

This Page Intentionally Left Blank

**Proceedings
of the
Ninth Annual
National Human Services
Training Evaluation Symposium**

The material in this volume is based on presentations and discussions at the Ninth Annual National Human Services Training Evaluation Symposium, held May 24–26, 2006, at the University of California, Berkeley.

*This volume is also available online at
<http://calswec.berkeley.edu/CalSWEC/CWTraining.html>.*

Co-sponsors

California Social Work Education Center (CalSWEC)
California Department of Social Services
National Staff Development and Training Association
of the American Public Human Services
Association, Washington, D.C.
American Humane Association, Englewood, CO

Published by

California Social Work Education Center
University of California, Berkeley
School of Social Welfare
Marchant Building, Suite 420
6701 San Pablo
Berkeley, CA 94720–7420
<http://calswec.berkeley.edu>



Editors
Barrett Johnson
Mark Thibedeau

Copyright 2007 by the Regents of the University of California

Table of Contents

Acknowledgments	1
-----------------	---

Introduction	2
--------------	---

The Accidental Human Services Training Evaluator <i>Jane Berdie and Dale Curry</i>	4
---	---

Training Evaluation as Quality Improvement: Moving from Theory to Reality <i>Henry Ilian</i>	11
Discussion	22

Comparing On-Line with Classroom Training <i>Cindy Parry and Susan Brooks</i>	25
Discussion	40

Using the Logic Model to Conceptualize and Enhance Training Evaluation Design <i>Alan J. Dettlaff and Joan R. Rycraft</i>	43
Discussion	55

Evaluating Trainees' Documentation Skills <i>Megan E. Paul</i>	58
---	----

Performance Based Contracting Aftercare Training and Evaluation—Lessons Learned in Aftercare <i>June Cairns, John H. Trudeau, Amanda Schreiber, and Maria G. Ramos</i>	73
Discussion	81

Using POLYDETECT in Assessing the Possible Number of Constructs in Social Work Surveys <i>Basil Qaqish</i>	83
Discussion	96

Using Interpersonal Process Recall (IPR) to Improve Child Welfare Worker Effectiveness <i>Patricia G. Taylor</i>	98
Discussion	107

Evaluating Values and Attitudes: Exploring the Impact of Training on Disproportionality and Disparity <i>Barrett Johnson, Leslie W. Zeitler, and Chris Mathias</i>	111
Discussion	128

Wrap-up and Evaluation	133
------------------------	-----

Program	143
---------	-----

Presenters and Participants	149
-----------------------------	-----

This Page Intentionally Left Blank

	Acknowledgments

The NHSTES would never be possible without a dedicated group of individuals and organizations to guide the planning and assist with the logistics and program. We are especially grateful to the co-sponsors of the event: the California Department of Social Services, the National Staff Development and Training Association of the American Public Human Services Association, and the American Humane Association.

Each year a very active steering committee provides sage guidance on the symposium program, as well as other aspects of the event. This year's members included Becky Antle, Jane Berdie, Sherrill Clark, Dale Curry, Midge Delavan, Bill Donnelly, Michelle I. Graef, Kyle Hoffman, Henry Ilian, Chris Mathias, Cynthia Parry, E. Douglas Pratt, Leslie Zeitler, and myself.

I would especially like to thank Leslie Zeitler, who again provided overall coordination of the symposium. Her organizational skills and guidance were truly indispensable. Other CalSWEC staff who assisted in the planning and implementation of the symposium included Monica Asfura, Amy Benton, Erika Cain, Terry Jackson, John Momper, Karen Ringuette, Mark Thibedeau, and Jennifer Uldricks. Mark Thibedeau provided great assistance in editing the proceedings.

Of course, special thanks also goes out to our presenters and facilitators, whose thought and preparation stimulated all of the great discussions.

Barrett Johnson, *LCSW*
Regional Training Academy Coordinator
CalSWEC
Principal Editor of the *Proceedings*

	Introduction: Measuring the Progress of the Symposium

Throughout my years at CalSWEC, I have come to look forward to the NHSTES, as each year it provides me one of the best opportunities to concentrate on the complexities of one specific area of training—evaluation—with a group of thoughtful and smart people. This year’s discussions caused me to focus my attention backward as well, perhaps because we are approaching our 10th anniversary.

Though we moved the symposium to a new venue this year, Jane Berdie and Dale Curry’s opening exercise helped tie us to our past symposia. They were able to formulate an entire trivia game based on our collective memories and stories, proving that this annual event has developed a robust culture of its own. Clearly we have made substantial progress toward building a community of training evaluators in the human services.

The presentations and discussion on the following days reinforced this for me, and also emphasized the great strides we have made as a distinct field of evaluation. Dr. Kraiger’s carefully tailored presentation pushed us to consider different models and types of evaluation based on the different needs of a given project or training. He also surprised me a bit when he said that upon reviewing our past proceedings, he felt our work in the human services was at least as complex as the training evaluation models in the private sector. All of our collective work has produced a great body of knowledge—evidence of the long way we have come since convening almost ten years ago and deciding to speak in a common language using Kirkpatrick’s levels.

The level of discussion this year focused on some of the most complex and nuanced issues of evaluation: how bias impacts the responses to knowledge questions; how we begin to measure the impact of training on values and attitudes; how the relationship between the evaluators and the administrators can impact the effectiveness of an evaluation plan; how we consider the ethical and practical concerns of high-stakes knowledge testing. Each of these topics pushes all of us to advance our training evaluation

skills to another level. They also assist in clarifying the role of training in relation to practice, by helping us to begin parsing out the role of training in areas that involve many other personal and organizational factors.

So, enjoy the reflection that the 2006 Proceedings provokes—and we hope to see you all for our 10th anniversary in 2007.

Barrett Johnson, *LCSW*
Regional Training Academy Coordinator
CalSWEC
Principal Editor of the *Proceedings*

	The Accidental Human Services Training Evaluator

Jane Berdie, *M.S.W.*, and Dale Curry, *Ph.D., LSW*

What do we know about ourselves as a training group? What is the career pathway to becoming a human services training evaluator (HSTE)? Currently, we know of no formal academic degree programs specifically pertaining to human services training and development (HSTD). There also seems to be little emphasis within human services degree programs, such as social work, concentrating on training and development. There are a growing number of human resource development programs in the academic arena, but these programs do not emphasize human services. So where do HSTD practitioners receive their training? More specifically, where does an HSTE receive one's training? What is the route (career path) taken to become a competent HSTE?

Given the lack of training programs (and the lack of research regarding the career paths of current HSTEs), we surmise that most have arrived in the field "accidentally." We know of no current HSTEs who as a children told their parents that one day they would be an HSTE. (*This notion was reinforced during the opening icebreaker exercise of the Ninth Annual National Human Services Training Evaluation Symposium.*)

In general, available training opportunities in HSTD/HSTE is not extensive; however there are some useful models emphasizing the role of the trainer. For example, statewide training programs in Ohio have developed trainer certification processes (Curry, Lawler, & Bernotavicz, 2006). In Ohio, a collaborative partnership of the Northeast Ohio Regional Training Center and Kent State University developed an 18-month Human Services Trainer Development Certificate Program. In Pennsylvania, the University of Pittsburgh School of Social Work administers the statewide child welfare training program as well as coordinates the Title IV-E-funded child welfare BSW and MSW programs. Although these programs have emphasized the development of human services

trainers, these types of collaborative agency-university relationships provide potential for future development of HSTEs. However, training for HSTEs has not yet been emphasized by these statewide training programs and university partnerships.

The National Staff Development and Training Association (NSDTA) also provides training for HSTD practitioners. The NSDTA Standards Committee has developed a competency model that includes nine roles and functions, including: administrative support, communications specialist, evaluator/researcher, human resource planner, instructional media specialist, manager, instructor/trainer, organizational development specialist, and training program and curriculum designer (Bernotavicz, 2004). The importance of training evaluation has been recognized by the NSDTA Board, resulting in the establishment of an Evaluation Committee that has (1) sponsored several full-day pre-Institute sessions on training evaluation, (2) assisted in the development of the publication *Training Evaluation in Human Services* (Parry & Berdie, 1999), and (3) co-sponsored the National Human Services Training Evaluation Symposium (NHSTES) since its inception in 1998.

The annual NSDTA Conference/Institute also serves as a valuable forum for participants to share knowledge of staff development and training issues. Workshops on training evaluation are common events at the Institute. The NSDTA has begun to build upon the competency model by establishing a Certification Committee charged with exploring the possibility of an NSDTA trainer certification process. The committee's work could result in a systematic process for HSTD professional development for the instructor/trainer role. However, discussions regarding a systematic career path for the HSTE have not yet been broached.

Training opportunities are fewer for HSTE professional development; the annual National Human Services Training Evaluation Symposium appears to provide the most relevant opportunity.

Systematic Body of Knowledge

According to Curry, Lawler, and Bernotavicz (2006), much of the HSTD knowledge base has been borrowed from other fields and adapted for HSTD application. For example, they cite several well-established scholarly journals within the broader field of

human resource development, including *Human Resource Development Quarterly*, *Human Resource Development International*, and *Advances in Developing Human Resources*. Other related journals include *The Training and Development Journal*, *Training*, *The OD Journal*, *International Journal of Training and Development*, *Adult Education Quarterly*, and *Performance Improvement Quarterly*. Several other training and development magazines and newsletters as well as a variety of books also inform practice; the American Society for Training and Development has an extensive listing of such publications. Much of what we know about needs assessment, training objectives, curriculum development, implementation of training (including presentation skills), transfer of learning, and training evaluation has built upon or been adapted from these and other related sources.

The body of knowledge specific to the HSTD field appears to be limited to a few journals such as *Journal of Social Work Education* (includes training and development articles), *Professional Development: The International Journal of Continuing Social Work Education*, and the NSDTA journal *Training and Development in Human Services*. NSDTA also has developed other publications such as *The NSDTA Code of Ethics for Training and Development Professionals in Human Services: Case Scenarios and Training Implications*, the *Human Services Staff Development and Training Roles and Competencies*, and *A Key to Success: Guidelines for Effective Staff Development and Training Programs*.

HSTD articles also continue to be published in human services journals such as *Child Welfare* and *Children and Youth Services Review*. The journal *Protecting Children* recently devoted a special issue to training and development issues. In addition, there are a variety of newsletters such as the *Protection Connection* by the Protective Services Training Institute of Texas and *Common Ground* by the Ohio Child Welfare Training Program. *CYC-net*, an international website for child and youth care professionals, devotes a special section to training in its online monthly magazine *CYC-Online*. The *Proceedings of the National Human Services Training Evaluation Symposium* has also provided valuable information specific to HSTD research and evaluation. HSTD-specific knowledge has been generated in many areas such as:

models of transfer of learning (Curry, Caplan, & Knuppel, 1994; Curry, 2001); human services trainer development (Curry & Chandler, 1999; McCowan, Wegenast, & McCowan, 2000; Peterson & Fox, 2001); evaluating transfer of learning (Curry, 2001); needs assessment; personnel selection (Curry & Rybicki, 1995; Bernotavicz & Wischmann, 2001; Graef, Rohde, & Potter, in press); models for statewide training evaluation (Parry & Berdie, 1999; Leake, Berdie, Parry, & Jones Kelley, 2001; Lindsey & Ahmed, 2002); agency-university education and training partnerships; staff retention; and training in specific content areas such as risk assessment and decision-making (Ayers-Lopez, S., Sanderson, M., & Richardson, J., 2001), case planning, and court testimony training (Philp, Cearley, Wright, & Burry, 2002; Wright & Ross, 2001).

Though far from systematic, the HSTD knowledge base is growing. The federal government has demonstrated interest, as evidenced by several projects funded by the U.S. Department of Health and Human Services Children's Bureau. This includes annual child welfare training grants, the child welfare workforce and training resources section of the National Clearinghouse on Child Abuse and Neglect Information (now incorporated into the Child Welfare Information Gateway); and the National Resource Center on Child Welfare Training and Evaluation at the University of Louisville. Continued investment in HSTD evaluation, research, and knowledge dissemination is essential for further development of the growing knowledge base.

Although the HSTD knowledge base is developing, many HSTEs most likely received their core evaluation knowledge and skills from non-HSTD sources. A relatively vast array of resources is available within the evaluation research field. Numerous evaluation journals and newsletters exist, including the *American Journal of Evaluation*, *New Directions for Evaluation*, and the *Canadian Journal of Evaluation*. A variety of professional associations pertaining to evaluation also exist, including the American Evaluation Association and the Canadian Evaluation Association. Evaluation programs also exist in higher education. The American Evaluation Association web site provides a listing of many of these programs.

Moving forward, although the field of HSTD is probably a long way from systematic development of career pathways for

HSTEs, the important task of training evaluation should not be left to “accidental human services training evaluators.” Some immediate steps toward understanding and further developing these pathways include:

1. *Conduct research regarding the backgrounds of current HSTEs and their pathway(s) to becoming an HSTE.* This may provide some valuable needs assessment information for recruitment, retention, and on-going training and development.
2. *Recognize and value the diversity of the educational and experiential backgrounds of HSTEs while helping to promote a professional identification with the field of HSTD.*
3. *Continue to strengthen the HSTE knowledge base while also continuing to create linkages to other relevant fields such as human resource development (HRD) and program evaluation.* As stated previously, we borrow from other fields. However, the field of HSTD also has made significant developments that appear unrecognized by other fields, such as HRD. A plan to disseminate HSTE and HSTD knowledge to other related fields is also indicated.

