effective mediators. By the end of the course, the participants will have practiced an entire mediation simulation, from beginning to end.

By the time the class participants have completed the 40 hours of training in civil mediation, they will have learned how to conduct a mediation from set-up and opening remarks through closure, they will understand the ethical standards and how to deal with bias, they will have learned how to set up a mediation program or business, and they will have had the opportunity to act as mediators in simulated mediations.

Evaluation Methodology

The purpose of this evaluation was to assess what participants learned at the Civil Mediation course. In particular, this research sought to determine how the participants felt about the training and to assess what they learned.

Evaluation Questions

As with the other evaluations, the Civil Mediation evaluation followed our expanded version of Kirkpatrick's model. As described earlier in this report, the first level, reaction,

measures what the participants felt about the training, while the second level, learning, measures what knowledge the participants gained from the experience. This evaluation examined the following evaluation questions (see Exhibit 10-1):

- How did the training participants react to the training? What was their assessment of the Civil Mediation course in terms of meeting their goals and expectations? Would they recommend the course to others?
- What did the training participants learn in terms of information and skills?
- What new skills can the participants perform on the job after the training?

Exhibit 10-1: Evaluation Questions and Data Collection Tools Grouped by Evaluation Level

Evaluation Level	Goal	Evaluation Questions	Data Collection Tool
1: Reaction	Determine immediate reaction to training	How did the participants react to the training? Were they satisfied?	Student course evaluation
2: Learning	Transfer new knowledge and skills to participants	What information and skills were learned?	Participant information sheet Pre-post knowledge test Follow-up survey
3. Behavior Change	Assess if the trainees can perform the new skills	What new skills learned at training can be performed back on the job?	Follow-up survey

Data Collection Methods and Framework

This section outlines the evaluation plan and data collection framework. The discussion focuses on the participants, evaluation design, evaluation framework, and study strengths and weaknesses.

Evaluation Plan

Participants

The NJC Civil Mediation course participants were mainly civil and criminal court judges, although additional participants could include tribal and military judges, mediators, and court personnel. A total of 69 participants attended the two Civil Mediation classes offered in 2004 at the National Judicial College campus in Reno. The first class was offered from February 29-March 5 (n=25) and the second from October 31-November 5 (n=44). The evaluation team served as participant observers in the first class, including participating in class exercises and role plays in an effort to develop a better understanding of the class, materials, participants, and trainers. As a result of this participation, the evaluation team was able to develop all of the data collection instruments except the standard student course evaluation already in use by the NJC. A consequence of participation in this class, however, was that most of the data collection did not take place until the second class.

The majority of the Civil Mediation course participants were male (79.1 percent), employed as judges (93.0 percent), and worked in general jurisdiction or district courts (50.0 percent). Twenty-five percent of the participating judges worked in administrative, probate, special, and related types of courts, while 22.5 percent worked in circuit or trial courts. Only one participant worked at the appellate level. Three participants were not judges; one worked as a hearings examiner, one as an attorney-mediator, and one as a court-appointed mediator. The years of experience as a judge ranged from a low of one year to a high of 32 years. The mean was 14.88 years, and the median and mode were both 15 years.

Design

The design for the Civil Mediation course evaluation was a pre-post within-subjects design with a follow-up conducted over a 10-month period (see Exhibit 10-2). Before the start of training, participants completed a pre-knowledge test and a participant information sheet. The pre-knowledge test assessed participant knowledge of civil mediation issues that would be addressed in the class. The participant information sheet asked about demographics (e.g., gender, occupation, years of experience), prior civil mediation experience and training, and reason for taking the course. At the immediate conclusion of training, participants completed a post-knowledge test and a student course evaluation. The post-knowledge test, identical to the pre-test, measured learning, while the student course evaluation assessed participant feedback

about the course and instructors. Ten months after the conclusion of the training, evaluators sent a follow-up survey, which asked participants about progress toward accomplishing their goals, and included questions about expectations and recommended improvements.

Exhibit 10-2: Pre-Post Test Within-Subjects Design of the Civil Mediation Training

Pre-Test	Training	Post-Test	2 nd Post-Test
O _a O _b		O _c O _d	O _e
T ₁	T ₂	T ₃	T ₄

Where:

- O_a = Pre-knowledge test
- O_b = Participant information sheet
- O_c = Student course evaluation
- O_d = Post-knowledge test
- O_e = Follow-up surveys
- T = Time periods

Data Collection Framework

Originally, the evaluation team had planned for a Level 4 evaluation assessing results of the training.⁴⁰ By the time the participants completed Civil Mediation, they would, in theory, have the required knowledge and skills to work as mediators. Measuring Level 4 results would have entailed assessing these impacts on the court organizations. This would have involved assessing the impact that the trained judge had when he or she returned to their courts—measuring the extent to which the number of successful civil mediations increased.

Many of the Civil Mediation participants, however, took the course not only to improve their own knowledge, but also to provide employable skills later—after they retired from the bench. Thus, it was not timely to conduct a Level 4 evaluation; and this also made a Level 3 evaluation challenging. Unfortunately, the evaluation team was not aware of the judges' motives in taking the course until we had already committed to the evaluation. The evaluation team

⁴⁰ In selecting each of the four trainings to evaluate as a test of the training evaluation model, the evaluation team established clear criteria for selection, including trainings that teach a measurable skill. In the case of the Civil Mediation course, participants were learning how to be mediators.

collected some Level 3 data, however, not in the degree of detail originally intended to assess changes in participants' behavior resulting from the course.

Level 1: Reaction

The key evaluation question for a Level 1 evaluation of the Civil Mediation course was: How did the participants react to the training? Answering this involved student course evaluation surveys focusing on the participants' reaction to the training. The survey instrument was developed by NJC staff and is standard in all NJC trainings (see Appendix 10-A).

Level 2: Learning

The key evaluation question for a Level 2 evaluation of the Civil Mediation course was: What information and skills were learned by the participants? Answering this involved conducting pre- and post-test surveys focusing on the knowledge obtained during the training. This survey was administered immediately before and after the training. The knowledge test was developed by the evaluation team after participating in the course and was reviewed for validity by the course instructors (see Appendix 10-A).

Level 3: Behavior

As mentioned above, we did not expect major behavioral changes from Civil Mediation participants because many took the course to improve their knowledge and obtain skills to use at a future date. We did, however, obtain some Level 3 data from a follow-up survey conducted with participants nine months following the course. The evaluation team developed the follow-up survey (see Appendix 10-A).

Study Strengths and Weaknesses

As with any evaluation, there are strengths and weaknesses of the design. Strengths of this evaluation of the Civil Mediation training include unprecedented access to the NJC staff and facilities, the Civil Mediation course instructors, and the training participants. The NJC staff and Civil Mediation instructors were open and responsive to this evaluation, allowing the evaluation team the freedom to observe and collect whatever data was possible. Additional strengths include the high response rate of the training participants. Seventy percent of the training participants returned the surveys. The participants did not require more than two or three follow-up telephone calls or emails in order to secure this response rate.

The main weakness of this evaluation was the lack of comprehensive Level 3 and adequate Level 4 measures. While we were able to provide strong measures of Level 1 (reaction) and Level 2 (learning), the lack of more comprehensive evaluation measures limits our ability to assess how participants used the training once back at work and how it may have changed their behavior and benefited their organizations. Additional weaknesses include the lack of a comparison group and the relatively low number of participants in the evaluation sample. While the pre-post design provided an adequate alternative to a comparison group, it assumes that changes in knowledge of mediation can be attributed to the course. The relatively low number of sample participants is a result of the NJC offering the Civil Mediation course only once or twice a year. While not ideal, the number of participants was sufficient for this design.

Evaluation Findings

Central to the evaluation design of the Civil Mediation course was determining whether the course met participants' expectations, assessing what they learned, and examining whether they have accomplished or will be able to accomplish their goals as a result of taking the course. By the end of the 40-hour training, participants would hold a certificate of completion and were expected to be able to conduct mediations from start to finish.

The Civil Mediation course evaluation findings were based on the following sources of data:

- Pre and post knowledge test
- Participant information sheet
- Student course evaluation
- Follow-up survey

Participant Reaction

In assessing participant reaction, the central questions were: What was the participants' assessment of the training? What did they think of the training and the course content? Findings were derived from the student course evaluations.

Student course evaluations were collected using the NJC's standard procedures for collecting student feedback on trainings. The course evaluation survey for Civil Mediation, developed by NJC, was administered at the end of the training by NJC staff, who collected the

surveys from the training participants. The survey questions dealt with a range of issues, including accomplishment of course objectives, course subject matter, and faculty and teaching assistants, as well as NJC services such as housing accommodations and food service. Data were provided to the evaluation team.

Overall, evaluations of the Civil Mediation course were high. On a seven-point scale, with seven being high and one being low, no ratings fell below 6.1. When asked to rate the overall value of the course, the mean score was 6.7. When asked to rate how successfully the course objectives had been accomplished, the participants were equally enthusiastic (see Exhibit 10-3).

Exhibit 10-3: Participant Reaction to the NJC Civil Mediation Training: Accomplishment of Course Objectives

Course Objective	Mean Score Scale 1 (low) to 7 (high)
Utilize the mediation process and describe useful techniques of mediation	6.6
Conduct mediation from opening remarks to agreement	6.5
Outline common standards of conduct and ethical considerations for mediators	6.6
Manage special problems that arise in mediation, and avoid classic errors	6.3
Avoid biases, including gender, and remain neutral when mediating	6.5
Evaluate your performance and the performance of other participants during the role playing exercises	6.4
Master the interpersonal dynamics of mediation	6.1

Participant comments on the course were very positive. Responses included the following comments:

- “One of the most informative continuing legal education courses I have ever taken.”
- “Great, interesting, educational, fun course.”
- “The training was excellent.”
- “I really enjoyed the week.”

- “I can’t think of a better combination of textual material, lectures, and role-playing activities to become familiar with the dispute resolution process.”
- “I learned valuable skills that I can use on the bench.”
- “Simply a good course.”
- “This was one of the best courses I have ever attended. Well organized, intensive and well presented.”