References

- Ayers-Lopez, S., Sanderson, M., & Richardson, J. (2001). "Advanced risk assessment training: Does it make a difference?" In B. Johnson, K. Kietzman, & K. Ringuette (Eds.), *Proceedings of the Fourth Annual National Human Services Training Evaluation Symposium*, (pp. 132-149). Berkeley, CA: California Social Work Education Center.
- Bernotavicz, F., & Wischmann, A.L. (2001). "Hiring child welfare caseworkers: Using a competency-based approach." In B. Johnson, K. Kietzman, & K. Ringuette (Eds.), *Proceedings of the Fourth Annual National Human Services Training Evaluation Symposium*, (pp. 12-27). Berkeley, CA: California Social Work Education Center.
- Curry, D. (2001). Evaluating transfer of learning in human services. *Journal of Child and Youth Care Work*, 15-16, 155-170.
- Curry, D.H., Caplan, P., & Knuppel, J. (1994). Transfer of training and adult learning (TOTAL). *Journal of Continuing Social Work Education*, 6, 8-14.
- Curry, D., & Chandler, T. (1999). The human services training effectiveness postcard (HSTEP): A tool for research and evaluation of human services training. *Professional Development: The International Journal of Continuing Social Work Education*, 2, 43-56.
- Curry, D., Lawler, M., & Bernotavicz, F. (2006). The field of training and development in human services: An emerging profession? *Training and Development in Human Services*, 3, 5-11.
- Curry, D., McCarragher, T., & Dellman-Jenkins, M. (2005). Training, transfer, and turnover: Exploring the relationship among transfer of learning factors and staff retention in child welfare. *Children and Youth Services Review*, 27, 931-948.
- Curry, D.H., & Rybicki, M., Sr. (1995). Assessment of child and youth care worker trainer competencies. *Journal of Child and Youth Care Work*, 61-73.
- Graef, M.I., Rhode, T.L., & Potter, M.E. (in press). An in-basket test for child protective services trainees. *Training and Development in Human Services*. National Staff Development and Training Association of the American Public Human Services Association. Washington, D.C.

- Leake, R., Berdie, J., Parry, C., & Jones Kelley, K. (2001). "A comprehensive evaluation of a statewide child welfare training system." In B. Johnson, K. Kietzman, & K. Ringuette (Eds.), *Proceedings of the Fourth Annual National Human Services Training Evaluation Symposium*, (pp. 117-127). Berkeley, CA: California Social Work Education Center.
- Lindsey, E.W., & Ahmed, F.U. (2002). Strategic planning for evaluation of training: A case study. *Training and Development in Human Services*. National Staff Development and Training Association of the American Public Human Services Association. Washington, D.C.
- McCowan, R., Wegenast, D., & McCowan, S.C. (2000). *Training and Development in Human Services, 1*, 82-85.
- Parry, C., & Berdie, J. (1999). *Training Evaluation in the Human Services*. National Staff Development and Training Association of the American Public Human Services Association. Washington, D.C.
- Peterson, N.A., & Fox, T.C. (2001). "Knowledge test construction and validation: Evaluation of New Jersey's interdisciplinary training for public agency caseworkers to improve child welfare services." In B. Johnson, K. Kietzman, & K. Ringuette (Eds.), *Proceedings of the Fourth Annual National Human Services Training Evaluation Symposium*, (pp. 32-59). Berkeley, CA: California Social Work Education Center.
- Philp, J.D., Cearley, S.T., Wright, L., & Burry, C.L. (2002). Writing for court: An evaluation of training outcomes for child welfare workers. *Training and Development in Human Services*.
- Wright, L., & Ross, J.W. (2001). "The logic model: An a-to-z guide to training development, implementation, and evaluation." In B. Johnson, K. Kietzman, & K. Ringuette (Eds.), *Proceedings of the Fourth Annual National Human Services Training Evaluation Symposium*, (pp. 153-167). Berkeley, CA: California Social Work Education Center.

	Training Evaluation as Quality Improvement: Moving from Theory to Reality

Henry Ilian, *D.S.W.*

Abstract

Training evaluation can function as a means of quality improvement aimed towards improving training content and delivery. However, the results of the evaluation must be put to use in order for training evaluation to function in this way. This is not automatic, and activities initiated by the training evaluator toward mainstreaming of evaluation are often necessary to ensure that results are used. This paper presents a case history describing a progression of steps, still in progress, toward mainstreaming training evaluation. It highlights the role of individual initiative in each step of the progression in shaping the outcome through the successive stages.

Introduction

In distinguishing between evaluation and quality improvement (QI), Weiman (2004) notes that evaluation is linear and time limited while QI is circular and continuous. With ongoing evaluating of a continually repeated training program, this difference evaporates. The aims of evaluation and QI become the same—to improve the content and delivery of the training. This corresponds with two of the purposes that Kraiger (2002) delineates for training evaluation: to provide feedback to course designers, trainers or trainees; and to facilitate decision making about continuing, modifying or dropping a course.

In order for training evaluation to function as a means of quality improvement, the results of the evaluation must be put to use. This is an area that is rarely addressed in the extensive literature on training evaluation. Models of training evaluation

described in recent surveys (Donovan & Hannigan, 1999; Eseryel, 2002; Alvarez, Salas & Garofano, 2004) generally assume that the evaluation, whatever it includes and whatever it measures, will be used. Yet, as Kopczynski and Pritchard (2004) report, in nonprofit organizations evaluation suffers from a general lack of use. Hatry, Wholey, and Newcomer (2004), extend this observation to government as well.

One reason that evaluations go unused is the isolation of the evaluator from other organizational activities. Bell (2004) points to several ways in which isolation can result in the evaluator's misunderstanding the organizational context, making it less likely that the findings will be relevant. Isolation can also result in the lack of a means for acting on evaluation. To remedy this, Eseryel (2002) maintained that "for evaluations to have a substantive and pervasive impact on the development of training programs, internal resources and personnel such as training designers, trainers, training managers, and chiefs of personnel will need to become increasingly involved as program evaluators." From the standpoint of the evaluators, Bell (2002) and Torres, Preskill, and Piontek (1996) urge close collaboration with program staff, which in the case of training evaluation translates to trainers and training managers. Activities to improve the usefulness of training evaluation and to promote the use of findings can be conceptualized as mainstreaming evaluation, defined by Sanders (2003) as "the process of making evaluation an integral part of the organization's everyday operations."

Mainstreaming Evaluation and Organizational Change from the Middle

In a traditionally structured organization, such as most government agencies or human service organizations, change normally originates from the top. Mainstreaming evaluation, however, originates with, and is the responsibility of, the evaluator. Moreover, mainstreaming evaluation requires autonomous action on the part of both the evaluator and those who collaborate in the process. This creates a challenge. When an initiative originates from above, the willingness and ability of appropriately situated individuals to become involved is assumed, if not always assured; when it originates elsewhere, involvement can be problematic. Nevertheless, while change initiatives in the public sector usually

originate at the top, this is not always the case. Anderson (1991), for example, describes the development of a structure to promote and recognize quality in state government. It was originated by an informal network of mid-level people in several state agencies along with individuals from the private sector.

While theorists such as Argyris (1955; 1957) have focused on the role of traditional organizational processes in fostering dependency, and consequently passivity, others have described a more dynamic environment in which autonomous action, and hence this type of innovation becomes possible. One contributing phenomenon is the degree to which individuals, in concert with coworkers and supervisors, are able to shape and redefine their jobs. Chatman (1989) noted that because of leaders' and subordinates' personal characteristics, tasks may change over time and can become more exceptional. This is underscored by Kohn and Schooler (1978), whose findings led them to suggest that individuals with a greater degree of intellectual flexibility may, over time, modify their original jobs to increase the level of complexity. Additionally, Miner (1987) describes what she calls "idiosyncratic jobs." These may either be "evolved jobs," whose duties reflect the interests or abilities of current employees, or "opportunistic hires," which reflect the interests of particular supervisors.

The ability to reshape a job can have various implications for the organization. Mechanic in 1962 (2003) drew a distinction between formal and informal power. He pointed out that "lower participants in an organization can often assume and wield considerable power which is not associated with their position as formally defined within these organizations." He discussed the considerable leeway for participants at organizational levels below top management, including those without formal power, to manipulate rules and trade favors to gain control of aspects of their jobs. This includes restructuring their jobs to their own advantage, to the organization's, or to both. Moreover, the fact of noncompliance with managerial initiatives or organizational policies underscores the latitude that individuals in complex organizations often have to define and control the context of their work. An example of this is Anderson and Johnson's (2005) description of differential compliance with a performance appraisal system by supervisors in three segments of a large state university.

In the opposite vein, Moorman and Blakeley (1995) describe what they called organizational citizenship behaviors. These are on-the-job behaviors “which are discretionary, not formally or directly recognized by the organizational reward system, yet promote the effectiveness of the organization.” Likewise, Cross, Baker, and Parker (2003) looked at the role of people they called “energizers” within organizations. These are people who are likely to have their ideas considered and put into action, to motivate others to act and to attract the commitment of other high performers. They contrast energizers with people they term “de-energizers,” who may have expertise and ideas, but whom others avoid and spend time complaining about.

While individuals within an organization may exercise autonomy for their own or the organization’s benefit, innovations within an organization can come in the form individuals working in concert to create what Gilmore and Krantz (1991) call “parallel processes.” These develop outside of established organizational roles and structures. They can have disruptive effects or “be used developmentally as scaffolding to enable the emergence of something more enduring” (Gilmore & Krantz, 1991).

This paper presents a case history (Stake, 1995) that describes an initiative originating at the mid-level of an organization to bring about small-scale organizational change. It describes a progression, still in process, of steps toward mainstreaming training evaluation and presents a much simplified account of a several-year period concentrating on trainee testing, one component of a range of evaluation activities. It recounts a progression in which the focus of testing broadened from a concern primarily with trainee performance to a means of both assuring the fidelity of the training to the written curriculum and of assessing the adequacy of the curriculum content. Recognizing that evaluations produced information on the content and delivery of training, but that no mechanism existed for systematically reporting or using that information, the Assessment and Evaluation Department (A&E) of the New York City Administration for Children’s Services’ (ACS) James Satterwhite Academy for Child Welfare Training worked to address the problem of the use of evaluation results. In doing so, it was able to mainstream evaluation so as to have a greater influence on how training is delivered.

Child Protective Services Training in New York City

Training for child protective workers in New York City is the product of a long history of collaboration and negotiation among many actors with differing priorities and viewpoints. Over time, training has been subject to frequent changes to meet new conditions, resulting in changes to the training plan and complete or partial redesigns of training content. Among those who have had a role in shaping or reshaping child welfare training have been the Satterwhite Academy itself, the New York State Office of Children and Family Services (OCFS), a succession of ACS commissioners, the ACS Division of Child Protection, the courts, child welfare advocates, and a variety of consultants brought in to develop and deliver various aspects of the training.

In 1996, ACS embarked on a period of comprehensive reform. One area cited in the original blueprint for this reform (*Protecting the Children of New York, 1996*) was under-skilled and inadequately trained front-line staff, caseworkers, and their immediate supervisors. Early in the reform cycle, the caseworker pre-service training period was lengthened, and supervisory and managerial training, which had heretofore been available sporadically, were reintroduced. This was followed by a comprehensive training plan for all front-line staff: the Common Core Training System (CCTS). Under the CCTS, all front-line staff, caseworkers, supervisors, front-line managers, and clerical workers would receive training in a set of human-service helping skills, the Common Core. Caseworkers received further training in a child welfare specialty, e.g. child protection, foster care, preventive services, etc. In 2001, Common Core training for child protective caseworkers, supplemented by specialty training in child protection, replaced the existing pre-service training. This has been, with some variation, the model for child protective worker training since that time.

The Test

Since 1995, high-stakes testing has been mandated at the end of the in-class segment of new CPS caseworkers' training. Currently, it constitutes the major component of the assessment of trainees' readiness for practice. Consequently, there has been concern at all levels over whether the test fairly assesses mastery of the content, whether too many or—especially given the need to

move a sufficient number of trainees to caseloads—too few trainees are passing the test. The need to move a sufficient number of trainees to caseloads contributes to this concern. A related concern has been the degree to which the average test score for a class as well as the answers to individual test items reflect the effectiveness of individual trainers and the Academy as a whole.

Stages in the Progression

During the first stage in the progression, A&E periodically initiated meetings with experienced trainers to ensure that test questions reflected the content trained. These were usually with one or two trainers or curriculum writers to review small numbers of newly written or rewritten questions. They were usually held every few months on no fixed schedule. Occasionally, there were meetings with larger groups to review sets of questions or the entire test, since training content and emphasis shifted over time. A&E's concern during this period was twofold, to be sure that questions reflected curriculum content and to replace questions that were not doing well. This stage lasted for several years.

At that time, there was a general awareness within the Academy that areas of content were trained differently by different training teams and some were routinely glossed over. This was apparent in test item analyses, which nevertheless did not influence what went on in the classroom. Instead, for a period, questions that a majority in a particular class got wrong were dropped from the scoring for that class. This was done so that trainees wouldn't be held responsible for what hadn't been trained.

In 2004, an increase in new caseworker hiring necessitated holding simultaneous classes. This formed the backdrop for the second stage, rethinking assumptions about the test and about individual test questions. In preparing item analyses for one pair of classes, the differences in percentages correct on a large number of questions was striking. For example, on a question that asked the best type of questions to use to help a 3-to-5-year-old clarify his/her concerns, 95% of one class answered correctly versus only 46% of the other class. A&E met with the two training teams to review the differences. It turned out that training teams rarely had the opportunity to compare notes on what they were training. This was the first time they had the opportunity to see what trainees were and were not understanding, as reflected by the test questions.

This experience was repeated a few months later with a different pair of teams.

The arrival of a new Training Director marked the beginning of the third stage. When A&E approached her to explain the Department's function and to go over recent item reports, it did so in a way that was perhaps more thorough or careful than previously with managers who were presumed to know what A&E did and had to offer. This began a collaboration in which the test could be used to bring attention to content areas insufficiently covered.

Rather than meet with training teams as they completed a class, A&E was now invited to meet with the teams immediately before a new class began to brief them on the content covered by questions that repeatedly had not done well. At first, A&E used PowerPoint slides that showed the stems of the questions along with the correct and most frequently chosen incorrect answers as well as the percentage of trainees who picked each. It also included a brief explanation of why the correct answer was correct. Later on, concern with "teaching to the test" led to presentations that focused more generally on item content without including the text of the question. As a result of the briefings, consistency improved among the teams.

A crisis within child protection in New York City necessitated the next stage in the progression. At the end of 2005, a series of deaths of children known to the agency led to an effort to strengthen its ability to respond to reports of abuse and neglect. One part of this effort involved relieving the pressure on casework staff created by high turnover. With the intention of stabilizing the workforce ACS condensed its existing Child Protective training from 36 to 29 days to bring new child protective staff to the front lines faster. This necessitated a major revision of the curriculum and the test. At A&E's initiative, the test writers joined the discussions on curriculum development from the beginning instead of being given a finished curriculum to write questions from. Because A&E was able to work more closely with the curriculum development team than ever before, there was closer calibration of the test with the curriculum.

Several months later a major breakthrough occurred, leading to the current stage. At the joint urging of the Director of Training and A&E, the Academy convened a panel on the test with

representatives from A&E, the Training and Curriculum Departments and the ACS Department of Child Protection. The purpose of the panel, which meets twice yearly, is to review the test and test development and to cement the link between the test, the training, and practice in the field.

Development of New Reporting Formats

During this time, A&E developed and refined a variety of reports directed toward individual trainers and managers. The reports covered trainee scores, item percentages by subject area (both for individual classes and across classes), and trainee ratings of various aspects of the training (for individual classes and comparatively). In order to assure that the reports were used A&E held a series of discussions with the Director of Training, which resulted in a system for distribution and, more importantly, review. A&E created an on-line folder and e-mailed training supervisors as the reports were ready. Trainers and supervisors would then get them from the folder and discuss them, allowing the information to feed into decisions on training strategy and emphasis for coming classes. Likewise, A&E also developed separate reports for managers to inform broader decision making.

Current Questions, Struggles:

The Role of Individuals in the Process

It is important to keep in mind that what has been described is a process which continues to develop. For example, many of the elements of the formal procedures for administering tests, analyzing and distributing results, and acting on information are in place. Others are not yet or are only beginning to be. Of the latter, a means to use test results to inform trainer activities in the classroom, and to fully situate training evaluation as a means of quality improvement, is only present in a rudimentary form. The test panel has met twice and is only beginning to prove itself.

The existence of willing partners among key decision makers who were able to see its value has been essential to the ability to carry the process forward. Without partners who were interested in what A&E could contribute, the described process would not have been possible. The process was also enhanced by the synergy that has developed over time among members of the A&E Department. A corollary of this is that until mainstreaming of training

evaluation is fully institutionalized, and possibly even after, a change in personnel or in organizational direction could jeopardize gains achieved over many years.

A further corollary is that the existence of willing partners was, at least in part, conditioned by A&E's efforts to seek partners and to educate key persons within and outside the Academy on the potential of evaluation to improve training. Organizational changes in direction and organizational crises pose both threats and opportunities. Prior groundwork by A&E made it possible, when a crisis necessitated changes in the training, to obtain a hearing and to be included in planning for the changes. Moreover, A&E's collaboration at that time led to better understanding of what evaluation was able to contribute, which built support for additional mainstreaming efforts.

At the outset of the mainstreaming process, A&E thought expansively about its role in the development and delivery of training. At the time, nobody else at the Academy saw training evaluation in that role. It was only when A&E began to conceptualize its role more broadly that it found means to actually broaden it, which benefited the Academy as a whole.

The mainstreaming process (working collaboratively over successive tries to discover which kinds of information can impact training, collecting data, and generating and reporting that information in the most usable format) is labor-intensive. This means that progress must necessarily come in small steps, with desirable steps put off in order to meet a more immediate priority. Nevertheless, some of what has been achieved for the CPS training, primarily because the test is high stakes, has generalized to training for other groups. Nevertheless, much remains to be done before mainstreaming is fully institutionalized for all of the Academy's training.

References Cited

- Anderson, D. (1991). The Minnesota quality initiative. *Public Productivity Management Review*, 15:2 (Winter), 175-178.
- Anderson, J.R., & Johnson, N.B. (2005). On the relationship between work contexts. Mandates and compliance behaviors of supervisors. *Journal of Change Management*, 5:4 (December), 381-390.
- Argyris, C. (1955). Organizational leadership and participative management." *The Journal of Business*, 28:1 (January), 1-7.
- _____ (1957). The individual and organization: Some problems of mutual adjustment. *Administrative Science Quarterly*, 2:1 (June), 1-24.
- Bell, J.B. (2004). In J.S. Wholey, H.P. Hatry, & K.E. Newcomer (Eds.), *Handbook of practical program evaluation (2nd ed.)*, (pp. 571-603). San Francisco: Jossey Bass.
- Chatman, J.A. (1989). Improving the interactional organizational research: A model of person-organization fit. *Academy of Management Review*, 14:3, 333-349.
- Cross, R., Baker, W., & Parker, A. (2003). What creates energy in organizations. *MIT Sloan Management Review*, 44:4 (Summer), 51-56.
- Eseryel, D. (2002). Approaches to evaluation of training: Theory & practice. *Educational Technology & Society*, 5(2), 93-98.
- Gilmore, T.N., & Krantz, J. (1991). Innovation in the Public Sector: Dilemmas in the abuse of ad hoc processes. *Journal of Policy Analysis and Management*, 10:3 (Summer), 455-468.
- Hatry, H.P., Wholey, J.S., & Newcomer, K.E. (2004). "Other issues and trends in evaluation." In J.S. Wholey, H.P. Hatry, and K.E. Newcomer (Eds.), *Handbook of practical program evaluation (2nd ed.)*, (pp. 670-684). San Francisco: Jossey Bass.
- Kohn, M.L., & Schooler, C. (1978). Reciprocal effects of substantive complexity of work and intellectual flexibility: A longitudinal assessment. *American Journal of Sociology*, 84:1 (July) 240-52.
- Kopczynski & Pritchard. (2004). "The use of evaluation by nonprofit organizations." In J.S. Wholey, H.P. Hatry, & K.E. Newcomer (Eds.), *Handbook of practical program evaluation (2nd ed.)*. San Francisco: Jossey-Bass.

- Mechanic, D. (2003). "Sources of power of lower participants in complex organizations." In J.S. Ott, S. Parkes, & R.B. Simpson (Eds.). *Classic readings in organizational behavior (3rd ed.)*, (pp. 410-417). Belmont, CA: Wadsworth/Thomson Learning.
- Miner, A.S. (1987). Idiosyncratic jobs in formalized organizations. *Administrative Science Quarterly*, 32:3 (September), 327-351.
- Moorman, R.H., & Blakeley, G.L. (1995). Individualism-collectivism as an individual difference predictor or organizational citizenship behavior. *Journal of Organizational Behavior*, 16:127-42.
- Protecting the children of New York: Reform plan for of action the Administration for Children's Services. (1996, December 19). Report to Mayor Rudolph Giuliani. New York: City of New York, Administration for Children's Services.
- Sanders, J.R. (2003). "Mainstreaming evaluation. New directions for evaluation." 99 (Fall). In J.J. Barnette & J.R. Sanders (Eds.), *The mainstreaming of evaluation*.
- Stake, R.E. (1995). *The Art of Case Study Research*. Thousand Oaks, CA: Sage Publications.
- Torres, R.T., Preskill, H.S., & Piontek, M.E. (1996). *Evaluation Strategies for Communicating and Reporting: Enhancing Learning in Organizations*. Thousand Oaks, CA: Sage Publications.
- Weiman, D. (2004). "Evaluation and Quality Improvement: Convergent—or Complementary?" Paper presented at the Evaluators Exchange of the Ohio Program Evaluators Group (OPEG), The Fawcett Center, Columbus, Ohio. June 4, 2004.

	Discussion: Training Evaluation as Quality Assistance

Theme: How Evaluation Brings Attention to Gaps in Training

Title: Training Evaluation as Quality Assistance: Moving from Theory to Reality.

Presenter: Henry Ilian, D.S.W.

Facilitator: Midge Delavan, Ph.D.

Dr. Ilian's presentation focus on the interpersonal dimension in ensuring that training evaluation is used. His presentation emphasized the role of partners among managers who are potential users the results of the evaluation. He argued for the need to seek out and establish an active partnership with those who are interested in exploring what the evaluation has to offer and can see a use for it. He maintained that evaluation results get used only when the evaluator can find managers who are interested in using them, which is not always the case, and even when it is, it is still necessary to develop relationships and to work with those managers as partners in order to refine the evaluation and the ways it is reported in order to make it maximally useful.

As an illustration, he discussed the possibilities for using test results as an instrument to measure training effectiveness. He discussed current progress that New York's Administration for Children's Services has made as well as challenges encountered.

Topics of Discussion

III. How to Promote Evaluation Use

- Promoting evaluation use requires evaluators to develop positive relationships with training administrators in order to form a partnership. Without these relationships, members of the training team are unlikely view the evaluation as a resource for quality control or for improving the training.

- Evaluators' relationships with training directors allow them to demonstrate the usefulness of evaluation. Only by understanding how their programs benefit from evaluation will training directors incorporate its use into the organizational process.
- After the partnership is established, evaluators and training administrators must work together to increase the evaluation's usefulness. For example, refining the type and format of reports enhance the usefulness of the evaluation.
- Privatized child welfare trainings in Florida demonstrate a partnership with evaluators. High trainee failure rates promoted dialogue between private providers and evaluators despite initial hesitance.
- Proving the connection between test results and job performance would support evaluators' promotion of partnerships.

I. How Tests Evaluate Training Effectiveness

- Training teams review how content of incorrectly answered test items is being emphasized during class and identify areas that are insufficiently covered. If two training teams have different results, it is clear that trainers are affecting the outcome.
- Trainers also use item analyses from pre- and post- tests to evaluate effectiveness. All trainees are new to the job and most arrive with little or no pre-existing knowledge. Do the multiple-choice tests accurately measure effectiveness? Testing bias is not controlled for. Another concern is that trainees may forget information during the 29-day training.
- Multiple trainers review written items to ensure that they represent the content. An item is put on the test to see how well trainees answer it. The item must fall in a range where most trainees, but not all, answer it correctly for it to be used in the future.

II. Challenges

- Using tests to measure training effectiveness as well as an individual diagnostic tool.
- High-stakes testing has resulted in improved trainee test scores. Previously, trainees did not take the tests seriously because they were not accountable for their test scores. Current trainees study more because their employability depends on passing the test.
- High-stakes testing also creates a need for other forms of evaluation because of increased accountability.
- Trainers may compete with each other or possibly cheat so that a higher percentage of their trainees pass tests. In response, there have been no cases of cheating. It is more appropriate to determine how much information trainees must leave with and identify curriculum biases that fail to adequately cover certain items. Through dialogue between trainers and evaluators, trainers become more aware of needing to cover all material that is on the test. The overall goal is to make sure that the class understands the content.

IV. Ethical Concerns

- Do child welfare providers have an ethical responsibility to provide competent workers to serve a vulnerable population?
- Is it ethical to not hire a trainee because he/she failed a test? In reality, only one to two trainees out of 25 fail the tests. Supervisor and trainer evaluations have also been used in trainee evaluations.
- Is it ethical not to hire trainees because they had an incompetent trainer?
- Standardization represents a safeguard because it holds trainers accountable to present specific information. However, increased standardization reduces the impact of trainers' personality driven characteristics.

	Comparing On-Line with Classroom Training

Cindy Parry, *Ph.D., Training and Evaluation Consultant*, and
Susan Brooks, *M.S.W., Director of Child Welfare Programs*,
Northern California Child Welfare Training Academy

Abstract

This paper discusses several issues associated with evaluating the effects of training modality for child welfare training in which participants self-selected into either classroom-based or on-line versions of the same class. The Northern California Child Welfare Training Academy's training on the Multi-Ethnic Placement Act was evaluated to test the effect on knowledge acquisition of delivery modality, on-line versus classroom. The evaluation process also illustrated several issues relating to quasi-experimental comparisons and differential item functioning. Results of the evaluation indicated that delivery modality did not make a difference in student learning, at least for classes of this type that address mainly knowledge and awareness level competencies.

Evaluation Background and Goals

The evaluation of the Northern California Child Welfare Training Academy's training on the Multi-Ethnic Placement Act (MEPA) had two main goals:

1. to determine the extent to which trainees gained knowledge and understanding of the requirements and application of MEPA legislation, regardless of training modality, and
2. to assess the effects of training modality on knowledge gained.

The training is administered in two modalities: on-line and traditional classroom training. The on-line course is asynchronous, meaning trainees complete the material at different times.

However, trainees do take the class within a specified time frame, and an instructor responds to questions and moderates class discussion through a bulletin board function. Both modalities use the same written curriculum, and within the Northern Academy, are taught by the same instructor. This evaluation also was conducted in cooperation with the Central California Child Welfare Training Academy (CCTA) and the Public Child Welfare Training Academy, Southern Region (PCWTA). Both of these academies conduct this training in the traditional classroom format, again using the same written curriculum.