The few negative comments were limited to particular aspects of the course. These comments included:

- “Too much repetition.”
- “Formulas too rigid.”
- “[I] did not agree with some theory.”
- “They spend too much time telling us not to state opinions, etc., which is foolish thinking in the real world.”

Knowledge and Skills Gained

In assessing learning, the central question was: What information and skills were learned? Findings were derived from the knowledge test participants took at the beginning and at the conclusion of the training. The pre-post design provided the evaluation team with a baseline measure of participants’ knowledge from which improvements could be determined.

As noted earlier, as part of the pre-testing before training began, participants were given an information sheet that asked for both background demographic data and inquired about the reason for taking the course and previous experience with mediation training. The majority of participants did not have any previous training in mediation, arbitration, or related topics (60.5 percent reported no training and 39.5 percent reported having had some previous training), although slightly more than half did have previous experience with civil mediation (56.8 percent); although these prior mediation experiences generally involved normal judicial duties in small claims court, pre-trial hearings, or settlement conferences.

Each participant was given a knowledge pre-test at the beginning of the training and the same test at the immediate conclusion of the training five days later. NJC staff administered both the pre- and post-tests. The questions were based on the course modules and included questions such as:

1. The five aspects of a court-connected mediation program are:
 - a. budget, code of conduct, administration, education, evaluation
 - b. policies, budget, neutrals, administration, evaluation
 - c. budget, case screening, code of conduct, standards, review
 - d. policies, budget, case screening, education, review
2. Why are the mediator's opening remarks critical to the mediation process?
 - a. ensures that all sides "meet and greet"
 - b. sets the tone and enhances the prospect of resolution
 - c. provides the only avenue for establishing the ground rules
 - d. it is NOT a "critical" part of the mediation process
3. The two primary models of mediation are:
 - a. restorative and representative
 - b. facilitative and arbitratative
 - c. curative and representative
 - d. facilitative and evaluative

Of the 44 participants who took the knowledge test, 44 completed pre-tests, and 42 completed post-tests. The tests were matched and analyses were conducted on the matched pre-post tests.

The results of the knowledge test showed that Civil Mediation participants improved in their knowledge of mediation techniques as a result of the course. There were 15 total questions on the knowledge test. On the pre-test, two participants scored a low of four, and three participants scored a high of 12. On the post-test, only one participant scored a low of eight, and two participants scored a high of 15. Statistical analyses indicate a significant improvement in the test scores of participants between the post-test ($M=11.79$, $SD=1.42$) and the pre-test ($M=8.41$, $SD=1.93$), $t(38)=10.136$, $p<.001$. Significantly, participants improved their scores by nearly 23 percent from the pre- to the post-test.

Participants were asked the reason they were taking the mediation course—prior to the start of the course and again in the follow-up survey. The open-ended answers given on the participant information sheet were used to provide multiple choice answers in the follow-up survey. The follow-up survey was sent to the 44 participants, with a return rate of 70 percent. Of those responding to the follow-up survey, 96.8 percent stated they took the course to improve their skills, 87.1 percent stated they took the course to better understand mediation, 61.3 percent took the course to provide part-time work in retirement, and 54.8 percent desired a certificate. Less than half the respondents took the course because it was recommended by a colleague or

supervisor (48.4 percent), because they wanted additional training (19.4 percent), or because they were looking for a job change (9.7 percent).

On the follow-up survey, participant feedback about the knowledge and skills learned was equally positive. Interestingly, comments reflected both specifics of mediation and stylistic issues, such as learning how a mediator's job is fundamentally different than a judge's. For instance, in the words of one participant, "I learned to assist and not dictate in decisionmaking." Many participants made statements to this effect. When asked what was the most important skill or most significant thing learned, participants reported the following:

- "Listening!"
- "To be a good listener."
- "Continued patience while fostering an environment for constructive communication among parties."
- "To be more patient with parties (and myself)."
- "Reframing issues in a way that is neutral."
- "Leveling the table, body language, and rephrasing."
- "Listening to the litigants."
- "How to properly mediate."
- "How to involve the parties."

When asked what was the most significant thing learned from the training, participants stated:

- "That you are the conduit for the parties reaching agreement."
- "How you approach other people's problems is the key to a pragmatic solution."
- "The practice sessions were most helpful, although pretty stressful!"
- "To stop taking so much personal responsibility for the outcome."
- "The difference in being a mediator from being a judge."

The vast majority (90.3 percent) of participants reported that they were both satisfied with what they had learned and that the course met their expectations. In addition, 90.0 percent of respondents stated they would recommend the Civil Mediation training to colleagues.

Behavior Change

Although not a valid Level 3 evaluation because not enough of the sample had even intended to use the new skills until a later date and because the instrument was self-report only, we did obtain some behavior change self-report information. When asked in the follow up surveys if they were doing anything differently as a result of the training, 71.0 percent of the respondents indicated they were. Comments included:

- “My ability to settle cases without giving my opinion has increased dramatically.”
- “Being a judge and a mediator require different skills. Possibly, I’m more patient and willing to listen.”
- “I’m now assuming a more passive role when conducting settlement conferences.”
- “I retired from the bench on 12/31/04 and now do arbitrations and mediations.”
- “Applied skills to mediation and settlement conference work.”

It is not unusual, especially in long trainings such as the Civil Mediation training, for participants to feel that some parts of the training are unnecessary. It is a sign of training quality when these comments are few, as was the case for this training. Only four survey respondents indicated they felt that some aspects of the training, exercises, or activities were a “waste of time.” In particular, some participants felt that too much time was spent on theory, and other participants felt that exercises were rushed. In general, most respondents would agree with the statement by one participant that “during the course there were times I felt there was a waste, but in the end it all came together.”

Discussion

The Civil Mediation course provided participants with the skills and tools necessary to work as mediators in civil cases. By the end of the training, participants worked through mock mediations; they learned all the techniques for conducting mediations from beginning to end; and for those seeking to start their own mediation business, they were provided a starting point. This evaluation sought to answer the questions of what the training participants felt about the training, what they learned, and if they used the learning on the job. We believe this evaluation provides meaningful feedback to NIJ, BJA, NJC, trainers, course participants, and the greater mediation community.

This evaluation collected five sources of data, including pre and post knowledge tests, information on participant demographics and course expectations, student course evaluations, and follow-up surveys approximately 10 months after training. Findings indicated that the participants learned more about civil mediation as a result of taking the training. Follow-up surveys indicated that most felt they had accomplished their goals and had gained both knowledge and skills. The three main goals attained as a result of the training, according to the participants, were to improve skills and knowledge, plan for a job change or retirement, and receive a certificate.

As noted earlier, keys to successful training include designing a high quality curriculum and finding capable instructors. The Civil Mediation course involved a transdisciplinary, problem-based learning style (Glenn, et al. 2003). The training curriculum was based on teaching the basic principles required to conduct a mediation. These skills were transmitted to participants via lecture, interactive discussion, and a series of exercises and role-plays that successively built upon each other. Thus, by the end of the 40-hour training, participants were able to conduct mediations of civil cases from initiation through conclusion.

In the Civil Mediation training, the instructors were active mediators with many years experience conducting successful mediations. In addition, the instructors created the course at the NJC and wrote the text used for the training—the course was based on their training delivered to the private sector. The instructors' dedication and thoroughness were key reasons why this training was successful. Many participants stated that the course was well organized, well run, and very informative.

As part of the survey process, participants were asked for their recommendations to improve the training. Most respondents stated “nothing,” or a variation of “if it isn’t broken, don’t fix it.” The suggestions that were offered to improve the training can be categorized into two major themes: (1) centered on the participants themselves and (2) practical aspects of the course. A number of participants recommended training judges and the non-judges separately. Particular comments included:

- “I felt somewhat constrained during the discussion periods because of the presence of non-judge students.”
- “Limit attendance to judges. Mediators, non-judges, whose work is reviewed by judges can have a chip on their shoulders.”

Participants also had a few minor complaints and recommendations to improve the training. These included:

- “Rotate seats so that partners during exercises are different.”
- “Learning should not be an endurance test.”
- “The days are too long.”
- “Smaller groups.”
- “[Add] a component regarding labor mediation.”
- “Shorten the course and [add] a follow-up about a year later.”
- “Add a Civil Mediation II section.”
- “Tape and critique actual cases being mediated.”

The evaluation team adds the following recommendations:

- The Civil Mediation training should constantly be improved through evaluation. A cycle of instruction, evaluation, and adjustment to curriculum, instructors, or methods is vital to keeping training fresh, current, and successful. Our expanded Kirkpatrick model provides a good framework for training providers to use to evaluate their courses.
- NJC might also strive to conduct a Level 4 training evaluation to show a return on investment. Additional key Level 4 questions include measures of cost benefit or cost effectiveness of the training. For example, can participants “buy” this same training elsewhere for less? These are important questions that require use of comparison groups and carefully planned pre- and post-training cost benefit measures.

Please fill-in your assigned identification number _____

APPENDIX 10-A

NJC Training Evaluation Materials

Course Evaluation

Participant Information

Knowledge Test

Follow-up Survey

**NATIONAL JUDICIAL COLLEGE
COURSE EVALUATION**

**CIVIL MEDIATION
FEBRUARY 29-MARCH 5, 2004**

EVALUATE THE COURSE OBJECTIVES LISTED BELOW.

Please rate the course objectives as to how successfully they were accomplished. Separate evaluations of subject matter, written material, faculty, and discussion groups are to be made later in this evaluation form. **Scale 7=High 1=Low**

-
- 1) Utilize the mediation process and describe useful techniques of mediation.

 - 2) Conduct mediation from opening remarks to agreement.

 - 3) Outline common standards of conduct and ethical considerations for mediators.

 - 4) Manage special problems that arise in mediation, and avoid classic errors.

 - 5) Avoid biases, including gender, and remain neutral when mediating.