Design and Instrumentation

The MEPA training was evaluated using a pre-test post-test quasi-experimental design. Amount of change from pre- to post-test for on-line participants was compared to the amount of change for traditional classroom training participants in a multiple regression framework, which also included several key variables on which the two groups differed, in addition to delivery modality. Initially, the study was planned to include only participants from the Northern Academy trainings. However, it became clear early on that most participants were electing to do the course on-line. Thus, the Central and Southern Academies were contacted and agreed to provide additional data from classroom deliveries of the same curriculum.

A 23-item test was developed based on the learning objectives, competencies, and content of the written curriculum. The test incorporated some existing True/False items from an earlier version developed by the Northern Academy, and a number of new multiple-choice items. Items were designed to test factual knowledge as well as understanding of the application of the law to child welfare placement decisions. The same test was used pre- and post-training.

Each academy assigned one trainer to conduct all classes in their region. In the Northern Academy, the same trainer taught both classroom and on-line versions of the course. Trainees were not allowed to refer back to notes or handouts when taking the classroom versions of the test. On-line participants did not have access to on-line course materials while completing the tests. However, it was possible that participants could have printed materials from the screens prior to taking the post-tests.

A form was also developed which collected background information from trainees regarding:

- Reasons for taking the course,
- Whether or not they had specific children or families in mind with whom to use the course information,
- A self-assessment of the extent to which they gained knowledge and understanding of MEPA provisions and applications to child welfare,
- Experience in child welfare,
- Experience in human services,
- Primary job responsibility,
- Age,
- Gender,
- Race/ethnicity,
- Educational level, and
- Degree field.

This information was collected once, post-training.

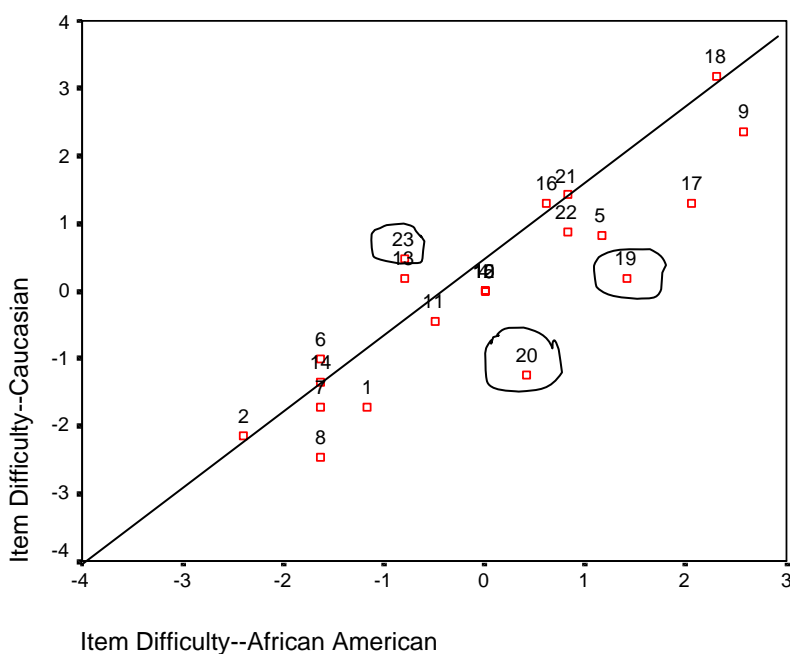
Item Functioning

Northern Academy staff familiar with the curriculum reviewed test items for clarity and appropriateness to learning objectives and content. Following data collection, items were analyzed statistically using a Rasch model framework. Any that did not display appropriate item discrimination indices¹ were not included when calculating test scores for use in the analyses. Of the 23 items originally included on the tests, 18 met item analysis criteria. An additional three items were excluded based on differential item functioning. Item difficulties were calculated separately for Caucasian, Hispanic/Latino, and African American trainees, and Pearson correlation coefficients were calculated. Overall correlations were high (.90 for African American with Caucasian and .94 for Hispanic/Latino), indicating that in general items that were difficult for one racial/ethnic group were also difficult for the other and vice versa. Difficulty estimates for African American

¹ In testing, the term discrimination refers to how a person's score on an item relates to his or her score on the test as a whole. Items should have positive discrimination indices, which means that someone who scores higher on the test has a greater probability of getting the item correct than someone who scores lower on the total test and vice versa.

and Hispanic/Latino trainees were also plotted against difficulty values for Caucasian trainees and outliers were identified for further scrutiny (see Figure 1 below).

Figure 1. Item Difficulty Scores for Caucasian Versus African American Trainees



Caucasian and Hispanic/Latino trainees were slightly more likely to answer correctly than African American trainees on two of the three items (19 and 20), while African American trainees were more likely to answer the third (item 23) correctly. Thus, 15 items made up the final set on which scores were based.

Items identified as potentially problematic were as follows (correct answers are underlined):

19. Which of the following state policies is least likely to conflict with the requirements of MEPA? Policies that promote
 - a. Racial or ethnic matching, as long as it does not exclude any child or prospective parent from consideration on the basis of race, ethnicity or national origin.
 - b. Racial or ethnic matching, as long as it does not deny or delay placement.

- c. Placement with adult relatives over unrelated adults, as long as they meet relevant child protection standards.
- d. Racial or ethnic matching when it is requested by parents who relinquish a child for adoption.

On this item 41.7 % of African Americans picked the correct answer, “c”, as compared to 67.5% of Hispanic/Latino trainees and 78.6% of Caucasian trainees. African Americans were more likely to choose option “a” or option “b” than either of the other two groups.

20. Tanya is a 12-year-old African American child who has been placed with a white foster family for approximately 15 months. The family moved into a predominantly white suburb about five months ago. Tanya continues to visit her biological mother and longs to return to her but she has also become close to her foster parents. Tanya is being teased and subjected to racial slurs by her white classmates and has few friends at school. She has frequently come home upset and in tears and has begun to have an upset stomach in the morning before leaving for the bus. The best option that complies with MEPA provisions would be to

- a. Place Tanya with an African American family living in the same school district as her mother and who can take her in six weeks time.
- b. Place Tanya with a Hispanic family living in a more diverse school district and who can take her immediately.
- c. Place Tanya with a white family living in a more diverse school district and who have had several foster children of other races, and who can take her immediately.
- d. Work with her current foster family to assess their capacity to meet Tanya’s need for permanency and to support Tanya in making friends and dealing with teasing and slurs from peers.

On this item, 62% of African Americans picked “d” compared to 88.3% of Hispanic/Latino trainees and 92.1% of Caucasian trainees. African Americans who didn’t pick “d” mainly picked either “a” or “b” (about evenly split).

23. Nathaniel is a 15-day-old Caucasian infant. He was born drug-addicted with many complications. He is due to be released from the hospital within the next few days. His mother left the hospital following his birth and has not returned to visit. She

did not come to the detention hearing. Her current whereabouts are unknown. The presumed father is deceased due to a drug overdose. Doctors say Nathaniel will probably need substantial ongoing medical and developmental services when he is released from the hospital. The best placement option for Nathaniel is

- a. A Caucasian foster family who have completed training in caring for special needs infants. They specialize in emergency placements and are not interested in adoption. They will take Nathaniel immediately.
- b. A lesbian couple, one of whom is an RN specializing in neonatal care. She is African American and her partner is a Cuban American preschool teacher. They are an approved adoptive family who are also licensed to do foster care. They can take him immediately.
- c. A Caucasian woman who is the godmother of the biological mother. She is willing to complete specialized training and meet with medical personnel regarding Nathaniel's care. She would like to facilitate visitation with the child's mother, if appropriate. She would be able to care for Nathaniel within three weeks.

On this item, the majority of all racial groups picked the correct answer, "b." A slightly larger percentage of African Americans chose the correct answer (83.3%) compared to Hispanic/Latino (70.1%) and Caucasian (73.3%) trainees. Hispanic/Latino trainees who did not pick "b" were evenly split between options "a" and "c." Caucasian trainees who did not pick "b" were more likely to pick "c" than "a."

Test reliabilities were also calculated as part of the item analysis and scaling process. Reliability at pre-test was .45 and at post-test was .59². These reliability estimates are lower than what is desirable, likely because the final version of the test was relatively short and somewhat easy for most participants.

Participants

The Northern, Central, and Southern Training Academies participated in this evaluation; however, complete sets of pre- and post-test data were available from only the Northern and Central

² Test reliabilities range from 0 to .99. Reliabilities of .7 to .8 are acceptable, .8 to .9 are good, and .9 and above are excellent.

Academies. A total of 313 sets of pre- and post-tests were available for analysis: 259 from classroom administrations and 54 from on-line administrations. All on-line data were collected in the Northern Academy-sponsored courses. The background information form was introduced after initial pilot testing of the data collection process. Thus, background information was available for approximately 250 trainees.³ This information is summarized below.

Some statistically significant differences were observed between participant groups from each academy. These were:

- A larger percentage of participants from the Northern Academy had children or families in mind with whom they planned to use what they learned (chi square = 5.44, $p < .05$).
- A larger percentage of participants from the Northern Academy agreed or agreed somewhat that they had a better understanding of the goals and requirements of MEPA following training (chi square = 9.54, $p < .05$).
- Northern Academy participants were more likely to have primary job duties involving case-carrying services, while the Central Academy had larger percentages in investigations, hotline, and non case carrying services (chi square = 11.08, $p < .05$).
- A higher percentage of participants in the Northern Academy classes were male (chi square = 7.45, $p < .01$), although the majority of participants in from both regions were female.
- Participants from the Central Academy were more racially and ethnically diverse (chi square = 19.44, $p < .01$).
- Participants in the Central region were more likely to have degrees in social work or human services, while participants in the Northern region had a substantial percentage with degrees in related fields, such as psychology or mental health (chi square = 11.19, $p < .05$).

No statistically significant differences between groups were noted in

- Experience in either human services or child welfare,
- Age,
- Educational level,

³ Numbers varied somewhat by question due to missing data.

- Reasons for taking the class, or
- Post training self ratings of understanding of MEPA's application to child welfare.

When background characteristics of those taking the course on-line were compared with those taking it in the classroom, statistically significant differences emerged in the following areas:

On-line participants were more likely to

- have less than 2 years of child welfare experience (chi square = 8.33, $p < .01$);
- have case-carrying services as their primary job duty (chi square = 8.19, $p < .01$);
- take the course for reasons other than being required to (chi square = 13.33, $p < .05$), and
- have a smaller percentage of college degrees overall, but a larger percentage of Master's or doctoral-level degrees (chi square = 13.41, $p < .05$).

On-line and traditional classroom participants did not differ in

- Experience in human services in general,
- Age,
- Race/ethnicity,
- Gender,
- Degree field,
- Post training ratings of knowledge of MEPA's goals and provisions or understanding of MEPA's application to child welfare, or
- Degree to which they had children or families in mind with whom they planned to use the training,

Results of the Evaluation

Data were analyzed using the Rasch model framework, which offers several useful properties for analysis⁴. Thus the mean pre- and post-test scores are reported here in logits⁵.

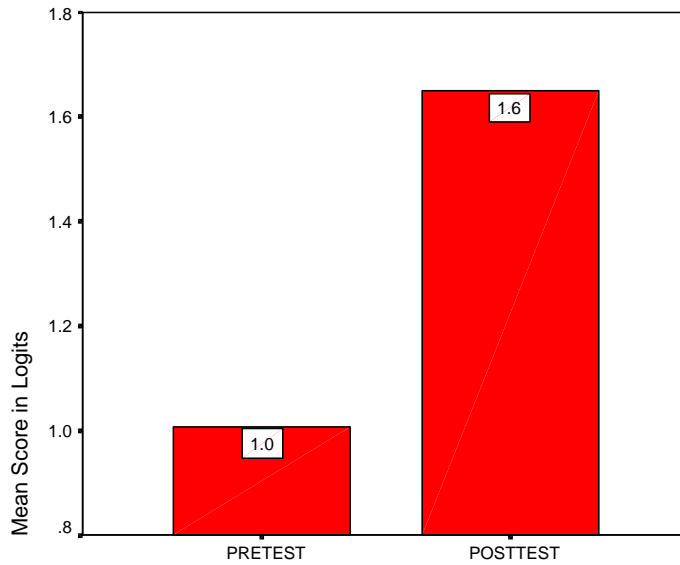
⁴ Rasch modeling is one of a number of item response theory measurement models. It offers what is referred to as item free and sample free measurement, which allows for easier interpretation of data across several samples and over time, and variations in items responded to (e.g., due to missing data or alternate test forms).

⁵ The Rasch model measures performance in logit scores. A person's logit score is the natural logarithm of their odds of success on items of a given difficulty level, in this case, items at the 0 point on the scale. Logits form an equal interval scale that is desirable for making statistical comparisons.

Goal 1: Assessment of participants gains in knowledge and understanding of MEPA

Overall, participants from both Academies made statistically significant gains in knowledge and understanding of MEPA from pre-test to post-test, as shown in Figure 2 below ($t = 9.100$, $p < .001$).