 - 6) Evaluate your performance and the performance of other participants during the role playing exercises.

 - 7) Master the interpersonal dynamics of mediation.
-

Comments:

Rate the overall value of the entire course

To the extent that issues of diversity (i.e., gender, race culture, sexual orientation, religion, or disability) arose in this course, were they adequately addressed?

Comments:

Please indicate any subjects/skills which should be:

1) Added:

2) Lengthened:

3) Shortened:

4) Deleted:

Rate the overall value of the entire course

Please comment on the professionalism and courtesy of The National Judicial College staff with whom you had contact during the course

Are there other ways NJC can help you to be a better judge?

Would you allow us to use your comments *with your name* in our NJC literature?

Yes

No

National Judicial College: Civil Mediation Course
October 31 – November 5, 2004
Participant Information Sheet

1. What is your gender? Male Female

2. Are you a judge? Yes No
 If yes, what type of court? _____
 How many years have you been on the bench? _____

3. If you are not a judge, what is your current occupation? _____
 How many years have you been at this job? _____

4. Have you had any previous training on mediation, arbitration, or related subjects?
 Yes No

If yes, please list:

TITLE	Number of Hours

5. Do you have any previous experience with civil mediation? If yes, please describe.

6. What is your reason for taking this course? Why are you here?

National Judicial College
Civil Mediation Course
October 31–November 5, 2004
Knowledge Assessment

Please fill-in the assigned identification number given to you at the beginning of the course _____

The Institute for Law and Justice is conducting an independent evaluation of the NJC’s Civil Mediation class. We are asking that you complete this survey to help with this evaluation. Please use your unique identification number assigned to you at the beginning of the class. By using this number, your answers will remain anonymous, but can be linked to other evaluation assessments you may be completing. For each question, there is just one “best” answer. Please circle the best answer for each question. Thank you for your participation.

1. The greatest degree of finality in the resolution of a case can be achieved through:
 - a. arbitration
 - b. litigation
 - c. summary jury trial
 - d. mediation

2. Classic mediation embodies the following concepts, EXCEPT:
 - a. empowerment of the parties
 - b. principled negotiation
 - c. non-neutral decision making
 - d. enhanced communication

3. The five aspects of a court-connected mediation program are:
 - a. budget, code of conduct, administration, education, evaluation
 - b. policies, budget, neutrals, administration, evaluation
 - c. budget, case screening, code of conduct, standards, review
 - d. policies, budget, case screening, education, review

4. The most widely used form of ADR is:
 - a. negotiation
 - b. conciliation
 - c. interviewing and counseling
 - d. mediation

5. Why are the mediator's opening remarks critical to the mediation process?
 - a. ensures that all sides "meet and greet"
 - b. sets the tone and enhances the prospect of resolution
 - c. provides the only avenue for establishing the ground rules
 - d. it is NOT a "critical" part of the mediation process

6. Confidentiality is one of the hallmarks of mediation. Circumstances in which a mediator can breach confidentiality include:
 - a. failure to reach resolution
 - b. death of one of the parties
 - c. admission of child abuse
 - d. violation of the agreement

7. The two primary models of mediation are:
 - a. restorative and representative
 - b. facilitative and arbitratve
 - c. curative and representative
 - d. facilitative and evaluative

8. Private caucuses with participants are NOT used by a mediator to:
 - a. gain insight into the dispute
 - b. manage emotions
 - c. determine who's "right"
 - d. reality test ideas

9. A mediation session in which one party is represented by counsel and the other side is not, is:
 - a. considered unethical by the Uniform Mediation Act
 - b. a common occurrence in insurance disputes
 - c. challenging for the mediator and may require the mediator to determine if mediation is the appropriate process
 - d. usually quickly settled
10. The primary causes of conflict between parties includes control, personal values, and:
 - a. actions
 - b. emotions
 - c. change
 - d. culture
11. “Traps” that new mediators fall into include the following, EXCEPT:
 - a. Forgetting to “reality-test” solutions
 - b. Forgetting the obvious: a sincere apology
 - c. Forgetting a formalized written memorandum of settlement terms
 - d. Forgetting to make an opening statement
12. A mediator is allowed to dictate the terms of a settlement:
 - a. never
 - b. only after the single-text method has failed
 - c. when the parties have reached agreement on all but one issue
 - d. as needed to end the mediation
13. An important step in the mediation process involves identifying the issues that are in dispute. Issue identification involves the following, EXCEPT:
 - a. agenda creation
 - b. encouraging discussion
 - c. reframing the problems
 - d. prioritizing plaintiff’s issues first

14. Words that should be avoided by the mediator during a session include:
 - a. “you” and “and”
 - b. “what” and “how”
 - c. “we” and “why”
 - d. “describe” and “issue”

15. Effective communication requires:
 - a. non-judgmental listening, information gathering, and trust
 - b. active listening, evaluation, and empathy
 - c. critical listening, appraisal, and compassion
 - d. assessment, critique, and conviction

National Judicial College: Civil Mediation Course

October 31–November 5, 2004

Follow-up Survey

1. When asked why you were taking the Civil Mediation training, the following responses were provided. Please place an X in the box next to the responses that (1) are the reason you took the course, and (2) met your satisfaction.

Please mark all that apply.

Goal	Is a reason I took the course	Goal obtained <u>or</u> Will be accomplished
Desire certificate	[]	[]
Improve my skills and/or learn new skills	[]	[]
Want to better understand the mediation process	[]	[]
Am currently working as a mediator and wanted additional training	[]	[]
Recommended by colleague, chief judge, or prior attendee	[]	[]
Looking for a job change	[]	[]
Obtain mediator work upon retirement from current position	[]	[]

2. In retrospect, did the Civil Mediation course meet your expectations? YES [] NO []

Please describe why or why not.

3. Please describe the most important **skill** that you learned in the course.

4. Please describe the **most significant** thing that you learned in the *Civil Mediation* course.

5. Overall, are you satisfied with what you learned? YES [] NO []

Please explain your answer.

6. Have you done anything differently at your job as a result of taking the *Civil Mediation* course? YES [] NO []

Please explain your answer.

7. Do you think any of the lecture topics, exercises, or activities were a **waste of your time**?
YES [] NO []

Please explain your answer.

8. What changes do you think could be made to the Civil Mediation course to make it more effective?

9. Would you recommend the course to your co-workers? YES [] NO []

Why or why not?

APPENDIX A

NW3C FIAT Course

Instructor Classroom Training Observation

Assessment Instrument

Name of Course/Topic: _____

Location of Training: _____

Dates of Training: _____

Name of Instructor: _____

Instructor's Agency: _____

Name of Observer: _____

Date(s) Observed: _____

Training audience (check all that apply)

Type:

law enforcement prosecutor's office public defender's office

corrections other public safety _____
(describe)

law enforcement support _____
(describe)

Other: _____
(describe)

Level:

line level supervisors mid-level management

top-level management Other: _____
(describe)

Training Methods

Ratings: 1=Unacceptable 2=Poor/must improve 3=Fair 4=Good 5=Excellent

Organizational Skills

Stated objectives clearly.....	1	2	3	4	5
Defined new/unusual terms.....	1	2	3	4	5
Explained directions.....	1	2	3	4	5
Adjusted pace for complexity & understanding.....	1	2	3	4	5
Stressed important content.....	1	2	3	4	5
Managed time well.....	1	2	3	4	5
Handled materials easily.....	1	2	3	4	5
Linked to previous learning.....	1	2	3	4	5

COMMENTS: _____

Delivery Methods

Anecdotes/examples given.....	1	2	3	4	5
Case study/scenario.....	1	2	3	4	5
Demonstration.....	1	2	3	4	5
Facilitation.....	1	2	3	4	5
Learning activity.....	1	2	3	4	5
Lecture.....	1	2	3	4	5
Panel discussion.....	1	2	3	4	5
Power point presentation as outline only.....	1	2	3	4	5
Interactive power point.....	1	2	3	4	5
Questions and answers.....	1	2	3	4	5
Role playing.....	1	2	3	4	5
Video.....	1	2	3	4	5
Workbook.....	1	2	3	4	5
Problem based (begin session by presenting a problem).....	1	2	3	4	5

COMMENTS: _____

Teaching Strategies

Provided overview of lesson.....	1	2	3	4	5
Provided feedback to learners.....	1	2	3	4	5
Created transition from phase to phase.....	1	2	3	4	5
Conducted periodic review/summary of material.....	1	2	3	4	5
Included examples of references to use on-the-job.....	1	2	3	4	5
Was able to use hardware and software easily.....	1	2	3	4	5

COMMENTS: _____

Questioning Skills

Answered questions thoroughly and clearly.....	1	2	3	4	5
Questions asked were clear.....	1	2	3	4	5
Put appropriate challenge in questions.....	1	2	3	4	5
Invited questions.....	1	2	3	4	5
Handled incorrect answers supportively.....	1	2	3	4	5
Used open-ended questions where possible.....	1	2	3	4	5

COMMENTS: _____

NCLETTTC Instructor Observer Ratings (developed by EKU)

Rating Category/Item	Unacceptable	Poor	Fair	Good	Excellent	% Agreement
<i>Organizational Skills</i>						
1. Stated objectives clearly						
2. Defined new/unusual terms						
3. Explained directions						
4. Adjusted pace for complexity & understanding						
5. Stressed important content						
6. Managed time well						
7. Handled materials easily						
8. Linked to previous training						
<i>Delivery Methods</i>						
1. Anecdotes/examples given						
2. Case study/scenario						
3. Demonstration						
4. Facilitation						
5. Learning activity						
6. Lecture						
7. Panel discussion						
8. PowerPoint presentation as outline only						
9. Interactive power point						
10. Questions and answers						
11. Role playing						
12. Video						
13. Workbook						
14. Problem based						
<i>Teaching Strategies</i>						
1. Provided overview of lesson						
2. Provided feedback to learners						
3. Created transition from phase to phase						
4. Conducted periodic review/summary of material						

5. Included examples of references to use on the job						
6. Was able to use hardware and software easily						
<i>Questioning Skills</i>						
1. Answered questions thoroughly and clearly						
2. Questions asked were clear						
3. Put appropriate challenge in questions						
4. Invited questions						
5. Handled incorrect answers supportively						
6. Used open-ended questions where possible						

APPENDIX B

Training Evaluation Model Project Evaluability Questions for Site Selection Screening

A. Grantee Information

1. Grant name & number
2. Grant duration
3. Current award amount
4. Contact information

B. Training Information

1. Who is the target audience? Demographics?
2. Where does the target audience/trainees come from?
3. How is the target population identified?
4. How many trainees are there per agency?
5. What are the training modalities?
6. How many trainings are being provided? How often? When?
7. What numbers of trainees are expected, per training, and over the next year?
8. What is the structure of the training? Type of training? Purpose? Length? Content?
9. At what stage of implementation is the training?
10. How do they know the training is effective?
11. What is the training history? How long has this training been provided?
12. What other trainings does the agency do?
13. Are there any shortcomings in the training and/or delivery? If so, what?