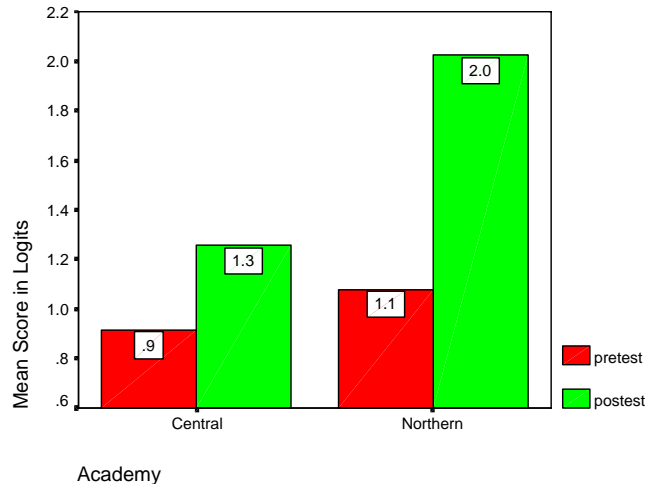
Figure 2. Overall Pre-test to Post-test Change



As shown in the graph, the average pre-test score was approximately 1 logit, and the post-test score was approximately 1.6 logits.

When considered separately, both the Central and Northern Academy participants showed statistically significant gains ($t = 4.90$, $p < .001$, and $t = 6.91$, $p < .001$, respectively). *These are shown graphically in Figure 3, below.*

Figure 3. Overall Pre-test to Post-test Change by Academy



Northern Academy includes extension course held in Ventura County

As shown in the graph, while both sets of participants experienced significant gains, participants trained by the Northern Academy experienced larger gains on average. This difference was also statistically significant ($t = 3.93, p < .001$).

Since all on-line participants were from the Northern region, and since there were some statistically significant differences between participants from the Academies on background variables, further multiple regression analyses were conducted to assess the relative contributions of these differences to the total picture. These are discussed under goal 2 below.

Goal 2: Assessment of the effects of training modality on gains in knowledge and understanding

Multiple regression analyses were conducted to estimate the relative contribution of modality while controlling for differences among participants related to demographics and the academy through which they received their training. Post-test scores were predicted from pre-test scores, modality, and all background variables, with the exception of time in human services, which was very closely related to time in child welfare, and the two variables measuring self-assessments of learning.

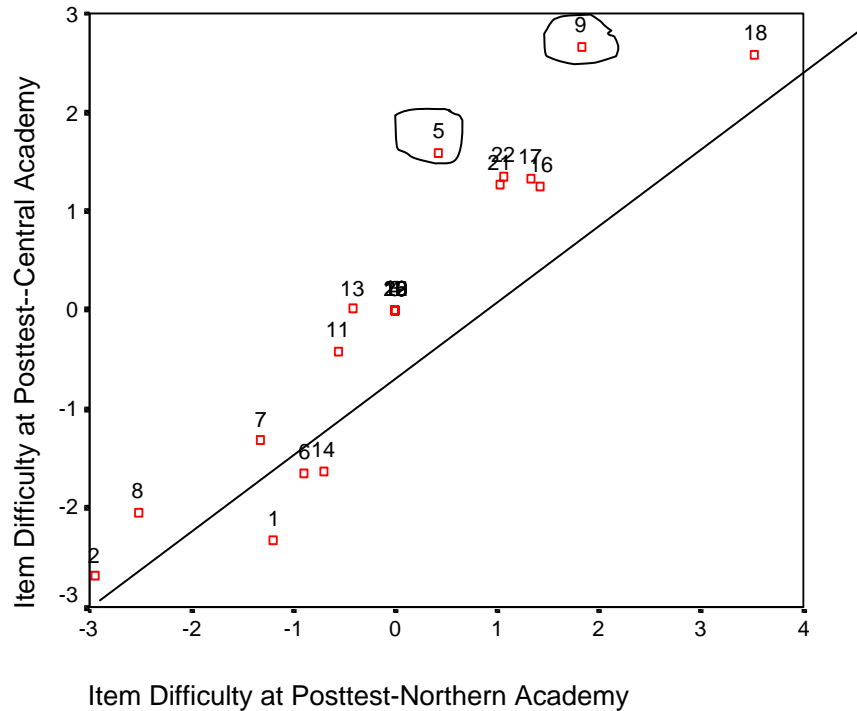
This analysis revealed that modality did not have a statistically significant effect on post-test scores, when the effects of pre-test score and other group differences were controlled for. Thus, for this course, both on-line and classroom delivery methods were equally effective.

Pre-test score, as might be expected, significantly predicted post-test scores ($t = 2.13$, $p < .05$). In other words, those who had higher scores on the pre-test were more likely to have higher scores on the post-test and vice versa. The academy delivering the training also significantly predicted posttest scores ($t = 4.32$, $p < .001$), even when other differences among participants were controlled for, suggesting that there may have been differences in the way in which the two trainers approached the curriculum. No other demographic or background variables were significantly related to posttest scores.

In an effort to better understand the role of possible trainer/course delivery differences in the differences between academies, separate analyses were run for each, and the item difficulty values at post-test were compared for each academy. If training is being delivered with somewhat different emphases, then individual questions might be easier or more difficult depending on who delivered the training. *This comparison is shown in Figure 4.*

With the exception of a few items, most item difficulties are comparable for the two groups, as indicated by a pattern clustered around the diagonal line on the graph and a high degree of correlation ($r = .93$). However, items 5 and 9 fall farther from the central line, indicating that they differed in difficulty depending on the academy offering the training. Both of these items were significantly more difficult for the Central Academy sample, or in other words, were answered incorrectly more frequently than expected.

Figure 4. Item Difficulty by Academy



These items are:

5. An agency may use race to differentiate between otherwise acceptable placement options if doing so won't delay or deny the placement.

- a. True
- b. False⁶

and

9. An agency can assess the racial, ethnic, and/or cultural capacity of foster parents. The agency can do this by assessing the capacity itself, or as part of another assessment.

- a. True
- b. False

While additional discussion with the trainers and observation of the classes would be needed to definitively answer any questions related to differences in delivery of the material, these

⁶ Correct answers are underlined.

items suggest a possible starting point; both address when race may be considered in placement decisions.

Other Findings

Two items on the background information form dealt with self-assessments of learning: “I gained a thorough understanding of the basic goals and requirements of MEPA in this course” and “I understand the application of MEPA to child welfare practice.” Ratings of learning on these two items were recoded to “agree” versus “disagree or neutral” responses, and correlated with the amount of change from pre- to post-test as reflected in the test scores. Self-ratings on the first item “I gained a thorough understanding of the basic goals and requirements of MEPA in this course” were not significantly related to learning as assessed by pre- to post-test gains. Self-ratings on the second item “I understand the application of MEPA to child welfare practice” showed a small, but statistically significant relationship to pre- to post-test change ($r = .13, p < .05$). Participants who rated their understanding more highly on this item also tended to show a higher degree of learning when measured by pre- and post-tests.

Discussion and Recommendations

This study had several limitations: lack of random assignment to delivery modality, the need to involve other training academies and instructors in order to boost numbers of participants in the classroom condition, and the use of the same, relatively brief, test for both pre- and post-training assessment. However, these limitations serve to help illustrate the complexities associated with evaluations of training delivery modality, as well as some of the issues associated with valid measurement of knowledge acquisition regardless of delivery modality. They also suggest directions for future research.

A complicating factor in this, and other studies, is the lack of random assignment to training modality. As illustrated here, participants differ in several potentially important ways, in addition to the way their training is delivered. In this study an initial comparison of pre- to post-test change (without consideration of academy or other background variables) indicated an advantage for the on-line delivery over the classroom delivery modality. However, this difference appeared to be explained by the

academy delivering the training, rather than modality, when additional variables were included in the analysis. In quasi-experimental designs using comparison groups, at a minimum, careful attention needs to be given to the effects of group differences other than the variable of interest. These studies are still limited, however, by the possibility that additional important differences have not been identified and included in the analysis. Thus, future research is needed where trainees are randomly assigned to a delivery modality.

While participants from different academies differed in several demographic and other background variables, these variables did not appear to explain the differences found with respect to academy. The same written curriculum was used for all academies. The fact that each academy employed a different trainer, and that there were differences in the relative difficulty of two of the items based on academy, indicates that there may have been differences in relative emphasis on, or interpretation of, some of the information. This suggests a continued need, as more and more courses are being evaluated at knowledge and skill levels, to minimize variation in content coverage, at least for key points. While it is neither realistic nor desirable to reduce trainers to automatons, meaningful interpretation of evaluation results over time, regionally or statewide, demands a level playing field where all trainees are exposed to key content in the same way.

Analyses of differential item functioning in this study also identified an interaction between the race of the trainee answering the question and the interpretation of some of the scenario-based test items. It appears that in an area such as this, trainees' backgrounds and previous experiences may play a larger part than initially expected in how they interpret what the law requires and what the best placement option is for a given child. Analyses of differential item functioning are not done routinely in evaluations of child welfare training, but the results obtained in this study suggest that they can be important in uncovering cultural differences in how trainees absorb and act on what they are taught. Differential functioning analyses are generally used to make tests a more fair assessment of knowledge or ability. However, this type of feedback may allow for not just improvement of tests themselves, but also of the training curricula and delivery, by

suggesting areas where more discussion is needed to convey important concepts.

Finally, in this study, the same test was used for both pre- and post- testing, leaving open the possibility that trainees benefited on the post-test from having seen the items before on the pre-test. This is common practice in child welfare evaluation, in part due to time and budget constraints and in part to the technical complexity of developing alternate test forms equivalent in difficulty and adequacy of content coverage. Item response theory models, like the Rasch model, offer a promising way to address the need for alternate forms of a test by scaling all items together and enabling subsets of the items to be interpreted within the same framework. If warranted by future enrollments, new items will be written for the MEPA course and scaled with existing items to create a pool from which different test versions can be developed. It is often possible to create two forms that have some overlapping content as well as several different items, thus minimizing the potential for inflated post-test scores without creation of a prohibitively large pool of potential items.

Conclusion

Training evaluation has come a long way in recent years. Earlier sample evaluation designs to measure trainee satisfaction have transformed into complex evaluations of knowledge and skill acquisition, and attempts (like this one) to examine alternate methods of training delivery. As training evaluation tackles increasingly sophisticated questions, evaluators face sometimes daunting technical challenges. It is important, however, to continue to identify these challenges and utilize the feedback from our evaluations to both improve our designs and to shed light on training content and instructional issues. By continuing to strive for more rigors in evaluation design and instrument development, we move forward both the evaluation of child welfare training and the knowledge of what makes training effective.

	Discussion: Comparing On-Line with Classroom Training

Theme: Factors that Influence On-Line Learners

Title: Comparing Trainee Learning in On-Line and Classroom Deliveries of MEPA Training

Presenters: Cindy Parry, *Ph.D.*, and Susan Brooks, *M.S.W.*

Discussant: Kurt Kraiger, *Ph.D.*

Facilitator: Todd Franke, *Ph.D.*

Topics of Discussion

Dr. Parry and Ms. Brooks presented their findings from developing an on-line training curriculum. The curriculum has been implemented to train child welfare workers in Northern California. Dr. Kraiger provided information about learning factors that impact the effectiveness of on-line trainings when used by adults. The discussion covered how trainees' affective characteristics influence their responses to test items and directions for future research to increase the effectiveness of on-line training.

I. On-Line Learning

- According to multi-media education research, animation and pop-ups are intended to maintain the student's interest. However, they result in reduced learning because they compete for attention with relevant information.
- Age and computer knowledge are two characteristics that impact the effectiveness of on-line trainings. Age impacts learning capacity. A 12-year-old student learns differently than a 25-year-old adult. Students who view computers as a resource will have more success than those who view a need to master a computer in order to complete a task.

II. Affective Characteristics: Why Do People Answer Questions in Certain Ways?

- Differential Item Functioning Analysis would benefit from further understanding the cognitive processes of why people choose certain answers which were incorrect.
- The group discussed differences in responses across groups, and the influence that group values may have on responses. Students may know the right answer according to the curriculum but answer according to their values that do not coincide with curriculum. According to decision-making research, individual values were greater decision-making predictors than knowledge of laws and policies. Thus, it is hard to teach someone to do something that conflicts with their values. This was particularly evident for these trainings that dealt with the Multi-Ethnic Placement Act (MEPA).
- Creating scenarios that promote critical thinking represents one solution to disentangling values from knowledge of the curriculum content. Scenarios should evaluate for students' curriculum knowledge and their values, and the course of action they would choose. The opportunity to express value-based disagreements may encourage students to answer curriculum questions correctly which they may have incorrectly answered because of conflicting values. On-line trainings can allow the opportunity for students to express their value-based disagreements.
- Student motivation also influences learning. Workers mandated to attend trainings demonstrated low motivation and resented being away from their jobs for an extended period of time. In this study, on-line students demonstrated higher levels of motivation, perhaps because of their personal motivation to take the course.
- Research designs could measure how values affect learning. Students could be divided into two groups based on their pre-training attitudes (motivated vs. not motivated/resentful). Based on a post-test, possible results could indicate how motivated students benefit from the training, thus a positive statement about training effectiveness.
- More research needs to be reviewed to meet the needs of current workers who are not motivated to learn. This is a

concern because they are currently working in the field but not learning new information.

- Tests can be used as a teaching tool. Scenarios can require students to address clients' situations in accordance with the intentions of specific laws. Then the student does not have the opportunity to answer in the context of their values. These scenarios may challenge students to think critically about why their values conflict with certain policies.

III. Other Item Difficulty Factors

- High-stakes testing would force people to answer correctly despite any value disagreements. Currently, no accountability is connected to test scores.
- Students may respond differently if testing was anonymous.

	Using the Logic Model to Conceptualize and Enhance Training Evaluation Design

Alan J. Dettlaff, *Ph.D., LMSW*, and Joan R. Rycraft, *Ph.D., ACSW*

Abstract

This paper addresses the application of logic models to training evaluation by drawing on the experience of developing a logic model for the federally funded project “Culturally Competent Systems of Care Practices with Hispanic Children and Families.” The paper reviews how the logic model was used as a tool to conceptualize both the goals of the project and the evaluation design. Throughout this project, the Principal Investigator and the Project Evaluator worked together in developing and revising the logic model to ensure that all aspects of the evaluation design were addressed in the conceptualization of the project. Issues that were addressed in these revisions included: additional operationalization of proposed outcomes, feasibility of the evaluation design; and enhanced linkages between goals, short-term outcomes, long-term outcomes, and the methods by which each of these would be evaluated. Each draft of the logic model is presented along with a discussion of the challenges and resolutions that each revision addressed in order to demonstrate the process of developing and refining the logic model to ensure consistency between outcomes and evaluation.

Introduction

Logic models are becoming an increasingly important planning and evaluation tool in health and human services programs in both the public and nonprofit sectors. As funding agencies continue to emphasize specificity and results, RFPs are requiring outcome-based evaluation, best practices programs, and sustainable organizations. Trying to describe these elements in a proposal narrative can be complex and sometimes boring—the death of fundable proposals. Too much detail can cause the

program to become lost in the verbiage. Sometimes it would be easier to just draw a picture.

For many grant writers the evaluation section at the end of a proposal is a daunting challenge. The components of writing the proposal such as developing goals and objectives, describing programs, and defining community need are often approached separately from the evaluation. Evaluators, however, have developed tools that can be very useful in planning grant projects and presenting goals and objectives. One such tool that is gaining wider usage is the logic model. Used correctly, a logic model can help in writing the body of the proposal, improve services, and provide the blueprint for a manageable and efficacious evaluation. The logic model has been used for more than 20 years by evaluators to describe logical linkages among program resources, activities, outputs, audiences, and short-term, intermediate, and long-term outcomes related to a specific problem or situation.

Purpose of the Logic Model

Evaluators use the logic model as a blueprint for the evaluation. A logic model can develop understanding about the program, what it is designed to do, and the measures it will use to determine success. A logic model can also monitor progress. It can serve as an evaluation framework to identify appropriate evaluation questions and relevant data.

The purpose of the logic model is to identify the short-term, intermediate, and long-term outcomes for the program and link them to each other and to program activities and inputs. A logic model provides stakeholders with a road map describing the sequence of related events connecting the need for the planned program with the program's desired results. Logic models help to identify what will be done and the results (outcomes) to be obtained. Developing appropriate and measurable indicators is key to a sound evaluation. The logic model helps to identify the practices of a program that will provide useful evaluation data, and to map out an appropriate sequence for collecting data and measuring progress.

Definition

A logic model is a management tool that can be used throughout the life of a program. It is a systematic and visual way to present and share your understanding of the relationships among

the resources of your program, the activities you plan, and the changes or results you hope to achieve. The most basic logic model is a picture of how you believe your program will work. It uses words and/or pictures to describe the sequence of activities thought to bring about change and how these activities are linked to the results the program is expected to achieve. Logic models are typically diagrams, flow sheets, or some other type of visual schematic that conveys relationships between contextual factors and programmatic inputs, processes, and outcomes. The value of the logic model is that it visually expresses beliefs about why the program is likely to succeed.

Developing the Logic Model

The key to building any model is to prepare a working draft that can be refined as the program develops. Most of a logic model's value lies in the process of creating, validating, and then modifying the model. An effective logic model will be refined and changed many times throughout the evaluation process as more information is collected about the program and how it works. Because logic models are pictorial in nature, they require systematic thinking and planning to better describe programs. In a logic model, one can adjust approaches and change courses as program plans are developed. The layout of the boxes and arrows represent how a program is assumed to function. The logic model is an iterative tool, so it can change over time as the program develops and a more thorough understanding of the program is gained. The remainder of this paper will describe the authors' process of developing a logic model for a federally funded training grant, and how this process facilitated the development of the evaluation plan.

Culturally Competent Systems of Care Practices with Hispanic Children and Families

Culturally Competent Systems of Care Practices with Hispanic Children and Families is a three-year training grant funded by the U.S. Department of Health and Human Services Administration of Children and Families' Children's Bureau. The project is a collaboration between the University of Texas at Arlington (UTA) School of Social Work, Texas Christian University Department of Social Work, and the Texas Department of Family and Protective

Services. The goals of this project are to: (1) develop a curriculum to improve the knowledge and skills of Texas child welfare supervisors and caseworkers in Hispanic culture, cultural competence, and the development and utilization of systems of care for Hispanic children and families, and (2) to enhance the University of Texas at Arlington social work curricula with content focusing on services to Hispanic children and families. The project builds on the longstanding partnership of the three entities involved and incorporates input from a Community and University Panel of Experts in curriculum development and Hispanic culture.

This training project was developed in response to the changing demographics of the Hispanic population in Texas and the need to prepare child welfare staff to address the needs of this population. Within Texas, 35% of the population is Hispanic (U.S. Census Bureau, 2004) and Hispanic children comprise 42% of confirmed victims of abuse and neglect (TDFPS, 2005). However, as the Hispanic population has grown in Texas, little has changed in child welfare practice to respond to this change in demographics. Historically, services to children and families have been selected from a limited set of choices that have not considered families' culture, values, or experiences. This training curriculum is built upon the premise that in order to improve outcomes of safety, permanence, and well-being, interventions must respond to individual and cultural differences of clients. The training incorporates a Systems of Care framework as a practice model for effectively meeting the needs of Hispanic children and families.

The Logic Model

The first draft of the logic model for this project (*Figure 1*) includes 8 headings: 1) Principal Participants, 2) Approach, 3) Inputs, 4) Activities, 5) Short-term Outcomes, 6) Evaluation, 7) Activities, and 8) Long-term Outcomes. The basic elements of a logic model are included in this draft, as the model describes who will be involved (participants, inputs), what will occur (activities), and the expected outcomes (short-term and long-term). The short-term outcomes in this model match the stated goals of the project and the long-term outcomes state the desired result of the short-term outcomes.

This draft of the logic model was used to develop the foundation for the evaluation design, as it was clear that each of

the four outcomes (short and long) would need to be evaluated. However, in order to strengthen the evaluation design, the Principal Investigator and Project Evaluator worked together to revise the logic model to ensure that the goals and outcomes of the project were addressed in the evaluation design and that this design would adequately measure the proposed outcomes. The primary issues addressed in the revisions of the logic model included further operationalization of the proposed outcomes and improved linkages between goals, short-term outcomes, and long-term outcomes.

First-Round Revisions

As a framework for the first round of revisions, the goals and outcome objectives of the project were reviewed. These goals and outcome objectives, as stated in the grant proposal, include:

- *Goal 1:* To develop a curriculum to improve the knowledge and skills of Texas child welfare supervisors and caseworkers in Hispanic culture, cultural competence, and the development and utilization of Systems of Care.
- *Outcome Objective 1:* To increase the knowledge and skill of 30 supervisors and 120 caseworkers in Hispanic culture and the development and utilization of culturally competent Systems of Care in the Hispanic community.
- *Goal 2:* To enhance the UTA School of Social Work curricula to improve the knowledge and skill of social work students regarding Hispanic culture and the development and utilization of Systems of Care for children and families in the child welfare system.
- *Outcome Objective 2:* To increase the knowledge and skills of UTA School of Social Work students in Hispanic culture, and the development and utilization of Systems of Care for children and families engaged with the child welfare system.

Following the review of goals and outcome objectives, the logic model was reviewed to ensure consistency with these goals and objectives. Linkages and directional flow within the logic model were also reviewed to ensure accuracy of the proposed logic. This process resulted in the following findings:

- The directional flow within Activities needed to be revised to more accurately show the flow of activities from curriculum development to pilot testing and training delivery.

- The linkages between Activities and Short-Term Outcomes needed to be revised to demonstrate how the activities lead to the outcomes, rather than indicating that all activities lead to all outcomes.
- The linkages between Short-Term Outcomes and Long-Term Outcomes needed to be revised to better demonstrate how the short-term outcomes lead to the long-term outcomes.

As a result, the logic model was revised to address these issues. These revisions are reflected in Figure 2 and described below:

- Principal Participants and Inputs were combined into one category labeled “Inputs” to reduce the overlap between these two columns. The column titled “Approach” was eliminated as the partnership between the inputs is clear from the directional arrows leading to the activities.
- Directional arrows between activities were revised to accurately reflect the flow of activity. Knowledge development was added as an activity to reflect the process of gathering the necessary information prior to developing the curriculum.
- Additional activities were added to more effectively address the second goal of enhancing the UTA School of Social Work curricula. The activities of “Curriculum Revision” and “Dissemination” that were included in a separate column in the first draft were incorporated into this column with directional arrows reflecting their order in the process.
- A column entitled “Outputs” was added to demonstrate the outputs of each activity.
- Short-Term and Long-Term Outcomes were expanded to assist with the evaluation design. Rather than four broad outcomes, outcomes were further operationalized and defined.
- Directional flow between short-term outcomes and long-term outcomes was improved to accurately reflect the process of how each short-term outcome would contribute to a long-term outcome.

Second-Round Revisions

Once the evaluation plan was developed, the logic model was reviewed against the evaluation plan to check for consistency. Similarly, goals and objectives of the project were reviewed to

ensure that each goal and outcome objective was clearly addressed in both the logic model and the evaluation plan. This process resulted in the following findings:

- Outcome Objective 2 was included in the evaluation plan, but missing in the logic model.
- The output of Outcome Objective 2 needed to be more clearly operationalized in the logic model in order to facilitate the evaluation of this objective.
- Dissemination of curriculum as an Intervention/Activity should follow the curriculum revision rather than originating from the Inputs.
- The directional flow of the Long-Term Outcomes needed to be improved to accurately reflect the process that would occur. The directional flow indicating that increased cultural competence would lead to utilization of systems of care is inaccurate.

Outcomes of the Revision Process

The final logic model resulting from each revision process is shown in Figure 3. Each input, activity, output, short-term outcome, and long-term outcome is reflected in the logic model with directional arrows reflecting the logic of the overall project. This process of revising the logic model facilitated the development of the overall evaluation plan by clearly demonstrating each component of the evaluation that would be necessary. As boxes in the logic model were further operationalized, components of the evaluation plan were operationalized. As specific outputs were added to the logic model, these outputs were used to refine and provide direction to the evaluation plan. The resulting evaluation plan includes each output, short-term objective, and long-term objective, along with methods for the evaluation of each component.

Conclusion

This paper demonstrates how the development of a logic model can be used to facilitate and enhance training evaluation design. When used correctly, the logic model clearly specifies the activities, outputs, and outcomes associated with a training project. Once activities, outputs, and outcomes are operationalized and defined within the logic model, a plan for the evaluation of each component can be developed. The logic model also facilitates the

process evaluation as the flow of proposed activities is clearly demonstrated in the model.

References

- Texas Department of Family and Protective Services. (2005). *2005 Data Book*. Austin, TX: Author.
- United States Census Bureau. (2004). 2004 American Community Survey. Retrieved February 26, 2006, from <http://factfinder.census.gov>