C. Training Model

1. Is it possible to deliver a pre/post training survey (for a Level 2 evaluation)?
2. Is it possible to measure changes in the trainee's behavior (for a Level 3 evaluation)?
3. Is it possible to survey and/or interview trainees pre-training and 3+ months later (for the Level 3 evaluation)?
4. Is it possible to survey and/or interview multiple sources about trainees' performance pre-training and 3+ months later (for a Level 3 evaluation)?
5. Are training results such that there are measurable & observable outcomes (for Level 4 evaluation)?
6. Are the Level 4 outcomes measurable in the evaluation time frame?

D. Data/Measurement

1. What type of data systems exist that would facilitate evaluation?
2. What are the key elements contained in these systems?
3. Any data available for comparison/control group?
4. In general, how useful is the data systems to an impact evaluation?
5. Is random assignment possible?
6. What threats to a sound evaluation are likely to occur?

7. For the 1st site (completion of evaluation by February 2004), how large of a target and comparison sample would there be?
 8. For the other 3 sites, is completion possible by December 2004? What is the size of the target and control groups? Is there time for an organizational impact study?
 9. What is the degree of front-end planning?
 10. What are the project's outcome goals? What outcome measures are being collected?
 11. Is it possible to measure program costs? What are these measures?
 12. Is it possible to measure program benefits? What are these measures?
- E. Cooperation/Evaluation planning
1. Is the grantee interested in being evaluated?
 2. Have any evaluations already been conducted of these trainings?
 3. Is the grantee willing to cooperate with a Level 3-4 evaluation?
 4. Will the grantees' agencies cooperate with a Level 3-4 evaluation?

APPENDIX C

Synthesis Report on Evaluability Assessments of Training Programs

The purpose of this report is to provide a synthesis of the 10 evaluability assessments conducted on the potential training programs selected from the list provided by NIJ and BJA. To recap our main goal, we will be conducting four impact evaluations in order to develop and test the training evaluation model. NIJ has requested that the first impact evaluation be completed during the first year of this project (i.e., by March 2004). The final three evaluations are targeted for completion by December 2004.

Training Comparison

Project staff were assigned to review grant application materials and contact the project directors for the 10 training projects. Staff conducted lengthy interviews by phone and prepared brief evaluability reports. ILJ staff then created a spreadsheet, based on the selection criteria factors listed below, and prepared this synthesis.

The 10 training programs are:

- National Judicial College
- Simon Wiesenthal Center
- Law Enforcement Innovations Center
- Center for Task Force Training
- National Law Enforcement & Corrections Training Center
- University of Mississippi School of Law
- Littleton Area Learning Center
- Roger Williams University
- National Training & Information Center
- National White Collar Crime Center

Criteria for Selection:

The selection criteria used to judge the 10 projects are listed below.

1. Diversity of subject matter: law enforcement, courts, corrections, etc.
2. Diversity of type of audience: line level, management/leadership, administrative, technical (info tech, etc.)

3. Diversity of training objectives: skill-based, knowledge-based, behavior-oriented, etc.
4. Diversity of training modalities: classroom, self-paced on PC/CD, distance learning, etc.
5. Level of cooperation and commitment from grantee—willing to engage in Level 3-4 evaluation; also, commitment from trainees' agencies.
6. Degree of front-end planning: course based on needs assessment, selection criteria for trainees.
7. Focus of analysis: individual versus organization—number of trainees per unit at agency. In other words, if one patrol officer is being trained out of 100 patrol officers, this will not show much impact back in agency unless training is a specialty and agency is small.
8. Adequate training dosage.
9. Involvement and commitment of trainees' agencies in trainees' acquisition of the training; realistic expectations.
10. Trainees have time and opportunity after training to practice and apply new learning.
11. Attrition factors—trainees dropping out of training before finishing.
12. Cost, if any, to trainee's agency.
13. Is there a local evaluation?

Selecting the First Evaluation Site

As noted above, we are required to complete one full impact evaluation within the first year of this project. The National White Collar Crime Center (NW3C) looks promising for this first evaluation. The NW3C functions as a support mechanism for law enforcement, prosecutors, and other organizations with an active interest in the prevention, investigation, and prosecution of economic crime. The NW3C accomplishes these goals through research, training, seminars, and analytical work. The core focus of the NW3C has always been to support state and local level agencies.

The NW3C has recently developed a new training course entitled Foundations for Intelligence Analysis (FIAT). The purpose of FIAT is to develop a basic analytical intelligence training curriculum for local, state, and federal law enforcement and regulatory personnel. FIAT training is scheduled to begin in October with three classes of approximately 25 students each completed by the end of 2003. The evaluation potential seems promising for the following reasons: (1) we are able to get in during the development stages of a new course and evaluate the

process; (2) the training has technical aspects that can be applied and measured; and (3) NW3C is willing to cooperate on all evaluation levels. The ILJ research team has conducted several site visits to the NW3C and has negotiated a memorandum of understanding with them. We are currently developing the evaluation design.

Paring Down

Of the nine programs left for consideration, three need to be selected for a total of four impact evaluations. Based upon the evaluability assessments and discussion with the program interviewers, there are four programs that can be ruled out. These are:

- **National Training and Information Center**—because of the focus on “community building,” it is hard to isolate a training component. Furthermore, it would be extremely difficult to do a Level 3 or 4 evaluation.
- **Littleton Area Learning Center**—serves as a medium for training delivery, but they do not have control over training content, course selection, or attendees.
- **Roger Williams University**—the only sizeable number of trainees comes from the Providence, Rhode Island, Police Department. This would not be an adequate sample size.
- **Law Enforcement Innovations Center**—provides two types of trainings: Command College and technology training, neither of which are well suited for this evaluation. The Command College training time extends beyond the scope of this evaluation. The technology training is small in scope and it is unclear if the law enforcement personnel trained will have the resources to put their training to use.

Selecting the Second through Fourth Evaluation Sites

Thus, there are five possible training evaluation programs, from which we need to choose three programs: National Judicial College, Simon Wiesenthal Center, Center for Task Force Training, National Law Enforcement & Corrections Training Center, and University of Mississippi School of Law. The strengths and weaknesses of each are presented below.

National Corrections and Law Enforcement Training and Technology Center

Strengths:

- Corrections component would allow for model testing in another criminal justice area besides law enforcement
- Leadership component allows for broader model testing

- Trainees include middle and upper management who might tend to be cooperative with a follow-up evaluation
- On-line training component could provide a natural comparison of the same courses provided in the classroom setting
- Executive director is willing to facilitate Level 3 and 4 evaluations, including providing the trainees' incentives to participate
- Large sample size: expecting a significant increase in trainees from 2,000 in 2003 to 3,000 in 2004

Weaknesses:

- May not have enough trainees per same agency
- Is it possible to do a cost benefit analysis? It is possible to get data on costs and benefits?
- Random assignment would be difficult

National Judicial College

Strengths:

- Judicial training would allow for model testing in another criminal justice area besides law enforcement
- Already collecting data for a Level 3 type of evaluation, but are currently only using it to provide feedback to the individual judges and not using the data in an aggregate form
- Recently hired an evaluator who may be interested in a cooperative evaluation
- Have an extensive database, including demographic information on trainees, history of courses taken, size of jurisdictions, whether they are legally trained, type of cases they hear, etc.
- Agreeable to a full impact evaluation and feel "judges would be very happy to cooperate"

Weaknesses:

- Random assignment would be difficult
- Could be difficult to obtain an adequate comparison/control group
- Conducting a Level 4 evaluation or cost benefit analysis could be difficult; benefit measures might be soft

Simon Wiesenthal Center

Strengths:

- Broad criminal justice audience
- Different and unique subject matter; would add depth to training evaluation model
- Have attempted a Level 3 evaluation through extensive post-training interviews and a plenary session devoted almost entirely to collecting Level 3 assessment information
- Have collected data on impact of training on the job
- Level 3 evaluation would involve measuring if trainees form a coalition (with the prosecutor's office, schools, law enforcement, and community leaders) as a result of this training
- Have trained over 50,000 law enforcement personnel and have "saturated" some communities
- Are adding an e-learning component to the training
- While random assignment may not be possible, finding a comparison community might be possible
- Willing to cooperate, understand the importance of evaluation, and have been "waiting for a call" from evaluators

Weaknesses:

- Level 4 assessment would be a challenge
- Director does not believe pre/post (Level 2) testing would accurately measure the impact of the experience of training, however, they are willing to cooperate with us on this

University of Mississippi School of Law, National Center for Justice and the Rule of Law (2 trainings: cyber crime and 4th amendment initiative)

Strengths:

- Different audience—prosecutors and U.S. Attorney General staff
- May be possible to do a Level 4 or cost benefit analysis (of cyber crime training); more difficult to identify a comparison group

Weaknesses:

- Usually only one trainee per agency
- Willing to cooperate, but have two partner organizations that would also have to agree
- No random assignment