FIGURE 1: LOGIC MODEL – DRAFT 1

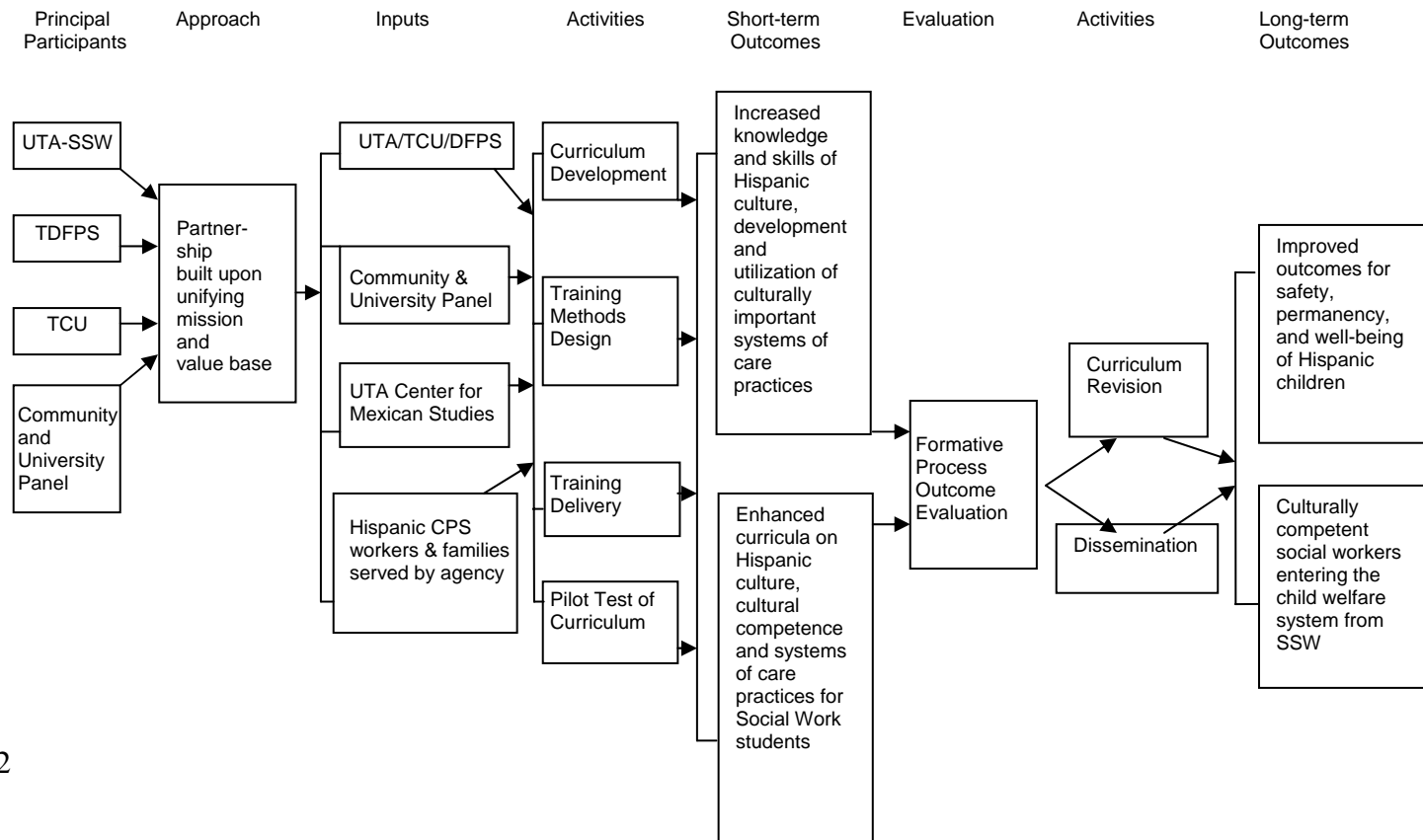


FIGURE 2: LOGIC MODEL – DRAFT 2

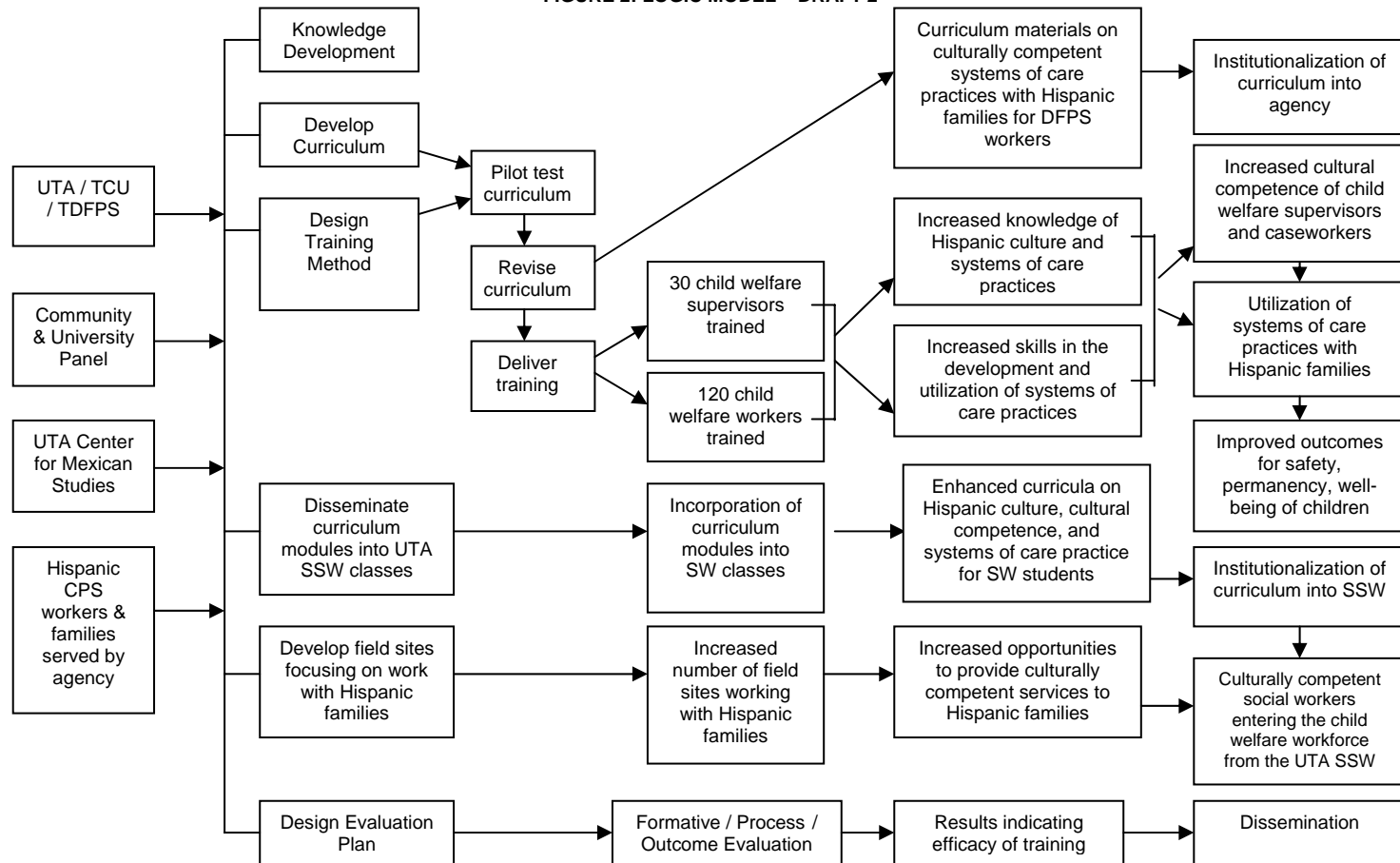
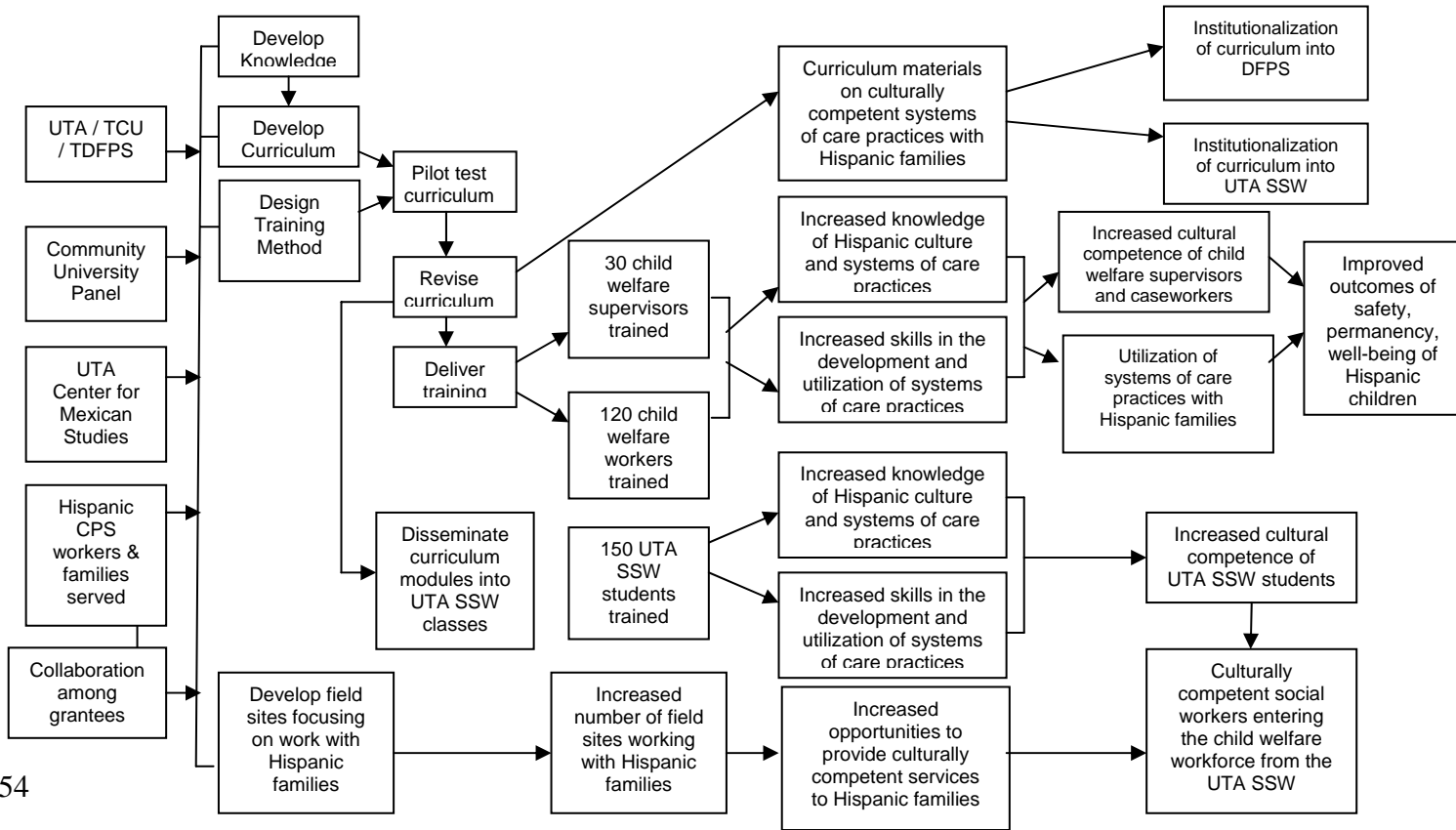


FIGURE 3: LOGIC MODEL – FINAL DRAFT



	Discussion: Incorporating Logic Models in Evaluation Design

Theme: Using a Logic Model to Conceptualize and Enhance Evaluation of a Training Program to Meet Needs of Hispanic Families

Title: Using the Logic Model to Conceptualize and Enhance Training Evaluation Design

Presenters: Allan Dettlaff, *Ph.D.*, and Joan R. Rycraft, *Ph.D.*

Facilitator: Norma Harris, *Ph.D.*

Drs. Dettlaff and Rycraft explained how they used a logic model to design a training evaluation program to develop child protection workers' cultural competency in working with Hispanic families in Texas. They discussed how they've benefited from using the logic model, its potential for future use and results from the pilot programs.

Topics of Discussion

I. Implementing a Logic Model

- Originally, the logic model was completed to satisfy requirements to receive a grant from the Children's Bureau. It was first revised after Dr. Dettlaff joined the project in the role of evaluator. It was further revised after receiving consultation from outside sources.
- One the benefits of the logic model is that it serves as a guide in which all components are mapped out. It visualizes how all the program components interact. Logic models do not serve to constrain action; rather, they are fluid and can be continually revised to incorporate new information.
- It is recommended that grant writers create the logic model before writing the grant. After the grant is written, the logic model can be refined to include any new components gleaned from the grant writing process.

II. Promoting Cultural Competence

- The instrument used to measure cultural competence is the Attitude, Skills, and Knowledge test given to Child Protective Service workers. Trained workers should have higher cultural competency scores and more successful outcomes than untrained workers.
- Dr. Patrick Leung's cultural competence measurement instrument was a good fit for the project because it pertained specifically to the child welfare field. Other instruments from the psychology field did not fit the project as well. Although Dr. Leung's model is not commonly used, its use is encouraged for child welfare research projects by the state of Texas.
- The project promotes the Systems of Care Model to be used with families willing to cooperate with child welfare professionals. This model is considered to promote cultural competency by incorporating formal and informal community partners to develop an individualized service plan to meet the family's needs.
- During the pilot, the Systems of Care Model was successfully implemented when supervisors promoted it to their workers. Implementation was considered successful based on workers' responses that it was useful. However, these same workers were concerned that they would not be able to implement it with all clients because of high workloads. To address these concerns, it was decided that the Systems of Care Model would be used with clients who demonstrated a high probability of benefiting from the interventions.
- Implementation was not successful when supervisors did not promote it to their workers.
- Supervisors attend the same trainings as their workers to gain an understanding of how the model works in practice. As of yet, supervisors have not evaluated changes in trained workers' behaviors but recognize the importance of doing so.
- Cultural competency is currently incorporated into existing trainings and built upon to focus specifically on Hispanics. This represents an enhancement of required trainings that workers automatically receive.

- Family complexities are measured informally by their desire to cooperate with child welfare workers. Workers will be steered towards cases that are considered appropriate for the Systems of Care Model.

III. Practice Implications

- Presently, agencies have not developed policies supporting the use of different practice models although some have expressed openness to using them as the project develops.
- Agencies have cooperated completely in providing access to documents and case conferences.
- The project also accounts for unintended outcomes that the Participant Action Plan Approach assessment tool may uncover.
- Outcomes for safety, permanency, and well-being are measured by indicators that coincide with the clients' situations per Child and Family Service reports.
- Potential exists for short-term outcomes to be linked to training content for the purpose of curriculum development, although this has not occurred yet.
- A condensed version of the Systems of Care Model provided a realistic illustration of what a child welfare worker is capable of doing. The pilot also highlighted which new skills workers needed to learn to implement the model. During the training, family selection will select cooperative families in order to minimize the new skills that need to be learned.
- Example skills can be used on action plans to provide more concrete concepts while not being too constrictive or vague.

	Evaluating Trainees' Documentation Skills

Megan E. Paul, *Ph.D.*

Abstract

This article provides an overview of the development and implementation of an evaluation tool for assessing the documentation skills of new Protection and Safety Workers. The content and process of the training are described, followed by the content and format of the final scoring rubric and the scoring and feedback process. Several issues emerge related to training evaluation, including: development issues related to creating an activity as the basis of an evaluation; defining performance standards; choosing rating scales and assigning points; and administering, scoring, interpreting, and reporting the results of the evaluation.

Background

The University of Nebraska–Lincoln, Center on Children, Families and the Law provides pre-service training to all new Protection and Safety (child protective service) Workers employed by the Nebraska Health and Human Services System. This six-month employment practicum period features a combination of classroom training, field training, and limited case assignments. This paper focuses on the evaluation of one of the *Electronic and Paper Records* units, which was designed to develop trainee skills in gathering and recording case information in electronic and paper formats.