- No comparison/control group
- Short training dosage, few participants

Institute for Intergovernmental Research Center for Task Force Learning (Offer three types of trainings: Narcotics Task Forces, Methamphetamine Investigation Management, and Raves/Club Drug Awareness)

Strengths:

- Multiple trainings being offered in 2003-04 (14 Narcotics; 20 Methamphetamine; and 22 Raves)
- Already do a Level 2 type evaluation; we may be able to build our questions into their survey, or vice versa
- Currently collect data on tangible costs, but not benefit

Weaknesses:

- May not have adequate numbers of trainees/agency because IIR discourages multiple personnel from any one agency
- Rave/club drug training is only one day, may not have adequate dosage. However, the other two trainings are both three-day workshops
- Trainings are designed to be general, not specialized; thus Level 3 evaluation may be difficult
- Level 4 evaluation would be challenging—how do you measure organizational impact and benefit?
- While they are agreeable to an impact evaluation, they sent mixed signals about how cooperative they would be (e.g., “we conduct our own evaluation, another may not be necessary”)
- Random assignment is unlikely
- No natural comparison/control group, but might be possible if trainees’ agencies are cooperative in providing a match

APPENDIX D

NW3C FIAT Training Evaluation Plan

Brief Background

The National White Collar Crime Center (NW3C) functions as a support mechanism for federal, state, and local law enforcement, prosecutors, and other agencies with an active interest in the prevention, investigation, and prosecution of economic crime. Recognizing a shortage of intelligence analysis training and a lack of standardized courses in this area, NW3C created, in conjunction with the International Association of Law Enforcement Intelligence Analysts (IALEIA), Law Enforcement Intelligence Unit (LEIU), and Regional Information Sharing Systems (RISS), a new training course entitled “Foundations for Intelligence Analysis Training” (FIAT). The purpose of FIAT is to develop a standardized basic analytical intelligence training curriculum for law enforcement and regulatory personnel. The goal of FIAT is to be the standard-bearer course in intelligence analysis training. A logic model of the FIAT course is presented in Exhibit 1.

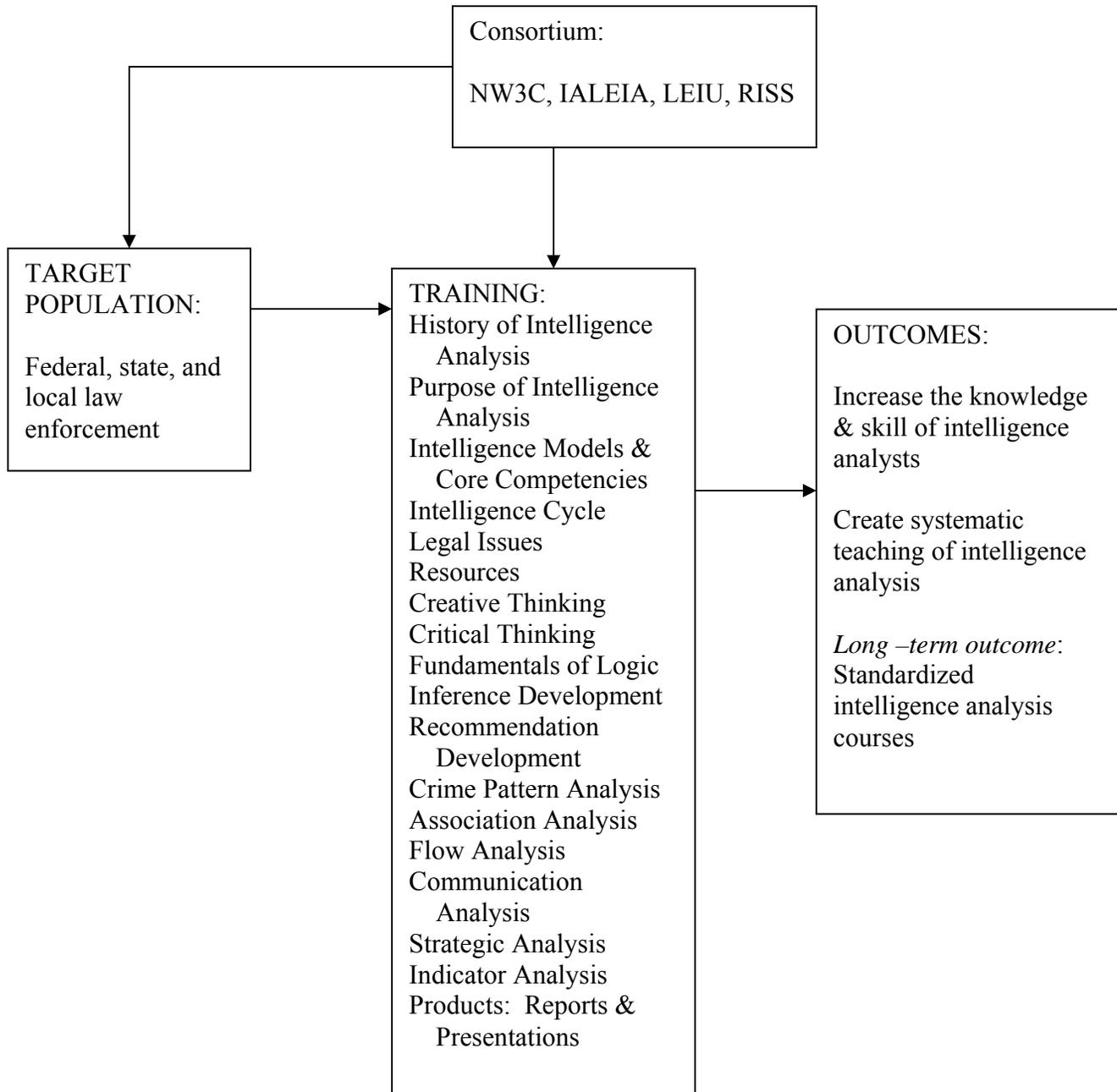
Process Evaluation

An important aspect of an impact evaluation is the process evaluation component. It is critical to determine if FIAT is performing its intended function before an impact evaluation can be completed.

Some of the key process questions that the evaluators will assess are shown on the following page.

Exhibit 1: FIAT LOGIC MODEL

GOAL: Develop standard-bearer foundational intelligence analysts course



Key Questions

- How many persons receive the training?
- Are they the intended targets?
- Are they receiving the proper amount, type, and quality of training?
- Are there targets who are not receiving training?
- Are members of the target population aware of the training?
- Is FIAT program staffing sufficient in numbers and competencies for the functions that must be performed?
- Is the FIAT program well organized?
- Does the FIAT program coordinate effectively with the other programs and agencies with which it must interact?
- Are FIAT program resources, facilities, and funding adequate to support program functions?
- Are FIAT program resources used effectively and efficiently?
- Are costs per service unit delivered reasonable?
- Is the FIAT program in compliance with applicable professional and legal standards?
- Was a needs assessment conducted?
- Are the FIAT courses redundant with other training already available?

Impact Evaluation

The key question in the impact evaluation is what are the measurable outcomes? That is, what are the FIAT participants learning (or being trained to do) that is observable and measurable? In order to measure actual behavior, and not just the participants' perception of their behavior, multiple sources of data will be collected. This includes surveys and/or interviews with the participants, their supervisors, and intelligence analyst colleagues.

Methods

Participants

The NW3C FIAT participants are drawn from local, state, and federal law enforcement and regulatory agencies who need training in criminal intelligence analysis. The NW3C offered

two classes in December 2003, averaging 21 students each. Thus far, they have scheduled two courses in 2004. Consequently, there will be at least 85 FIAT trainees and comparable matched comparison groups.

Design

The design for the NW3C FIAT evaluation will be a mixed within- and between-subjects quasi-experimental longitudinal design with a matched comparison group (see Exhibit 2). A quasi-experimental, rather than an experimental design, is necessary because the FIAT program participants are not randomly selected nor can they be randomly assigned to the training treatment condition. While true experimental designs are considered the best design for impact evaluation research, they are often extremely difficult to implement because of program restrictions and prohibitive cost. In the FIAT course, random sampling is not possible because the evaluation team does not have access to a pool of willing participants from which to sample. Rather, training participants voluntarily sign up for this course on a first come, first served basis. Accordingly, random assignment is also infeasible. In addition, true experimental designs are not generally recommended for evaluating the impact of programs that are in the early stages of implementation, such as FIAT, because new programs are often unstable as key features are modified (Rossi, Lipsey, & Freeman 2004).

The evaluation team proposes two sources of participants for the comparison groups, the first group drawn from intelligence analysts working in the same department as the FIAT participant, and the second group drawn from participants who have taken a different intelligence analysis foundational course. The first comparison group of intelligence analysts working in the same units as the training participants will be matched by key characteristics such as years of experience. The advantages to selecting this comparison group include controlling for department dynamics and workload variation, and a high participation rate because department supervisors will have already agreed to allow participation by analysts in the treatment condition. The disadvantage of selecting a comparison group from the same agency as the treatment group is the potential for contamination of the comparison group via the sharing of training materials by the treatment group. This may be compounded by the fact that the FIAT course encourages trainees to share the training materials and use a team approach in their case analysis. Consequently, the comparison group analysts may expand their repertoire of skills simply by

their exposure to the training participants, nullifying major group differences. Of further consideration is the motivation for taking the FIAT training. It may be that analysts who took the training are more self-motivated, or, it may be that they have a greater need for a foundational course in intelligence analysis because they are less experienced than their colleagues. In either case, it would be challenging to measure these differences, yet it can have a potentially profound impact on the results.

The second comparison group will be drawn from analysts who have taken a different intelligence analysis foundational course. The goal of FIAT is to be the standard-bearer intelligence analysis course. Consequently, a logical comparison group is other intelligence analysis courses. This comparison group will be comprised of participants in the “Analytical Techniques Training” course offered by the Regional Organized Crime Information Center (ROCIC), one of the six RISS centers in the United States. The ROCIC officials have agreed to allow ILJ access to their intelligence analysis course trainees. The advantage to this includes being able to evaluate if FIAT is achieving its goal as the standard setting course. The disadvantages include the possibility that the curriculum for both courses are so similar that meaningful group differences are not found. In addition, this comparison is limited by post-training surveys and interviews, rather than a pre-post design (see Exhibit 2).

Exhibit 2: Proposed Quasi-Experimental Design of the FIAT Program

	Pre-Test (knowledge survey; behavior survey)	Training	Post-Test (course reaction; knowledge survey)	2nd Post-Test (behavior survey)
Treatment	O _a	FIAT	O _a	O _a O _b
Comparison 1	O _a		O _a	O _a O _b
Comparison 2		ROCIC	O _a	O _a O _b
	T ₁	T ₂	T ₃	T ₄

Where:

O_a = Study participant measures

O_b = Supplemental measures collected from supervisor(s) & colleagues of study participants

Instruments and Procedures

Sources of data for the impact evaluation of the FIAT program include pre-post training surveys, interviews with participants, interviews with participants’ supervisors, and anecdotal data from both the treatment and comparison groups.

Level 1: Reaction

The key evaluation question for Level 1 evaluation of the FIAT program is how did the participants react to the training? Answering this will involve a post-training survey focusing on the participants’ reaction to the training, immediately following the completion of the training (see Exhibits 2 and 3). This survey instrument is standard in all the NW3C courses and adapted for use in this evaluation.

Level 2: Learning

The key evaluation question for a Level 2 evaluation of the FIAT program is what information and skills were gained by the program participants? Answering this involves a pre-and post-test survey focusing on the knowledge obtained during the training. This survey is

administered immediately prior to and immediately following training (see Exhibits 2 and 3). NW3C staff developed this survey instrument.

Level 3: Behavior

The key evaluation question for a Level 3 evaluation of the FIAT program is the extent to which program participants have transferred learned knowledge and skills to their jobs through their on-the-job behavior? Answering this will involve surveys or interviews completed both before the training begins and at least six months after the training. To truly assess behavior change, it is imperative that we not only obtain the view of the participants, but include multiple sources of data, including surveys from other analysts and supervisors. These surveys will be administered to both the study participants and their supervisors and analyst colleagues six months after training (see Exhibits 2 and 3). This survey instrument is currently being developed in conjunction with NW3C.

Level 4: Organizational Impact

The key evaluation question for a Level 4 evaluation is what effect has the training had on the organization? This is the most difficult question to assess because it is challenging to ascertain the organizational impact of a training. Kirkpatrick (1998) suggests using a behavior checklist survey completed by other intelligence analysts and supervisors. However, the foundational nature of the FIAT program does not easily lend itself to a simple behavior checklist as a measure of organization change. While the evaluation team will attempt to measure organizational level changes due to FIAT participation, we expect that a Level 4 evaluation will be a challenge.

Exhibit 3: Evaluation Questions and Data Collection Tools

Evaluation Level	Evaluation Questions	Data Collection Tool
1: Reaction	How did the participants react to the training?	Course reaction sheets
2: Learning	What information and skills were gained?	Survey administered before and after training
3: Behavior	How have participants transferred knowledge and skills to their jobs?	Survey and/or interviews administered before and after training Supervisor and employee questionnaires Anecdotal data Comparison group
4: Organizational Impact	What effect has the training had on the organization?	Survey and/or interviews Comparison group

Important Dates

- **September 16 - 18, 2003:** **FIAT Train the Trainers (Fairmont, WV)**
- **December 1 - 5, 2003:** **First FIAT training (Hershey, PA)**
- **December 8 - 12, 2003:** **Second FIAT training (Phoenix, AZ)**
- **February 16 - 20, 2004:** **Third FIAT training (Fort Lauderdale, FL)**
- **March 1 - 5, 2004:** **Fourth FIAT training (Boston, MA)**

References

Kirkpatrick, Donald. *Evaluating Training Programs. The Four Levels.* San Francisco, CA: Berrett-Koehler, 1996.

Rossi, Peter H., Mark W. Lipsey, and Howard E. Freeman. *Evaluation: A Systematic Approach.* 7th edition. Thousand Oaks, CA: Sage Publications, 2004.

APPENDIX E

Memorandum of Understanding Between INSTITUTE FOR LAW AND JUSTICE and NATIONAL WHITE COLLAR CRIME CENTER

This Memorandum of Understanding (“MOU”) is made as of August 8, 2003, between the Institute for Law and Justice (“ILJ”), and the National White Collar Crime Center, Inc. (“NW3C”).

Introduction

ILJ is a private corporation dedicated to consulting, research, evaluation, and training in criminal justice. ILJ’s mission is to bring the best of criminal justice research and practice to the field. ILJ conducts research and evaluation on policing and criminal justice. ILJ’s staff, which has both high academic credentials and field experience, has conducted more than 250 major research and evaluation projects in law enforcement and criminal justice, including evaluation of criminal justice training programs.

NW3C is a non-profit, Virginia corporation whose mission is to assist state and local law enforcement in combating economic crime, high-tech crime, and related terrorist activities. It is funded by a grant awarded by the Bureau of Justice Assistance, Office of Justice Programs, U.S. Department of Justice. NW3C is authorized to enter into this MOU pursuant to the Department of Justice and Congressional funding mandate expressed in the Consolidated Appropriations Act of 2000. Congress expects NW3C to assist in forming alliances with the public and private sectors and in providing support for multi-agency projects which address economic and high-tech issues. This MOU is made to further that intent.

Background

The parties to the MOU share an interest in supporting the efforts of state and local law enforcement, prosecutors, and other criminal justice professionals by providing specialized training effectively and efficiently.

The parties desire to work together to evaluate NW3C’s new Foundations of Intelligence Analysis Training (“FIAT”) course (the “FIAT Evaluation Project”). Both the Bureau of Justice Assistance and the National Institute of Justice are supportive of the collaboration of these parties for this project.

Therefore, the parties intend to work together to set forth the following activities, agreements, and understandings necessary to accomplish the goals and objectives of this MOU.

Goals and Objectives

The parties agree to use reasonable efforts to accomplish the following Goals and Objectives of this MOU:

A. The parties will:

1. support the accomplishment of the mission and goals of the FIAT Evaluation Project;
2. share resources and expertise necessary to evaluate the FIAT course;
3. establish a working group to coordinate and monitor the FIAT Evaluation Project's activities, to identify and resolve policy, process, and operational issues, and to ensure that mutual objectives are achieved;
4. assist the working group in developing a concept for the operation of the FIAT Evaluation Project, and provide timely acknowledgement of, and agreement with, the concept; and
5. work together to maintain an environment which supports, encourages, and requires creativity, opportunity, and excellence in the FIAT Evaluation Project.

B. ILJ will:

1. coordinate the development of the FIAT Evaluation Project and act as the lead partner in the FIAT Evaluation Project; and
2. request that NW3C provide support for the FIAT Evaluation Project, as needed.

C. NW3C will:

1. coordinate meetings and communications between the parties during the initial development phase of the FIAT Evaluation Project;
2. designate the FIAT Evaluation Project as a pilot program;
3. provide specific and customized assistance to state and local agencies associated with the FIAT Evaluation Project by providing FIAT training at no cost to the FIAT Evaluation Project;
4. provide other types of assistance as needs are identified during the development of the FIAT Evaluation Project, including, but not limited to:
 - a) In the registration process, ask prospective trainees to list the name and contact information for an immediate supervisor.
 - b) On the registration screen, inform prospective trainees that they need to agree to participate in the follow-up evaluation as part of providing feedback on the effectiveness of the training.
 - c) In the follow up evaluation, send trainees a follow-up evaluation form. Allow ILJ to review the form and possibly add some questions or request some performance data. Allow ILJ staff to contact the trainee's immediate supervisor to verify any improvements in job performance due to the training (and possibly identify in-house comparison groups who did not receive the training).
 - d) Consider using pre-/post-course testing to assess an improvement in knowledge of trainees. This would give ILJ a performance measure to show that the objective of improving knowledge was achieved. This could be done with a voluntary sample who specifically agreed to be part of the test. ILJ would be willing to separately pay some of the SMEs to develop the test instruments.
 - e) Share information with ILJ on characteristics of trainees.
 - f) Share Level 1 course evaluation findings with ILJ.
 - g) Provide all course materials to ILJ.

- h) Permit ILJ to attend the training.
- i) Permit ILJ to hold limited and brief interviews with instructors and students after the training to gain feedback from them on course implementation

Ownership, Management, and Dissemination of Information

Each party shall retain the copyright in any original work provided to another party under this MOU.

Each party agrees not to use the name of the other party or any member of its staff in any form of promotion or publicity without the written consent of the other party.

Notices

All notices or other communications required by or related to this MOU shall be in writing and shall be delivered by Express Mail, or by certified registered mail, return receipt requested, with all postage or charges prepaid. Notices or other written communications shall be addressed as indicated below, or as specified by a subsequent written notice delivered by the party whose address has changed.

To ILJ:

Institute for Law and Justice
Attn: Edward F. Connors, President
1018 Duke Street
Alexandria, VA 22314
(703) 684-5300
(703) 739-5533 (fax)
econnors@ilj.org

To NW3C:

National White Collar Crime Center
Attn: Karla Snellings, Deputy Director
7401 Beaufont Springs Drive, Suite 300
Richmond, VA 23225
(804) 323-3563, ext. 331
(804) 323-3566 (fax)
kshellings@nw3c.org

MOU Coordinators

The following individuals will serve as coordinators for their respective organizations for this MOU and for the working group: Edward Connors for ILJ, and Karla Snellings for NW3C.

Dispute Resolution

Disagreements between the parties arising under or relating to this MOU will be resolved by consultation between the parties. Any operational issues that are not resolved at the FIAT Evaluation Project Coordinator level will be forwarded to the President of ILJ and the Director of NW3C to be jointly resolved.

Term and Termination

This MOU will become effective when signed by the parties and remain in force for a period of one (1) year. It may be extended by mutual written consent of the parties' authorized representatives.

The initial timeline for the course is July 28-August 1, 2003—Beta testing; September 16-19, 2003—Instructors' Training; October—December 2003—Classes (3).

This MOU may be terminated at any time upon thirty (30) days prior written notification to the other party. In the event of termination, the parties agree to ensure termination on the most economical and equitable terms for the parties.

Drug-Free Workplace

Each party shall provide a drug-free workplace in accordance with applicable federal and state law.

No Assignment

Neither party shall assign this MOU or any right or obligation hereunder or enter into any contract or subcontract for the performance of the FIAT Evaluation Project without the express written consent of the other party.

Virginia Law

This MOU is made under the laws of the Commonwealth of Virginia, is deemed to have been executed in Virginia, and will be enforced according to Virginia law.

Entire Agreement and Amendment

This MOU is the complete and exclusive statement of the agreement between the parties with respect to the subject matter thereof, and supersedes all written and oral proposals and other communications between the parties. It may be modified only by a written agreement signed by the parties' authorized representatives.

IN WITNESS WHEREOF, each of the parties has caused this MOU to be executed by a duly authorized representative as of the date written above.

INSTITUTE FOR LAW AND JUSTICE

By _____
Signature

Name

Title

NATIONAL WHITE COLLAR CRIME CENTER

By _____
Signature

Name

Title

References

- Alaska Judicial Council. (1999). *Mediation, ADR, and the Alaska court system*. Anchorage, AK: Alaska Judicial Council.
- Allen, M.J. & Yen, W.M. (2002). *Introduction to measurement theory*. Long Grove, IL: Waveland Press.
- Anti-Defamation League. (2003). *State hate crime statutory provisions*. New York, NY: Author.
- Anti-Defamation League. (2004). *Audit of anti-Semitic incidents*. New York, NY: Author.
- Arksey, H., & Knight, P. (1999). *Interviewing for social scientists*. Thousand Oaks, CA: Sage Publications.
- Banks, D. (2005, July). Workshop on evaluability assessments, 2005 National Institute of Justice Research and Evaluation Conference. Presentation for Caliber Associates, Inc. Washington, D.C.
- Basarab, D.J., & Root, D.K. (1992). *The training evaluation process*. Norwell, MA: Kluwer Academic Publishers.
- Benjamin, D. & Irving, C. (1995). Research in family mediation: Review and implications. *Mediation Quarterly*, 13, 53-71.
- Bens, I. (2000). *Facilitating with ease! A step-by-step guidebook with customizable worksheets on CD-ROM*. San Francisco, CA: Jossey-Bass.
- Bens, I. (1999). *Facilitation at a glance!* San Francisco, CA: Jossey-Bass.
- Berg, B. (2001). *Qualitative research methods for the social sciences*. Boston, MA: Allyn and Bacon.
- Berrill, K., & Herek, G. (1992). Primary and secondary victimization in anti-gay hate crimes: Official response and public policy. In G. Herek & K. Berrill (Eds.), *Confronting violence against lesbians and gay men* (pp. 289-305). Newbury Park, CA: Sage Publications.
- Blazer, W. K., Kihm, J. A., Smith, P. C., Irwin, J. L., Bachiochi, P. D., Robie, C., Sinar, E. F., & Parra, L. F. (1997). *Users' manual for the Job Descriptive Index (JDI; 1997 revision) and the Job in General (JIG) scales*. Bowling Green, OH: Bowling Green State University.
- Bloom, B. S., Englehart, M. D., Frust, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *A taxonomy of education objectives: Handbook I, the cognitive domain*. New York: McKay.
- Brand, R. F., & Peak, K. (1995). Assessing police training curricula: "Consumer reports." *The Justice Professional*, 9, 45-58.
- Brown, S.M. (1997). *Evaluating corporate training: Models and issues*. Boston, MA: Kluwer Press.
- Buerger, M.E. (1998). Police training as a pentecost: Using tools singularly ill-suited to the purpose of reform. *Police Quarterly*, 1(1), 294-314.

- Bumgarner, J. (2001). Evaluating law enforcement training. *The Police Chief*, 68(11), 32-36.
- Bureau of Justice Assistance. (1997). *A policy maker's guide to hate crimes*. Washington, D.C.: Bureau of Justice Assistance, U.S. Department of Justice.
- Capella, R. (1994). *Planning programs for adult learners*. San Francisco, CA: Jossey-Bass Inc.
- Christopher, W., Arquellas, J. Anderson, R. Barnes, W., Estrada, L., & Kantor, M., et al. (1991). *Report of the independent commission on the Los Angeles Police Department*. Los Angeles, CA: Diane Publishing.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.)*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Comstock, G. (1989). Victims of anti-gay/lesbian violence. *Journal of Interpersonal Violence*, 4, 101-106.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation. Design and analysis issues for field settings*. Boston, MA: Houghton Mifflin.
- Cordner, G. W., Scarborough, K. E., & Sheehan, R. (2004). *Police administration (5th ed.)*. Cincinnati, OH: Anderson/LexisNexis.
- Criminal Intelligence Training Coordination Strategy Working Group. (2004). *Minimum criminal intelligence training standards for United States law enforcement and other criminal justice agencies*. Washington, D.C.: U.S. Department of Justice.
- Dillman, D. A. (1978). *Mail and telephone surveys: The total design method*. New York, NY: Wiley.
- Edwards, T. D. (1993). State police basic training programs: An assessment of course content and instructional methodology. *American Journal of Police*, 12, 23-45.
- Fay, J. (1988). *Approaches to criminal justice training (2nd ed.)*. Athens, GA: Carl Vinson Institute of Government, University of Georgia.
- Federal Bureau of Investigation. (2004). *Hate crime statistics 2003*. Washington, D.C.: U.S. Department of Justice, Federal Bureau of Investigation.
- Finn, P., & McNeil, T. (1987). *Bias crime and the criminal justice response: A summary report (prepared for the National Criminal Justice Association)*. Cambridge, MA: Abt Associates.
- Freeman, R. M. (1999). *Correctional organization and management: Public policy challenges, behavior, and structure*. New York, NY: Butterworth Heinemann.
- Glenn, R.W., Panitch, P.R., Barnes-Proby, D., Williams, E., Christian, J., Lewis, M.W., Gerwehr, S., & Brannan, D.W. (2003). *Training the 21st century police officer: Redefining police professionalism for the Los Angeles Police Department*. Los Angeles, CA: RAND Corp.
- Global Intelligence Working Group. (2003). *Core criminal intelligence training standards for United States law enforcement and other criminal justice agencies*. Washington, D.C.: U.S. Department of Justice.
- Haberfeld, M. R. (2006). *Police leadership*. Upper Saddle River, NJ: Pearson Prentice-Hall.

- Hackett, D. & Martin, C. (1993). *Facilitation skills for team leaders*. Menlo Park, CA: Crisp Publications.
- Hate Crime Sentencing Enhancement Act (HCSEA) of 1994. Pub. L. No. 103-322, 108 Stat. 1796, H.R. 1796.
- Hate Crime Statistics Act (HCSA) of 1990. Pub. L. No. 101-275, 28 U.S.C. 534, H.R. 1048.
- Head, G. (1994). *Training cost analysis*. Alexandria, VA: American Society for Training and Development.
- Hensler, D.R. (2003). Our courts, ourselves: How the alternative dispute resolution movement is re-shaping our legal system. *Penn State Law Review*, 108(3), 165-197.
- Hoffman, B. (1999). Terrorism trends and prospects. In I. O. Lesser, B. Hoffman, J. Arquilla, D. F. Ronfeldt, M. Zanini, and B. M. Jenkins (Eds.) *Countering the New Terrorism* (pp.7-38). Santa Monica, CA: Rand Corp.
- Holton, E.F. III. (1996). The flawed four-level evaluation model. *Human Resource Development Quarterly*, 7(1).
- Houston, J. G. (1999). *Correctional management: Functions, skills, and systems*. Chicago, IL: Nelson-Hall.
- International Society for Performance Improvement. (2000, April). *Human performance technology primer*. Retrieved November 14, 2002, from <http://www.afc-ispj.org/hptprimer.htm>.
- Jenness, V. (1999). Managing differences and making legislation: Social movements and racialization, sexualization, and gendering of Federal hate crime law in the U.S., 1985-1998. *Social Problems*, 46(4), 548-571.
- Jenness, V. (2003). Engendering hate crime policy: Gender, the “dilemma of difference,” and the creation of legal subjects. *Journal of Hate Studies*, 2, 73-97.
- Justice Research and Statistics Association. (2003). *Evaluability assessment: Examining the readiness of a program for evaluation*. Program Evaluation Briefing Series #6. Washington, D.C.: Office of Juvenile Justice and Delinquency Prevention, U.S. Department of Justice.
- Kaner, S. (1996). *Facilitator's guide to participatory decisionmaking*. Gabriola Island, BC: New Society Publishers.
- Kirkpatrick, D. L. (1998). *Evaluating training programs: The four levels* (2nd ed.). San Francisco, CA: Berrett-Koehler.
- Kraemer, H.C. (1987). *How many subjects? Statistical power analysis in research*. Newberry Park, CA: Sage Publications.
- Krueger, R. A. & Casey, M.A. (2000). *Focus groups: A practical guide for applied research*. Thousand Oaks, CA: Sage Publications.
- Kuehnle, K., & Sullivan, A. (2001). Patterns of anti-gay violence: An analysis of incident characteristics and victim reporting. *Journal of Interpersonal Violence*, 16(9), 928-943.

- Lambert, E. G., Barton, S. M., & Hogan, N. L. (1999). The missing link between job satisfaction and correctional staff behavior. The issue of organizational commitment. *American Journal of Criminal Justice*, 24, 95-116.
- Lerda, L.W. (1967). Conference methods. In *Training and Development Handbook*, R. L. Craig and L. R. Bittel (Eds.). New York, NY: McGraw-Hill.
- Lewin, K. (1997). *Resolving social conflicts and field theory in social science*. Reprinted by the American Psychological Association.
- Marion, N. (1998). Police academy training: Are we teaching recruits what they need to know? *Policing: An International Journal of Police Strategies & Management*, 21, 54-79.
- Maslow, A.H. (1970). *Motivation and personality*. (2nd ed.) New York, NY: Harper-Collins.
- Maxfield, M.G. (2001). *Guide to frugal evaluation for criminal justice*. Report prepared for National Institute of Justice, U. S. Department of Justice.
- Maxfield, M. G. & E. Babbie. (1995). *Research methods for criminal justice and criminology*. Belmont, CA: Wadsworth
- Maxwell, J. A. (1996). *Qualitative research design: An interactive approach* (41). Thousand Oaks, CA: Sage Publications.
- McCain, D.V. (1999). *Creating training courses*. Alexandria, VA: American Society for Training and Development.
- McNaught, J. (1991, November). The ten rules for perfect evaluations. Originally published by *Data Training Magazine*, May 1991. Retrieved November 15, 2002, from www.karinrex.com/tc_evals.htm.
- Melamed, J. (2004). *What is mediation?* Retrieved April 6, 2006, from www.mediate.com.
- Merriam, S.B. & Caffarella, R.S. (1999). *Learning in adulthood: A comprehensive guide*. (2nd ed.) San Francisco, CA: Jossey-Bass.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis* (2nd ed). Thousand Oaks, CA: Sage Publications.
- Minor, K. I., Wells, J. B., Cobb, K., Lawrence, B., & Cox, T. C. (2005). *Test validity in justice and safety training contexts. A study of criterion-referenced assessment in a police academy*. Springfield, IL: Charles C. Thomas.
- Minor, K.I., Wells, J.B. & Angel, E. (2002). *Kentucky Department of Juvenile Justice Training Evaluation: Predictors of staff turnover – phase IV analysis*. Report prepared for the Department of Juvenile Justice, Frankfort, KY.
- Mowday, R. T., Steers, R. M., & Porter, L. (1979). The measurement of organizational commitment. *Journal of Vocational Behavior*, 14, 224-247.
- Murphy, W.A., and G. Gascón. (2001, November). New training program helps LAPD meet training mandates. Alexandria, VA: *The Police Chief*.
- National Center for State Courts. (1994). *National symposium on court-connected dispute resolution research: A report on current research findings*. Williamsburg, VA: National Center for State Courts.

- National Center for State Courts. (2005). *Mediation: Overview*. Retrieved April 6, 2006, from www.ncsconline.org.
- National Coalition of Anti-Violence Programs. (2005). *Anti-lesbian, gay, bisexual and transgender violence in 2004*. New York, NY: Author.
- National White Collar Crime Center, & Federal Bureau of Investigation. (2004). *IC3 2003 internet fraud report: January 1, 2003—December 31, 2003*. Richmond, VA: Author.
- Ness, J. J. (1991). The relevance of basic law enforcement training—does the curriculum prepare recruits for police work: A survey study. *Journal of Criminal Justice*, 19, 181-193.
- Nickols, F. (2000). *Evaluating training: There is no "cookbook" approach*. Retrieved November 14, 2002, from <http://home.att.net/~nickols/evaluate.htm>.
- Nolan, J. J., & Akiyama, Y. (1999). An analysis of factors that affect law enforcement participation in hate crime reporting. *Journal of Contemporary Criminal Justice*, 15(1), 111-127.
- Parry, S.B. (1997). *Evaluating the impact of training: A collection of tools and techniques*. Alexandria, VA: American Society for Training and Development.
- Parry, S. B. (2000). *Training for results*. Alexandria, VA: American Society for Training and Development.
- Peterson, M. (1997). The role of analysis in intelligence-led policing. In A. Smith (Ed.) *Intelligence-led policing: International perspectives on policing in the 21st century* (pp. 2-4). Lawrenceville, NJ: International Association of Law Enforcement Intelligence Analysts, Inc.
- Petrosino, C. (1999). Connecting the past to the future. *Journal of Contemporary Criminal Justice*, 15(1), 22-47.
- Phillips, J. (1996, April). ROI: the fifth level of evaluation. *Technical Skills & Training*. Alexandria, VA: American Society for Training and Development. Retrieved February 2005 from <http://www.astd.org/NR/rdonlyres/D0BCF259-880D-4EEC-BF89-7F1B9A88F430/0/phillips.pdf>
- Prince II, H. T., Halstead, J. F., & Hesser, L. M. (2003) (Eds.). *Leadership in police organizations*. New York, NY: McGraw-Hill.
- RAND. (1996). *An evaluation of mediation and early neutral evaluation under the Civil Justice Reform Act*. Los Angeles, CA: RAND Corp.
- Rees, F. (2001). *How to lead work teams: Facilitation skills*. (2nd ed.) San Francisco, CA: Pfeiffer.
- Reese, R. (2005). *Leadership in the LAPD: Walking the tightrope*. Durham, NC: Carolina Academic Press.
- Reiss, A. J., Jr. (1985). Shaping and serving the community: The role of the police chief executive. In W. A. Geller (Ed.), *Police leadership in America: Crisis and opportunity*. New York, NY: Praeger.
- Roehl, J. (2002). *Evaluating a weed and seed strategy*. Washington, D.C.: Executive Office for Weed and Seed, U. S. Department of Justice.

- Rosenbaum, D. & Hanson, G. (1998). Assessing the effects of school-based drug education: A six-year multilevel analysis of project D.A.R.E. *Journal of research in crime and delinquency*, 35 (4).
- Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2004). *Evaluation: A systematic approach* (7th ed.). Thousand Oaks, CA: Sage.
- Scarborough, K.E., Christiansen, K., Cordner, G., & Smith M. (1998). An evaluation of community-oriented policing training. *Police Forum*, 11(3), 11-15.
- Schwartz, R.M. (2002). *The skilled facilitator: A comprehensive resource for consultants, facilitators, managers, trainers, and coaches*. San Francisco, CA: Jossey-Bass.
- Seiter, R. P. (2002). *Correctional administration: Integrating theory and practice*. Upper Saddle River, NJ: Prentice-Hall.
- Seymour, A., Hook, M., & Grimes, C. (2002). Hate and bias crime. In A. Seymour, M. Murray, J. Sigmon, M. Hook, C. Edmunds, M. Gaboury, & G. Coleman (Eds.). *National Victim Assistance Academy* (pp. 1-12). Washington, D.C.: Office on Victims of Crime, U.S. Department of Justice.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.
- Simon Wiesenthal Center. (2002). *Tools for Tolerance for Law Enforcement East Coast Initiative and National Institute Against Hate Crimes and Terrorism*. Proposal submitted to the Bureau of Justice Assistance, U.S. Department of Justice. Los Angeles, CA: Author.
- Smith, A. (1997). Editor's introduction. In A. Smith (Ed.) *Intelligence-led policing: International perspectives on policing in the 21st century* (p 1). Lawrenceville, NJ: International Association of Law Enforcement Intelligence Analysts, Inc.
- Spencer, K. (1998). Purposeful teaching: Design and instruction for adult learners. Baltimore, MD: Johns Hopkins University. Retrieved June 2005 from www.rcmp-learning.org/docs/ecdd1140.htm.
- Stamper, N. H. (1992). *Removing managerial barriers to effective police leadership*. Washington, D.C.: Police Executive Research Forum.
- Strategic Business Ethics, Inc. (2004). *Strategies against hate crimes and terrorism: Consensus building and a blueprint for collaborative action*. Los Angeles, CA: Author.
- Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics*. New York, NY: Harper Collins.
- United States Department of Homeland Security. (2005). *Understanding the U.S. Department of Homeland Security*. Washington, D.C.: Author.
- United States Department of Justice, Bureau of Justice Assistance. (2005). *10 simple steps to help your agency become part of the national criminal intelligence sharing plan* (Revised 2005). Washington, D.C.: Author.
- United States Department of Justice. (2004). *The National criminal intelligence sharing plan*. Washington, D.C.: Author.

- United States Department of Justice. (2005). *National criminal intelligence sharing plan* (Revised 2005). Washington, D.C.: Author.
- United States General Accounting Office. (1998). *Performance measurement and evaluations: Definitions and relationships*. Washington, D.C.: Author.
- United States Office of Personnel Management. (2000). *A guide to strategically planning training and measuring results*. Washington, D.C.: Author.
- Usova, G. M. (1997). Effective test item discrimination using Bloom's taxonomy. *Education*, 118, 100-111. Retrieved June 5, 2000 from EBSCO host MasterFILE Premier (Item: 9711215349).
- Van Tiem, D. (2002) . *Do trainers need to use performance technology?* Alexandria, VA: American Society for Training and Development. Retrieved November 1, 2002, from http://www.astd.org/news_letter/October.
- Violence Against Women Act of 1994, Pub. L. No. 103-322.
- Waagen, A.K. (2000). *Essentials for evaluation: Evaluation and research*. Alexandria, VA: American Society for Training and Development.
- Watson, C.E. (1979). *Management development through training*. Reading, MA: Addison-Wesley
- Weaver, G.S., & Wittekind, J. E. C. (2002). Categorizing September 11: Casualties as acts of terrorism. *Homicide Studies*, 6 (4), 372-376.
- Welsh, B.C., Farrington, D. P. , & Sherman, L.W. (2001). *Costs and benefits of preventing crime*. Boulder, CO: Westview Press, Inc.
- Weiss, C. (1998). *Evaluation* (2nd ed). Upper Saddle River, NJ: Prentice Hall.
- Yates, B.T. (1999). *Measuring and improving cost, cost-effectiveness, and cost-benefit for substance abuse treatment programs: A manual*. Washington, D.C.: U.S. Department of Health and Human Services.
- Yeend, N.N. (2003). *Comprehensive mediation*. St. Petersburg, FL: The John Paul Jones Group.
- Yin, R. (2003). *Case study research: Design and methods* (3rd ed). Newbury Park, CA: Sage Publications.