Curriculum Sequence

In the sixth day of their employment practicum, trainees receive one-half day of training on the general content and style guidelines for written narratives. This is considered *Part I* of the *Electronic and Paper Records* unit. The next 23 days of training cover a variety of topics related to case management duties and include

practice on specific types of documentation. *Part II of Electronic and Paper Records* is delivered on the 24th day of training, during which trainees spend another one-half day engaging in an information-gathering and documentation activity. This activity allows them to practice and demonstrate their documentation skills. The activity and its evaluation are described next.

Curriculum Content and Process

The activity requires trainees to prepare for, conduct, and document a mock home visit. Trainees first search the training computer system to find historical information about a hypothetical family. After answering a series of questions to guide them through the information-gathering process, they prepare questions for an upcoming required contact with the parents. Before beginning the mock visit, they consult with the trainers as they would consult with a supervisor prior to a visit. The trainees conduct a group interview of the parents, who are role-played by trainers. Trainees take turns asking questions of the parents, while taking notes on the discussion. When the interview is complete, trainees document a narrative about the meeting in the state's training computer system. Finally, each trainee prints a copy of his or her narrative and submits it to the trainer.

Trainers respond to consultation questions from trainees and then assume the role of the parents, while following a loosely structured script to answer interview questions from trainees. They complete a checklist at the conclusion of the interview to document what issues were discussed during the interview. Trainers later submit the completed narratives and checklist to a training evaluator for scoring.

Final Evaluation

Purpose

The primary purpose of the evaluation is to assess trainee skill in documenting contacts with families. Although the activity requires trainees to demonstrate several other skills, including computer navigation and interviewing, the evaluation only focuses on the quality of their documentation.

Scoring Rubric

The final scoring rubric (see Appendix A for an excerpt) includes eight performance dimensions, which coincide with the dimensions outlined in the general *Content and Style Guidelines*

trainees received in *Part I* of the training several weeks before. The content dimensions include *who, what, when, and where*; the style dimensions include *organization, grammar, clarity, and use of names and pronouns*. Within these dimensions, there are a total of 39 items tailored specifically to the mock family scenario. Each of the 39 items is scored on either a 2- or 3-point rating scale, which measures whether, how often, or how thoroughly information was documented. For example, documented information such as the worker's name, job title, and the date and time of the visit are types of items rated on a 2-point scale, anchored by *No* and *Yes*. Other types of behaviors, such as using full sentences, spelling, punctuation, and use of pronouns, are repeated throughout the narrative and are rated based on how often the standards are met. For these, a 3-point scale is used, with different frequency anchors for different items. Finally, some other types of information can be documented with varying degrees of completeness, such as family information pertaining to visitation, employment, substance abuse, and discipline. For these, a 3-point scale is used.

The format of the scoring rubric is similar to that of a Behavioral Observation Scale (BOS) format, which focuses on highly desirable or undesirable behaviors. The primary difference here is that a BOS usually assesses only the frequency of behaviors, whereas this format includes other criteria (e.g., thoroughness) in addition to frequency. Some scales also include a *Not Applicable* option because the interview content is prone to vary across training groups, depending on what interview questions trainees ask. It is therefore possible that some criteria may not always apply.

For each item, the number of potential points earned varies from 0 to 2. For 2-point scales, trainees earn either 0 or 2 points. For 3-point scales, trainees earn either 0, 1, or 2 points. Given that there are 39 items, the maximum total number of possible points is 78 when all items are deemed applicable.

Scoring Process

The scoring process includes three basic steps. First, the *Trainer Checklist* is used to identify all the *Not Applicable* ratings. Anything that was not discussed in the interview is automatically given a rating of *Not Applicable*. Next, ratings are made by a single training evaluator who reads each narrative and enters ratings into the scoring rubric, which is in an electronic format. After this step has been completed for all trainees in a given training group, the

percentage correct for each trainee and the group average are calculated. A percentage is calculated because the *Not Applicable* ratings result in different *Total Points Possible* for each group, making the *Total Points Earned* meaningless.

Feedback

Feedback to trainees and their supervisors is given as soon as possible after the narrative scoring is completed. The trainee's narrative is first scanned to create an electronic file that can be more easily and quickly distributed than hard copies. The scanned narrative and completed scoring form are sent to the trainee's assigned Field Training Specialist (FTS), who is the primary liaison between the training unit and the field. The FTS forwards these documents to both the trainee and the trainee's supervisor. The trainee, the FTS, and the supervisor also discuss the trainee's performance in more detail at their next regularly scheduled meeting. If there are immediate concerns, the FTS will contact the supervisor as soon as possible to discuss the concerns.

EVALUATION DEVELOPMENT

Creating the Exercise

In creating the exercise that will form the foundation of the evaluation, at least three questions should be addressed: (a) Does it adequately simulate what's required on the job? (b) Is it practical to administer? and (c) Will it allow for standardized evaluation across trainees and trainers? It is important to strike a balance among these criteria.

Before developing an exercise to satisfy these criteria, it is wise to consider whether there are any existing activities that already meet these standards. In this case, an exercise already existed as part of the training, but trainees' performance had never been systematically evaluated. The next questions considered were (a) Have trainees had adequate time to acquire the knowledge and skills we intend to evaluate? and (b) Does the activity need to be and can it be revised in any way?

Regarding the first question, the timing was deemed not appropriate because trainees were learning the rules of documentation and immediately applying them during this activity on the same day. This afforded very little opportunity to become familiar with the rules and no opportunity to practice applying the rules prior to the evaluated activity. To rectify this, trainees received the training on the rules of narrative writing earlier in the

training sequence. The evaluation activity remained at the same time. The writing guidelines that were originally trained during *Electronic and Paper Records: Part II* were moved to *Electronic and Paper Records: Part I*. This gave trainees ample time to become familiar with and practice applying the guidelines before being evaluated in *Part II*.

When considering the timing of an evaluation, one must also consider the organizational culture and how the results will be used. Trainees need to feel that they have received a fair opportunity to learn and that the results will be used constructively. People familiar with both the trainees and the organization should therefore make such decisions.

Consideration of the second question, about the need for activity revisions, revealed very little need for change. The only necessary improvement was to reduce the variability of the exercise across training groups and to ensure that the trainers kept their responses as standardized as possible. Such changes are best achieved when there is a cooperative partnership between evaluators and trainers or curriculum developers.

Defining the Performance Standards

After the activity is chosen or revised, the training-evaluation team must determine what standards would be used to evaluate performance. Ideally, the curriculum clearly explains these standards, preferably through a well-written learning objective, so that they are obvious to everyone, especially trainees. In this case, the standards were not sufficiently apparent for evaluation purposes. Trainees needed clear criteria for what case information they should document and how they should document it. By working intensely with the trainer and gathering input from the agency's Quality Assurance Unit, a two-page set of *Content and Style Guidelines* was developed for trainees. It clearly defined the standards for written documentation and could be used during training and as a job aid in the field.

Although an evaluation could have been implemented without these changes, it would have been no surprise to find poor performance and would have led to unfavorable reactions from trainees. These sorts of changes are probably best characterized as *formative evaluation*, although this term usually refers to changes that are made before training is implemented. However, this type of evaluation and revision is absolutely necessary in order to avoid setting up a *summative evaluation* for failure. In this case, there

was no reason to implement the new scoring system if there was little reason to expect achievement of objectives. Whenever trainees are held accountable for their performance—and perhaps if they're not—evaluators should do everything in their power to ensure that the training prepares trainees to perform well before implementing a formal evaluation.

Operationalizing the Standards

Even though the performance standards were clearly described, they had to be further defined to match the exercise and allow for easier measurement. In doing this, one must consider several factors, including: exactly what behaviors are desired and undesired whether to focus on effective behaviors or ineffective behaviors, and what specific criteria best apply to each behavior. Such details are often overlooked and perhaps taken for granted but are important and should not be hastily chosen.

Sometimes it is important to assess whether a positive behavior was exhibited. In these cases, the main alternative to exhibiting the positive behavior is simply failing to exhibit the behavior. No noteworthy negative behavior would be exhibited instead. The assessment would target the presence or absence of the positive behavior. Examples of the presence of positive behavior include “documented the date of the visit” or “described the sequence of events.” The corresponding absence of these positive behaviors would be reflected in the statements “did not document the date of the visit” and “did not describe the sequence of events.” In other instances, the more important issue is whether a negative behavior was exhibited, and the same approach applies.

When deciding which approach to take, it is important to consider the relative consequences of each behavior. Some are more important and have more impact than others. It is also important to consider the difficulty associated with making each of these judgments. Some may be much harder to rate than others, and behaviors should be described in a way that minimizes rater error. Finally, it is important to also consider the difficulty associated with scoring the instrument. Some behaviors are described in a way that necessitates reverse scoring, which can be challenging to the scorer.

Choosing Rating Scales

After the target behaviors were identified the criteria on which they would be rated had to be determined. We considered whether we expected the behaviors to vary in quality, quantity, frequency,

degree, etc. and used this as a guide to choose rating scales. As previously mentioned, the final rating scales assessed whether, how often, and how thoroughly information was documented.

In deciding on the rating options, one must try to ensure that raters have a clear and shared understanding of what each anchor means. The anchors must be clearly distinguished from one another and must be likely to be selected during the rating process. Just as with a multiple choice test with distractors that no one ever chooses, one does not want a response option that no rater ever selects. Finally, the level of variability of each rated behavior must be considered as well as how well raters could detect this variability. Ultimately, we chose anchors that would best capture important differences, or the full range of meaningful behavior.

Subject matter experts should be consulted to address these issues. If possible, the evaluator should pilot the instrument to get first-hand experience with how the behaviors and rating scales work. Alternatively, evaluators may work closely with subject matter experts to get their feedback on the instrument's performance. Practice with the tool often reveals changes that need to be made.

Assigning Points

As previously mentioned, the number of potential points earned for each item varied from 0 to 2. Because these points were used only during this evaluation and not as part of any other point system, the values themselves were somewhat arbitrary. Performance was defined as a ratio of the total points earned over the total points possible. Therefore, any point system that reflected these relative values (e.g., 100, 101, and 102) would have resulted in the same scores.

All items in this evaluation were equally weighted, such that the maximum points were equal across items. Although there is something particularly compelling and intuitively appealing about the idea of giving more weight to certain items than others, the actual weighting is usually very different than the intended weighting. Moreover, with a fairly long evaluation (with many items), the weighting has little or no significance.

Implementation

One of the primary implementation struggles was getting trainers to adhere to the script. Because of their experience teaching the exercises, trainers were accustomed to having freedom to vary their responses as needed. Trainers had to

become more disciplined about their answers to minimize scoring difficulties.

Trainees were allowed to use their *Content and Style Guidelines* during the evaluation because they have access to these guidelines on the job. This ensured that doing well in training did not require greater knowledge, skill, or ability than that required to do well on the job.

Reliability and Validity

Whenever an evaluation requires subjective performance ratings, rater training should be included as part of the implementation. This minimizes errors (with one or more raters) and maximizes agreements (necessary only with two or more raters). Rater training usually addresses the following: (a) eliminating biases, or systematic judgment errors humans tend to make when evaluating performance, and (b) achieving agreement among multiple raters. A preferred method of rater training includes lecture, practice, group discussion, and feedback. We have not yet needed to do this because we only have one rater, who is already familiar with typical rating biases and how to avoid them. If additional raters are added, training and practice will be required to establish sufficient interrater agreement.

Another important part of implementation is assessing the measurement properties of the instrument. Although the development process included steps to maximize reliability and validity, more information about these properties can be gathered once the instrument is put to use.

Rater consistency is probably the best type of reliability in these types of measures, so assessing it is desirable. Although other types of reliability, such as internal consistency, could be assessed, they may not be appropriate, depending on the extent to which the items are homogeneous. The primary index of rater consistency is interrater reliability, but it can only be assessed with multiple raters. It should be noted that this statistic is different than the interrater agreement that was mentioned earlier. With just one rater, a sort of test-retest or intrarater reliability could be estimated if the rater rates the same target multiple times. Because of the limitations associated with having heterogeneous items and only one rater, no reliability estimates are available at this time.

By assessing nearly 100% of the content domain, we have clearly established content validity, which refers to the extent to which the content of the evaluation reflects the content of the

curriculum. Looking for other evidence of validity, such as whether these scores predict other training or job performance outcomes, is another possibility. We have done this yet; however, a post-training job performance measure targets documentation skills, and we could look at the extent to which later performance is predicted by training performance.

Feedback

The final implementation issue to address is feedback. First, it is important to consider to whom feedback should be given. As mentioned before, we give feedback to the trainee's Field Training Specialist, the trainee, the trainee's supervisor, and the trainers. The first three are informed so they can get an indication of how well the *trainee* is doing; the trainer is informed to get an indication of how well *training* and the *trainer* are doing.

Also important is deciding the nature of the feedback which is, in part determined by the format and content of the evaluation. One of the main issues is whether to interpret performance in a normative fashion or relative to some criterion. In training, the latter is probably preferable because the relative standing of each trainee does not matter if none of them achieve the desired proficiency. However, the struggle is in knowing exactly what level of performance constitutes proficiency. As mentioned before, we provide the total number of points earned, the total number of points possible, and the percentage earned. These indicate the trainee's performance relative to some known criterion (100% as the maximum), but we don't have an established cut score or passing score.

Another piece of information we give is the group mean, so that scores can be interpreted in light of one another. This is more of a norm-referenced score. Other indicators would be a percentile rank for each score, although these can sometimes be problematic with small groups and when there is limited variability in scores.

The method of giving feedback is also a challenge because trainees are only able to print one copy of their narratives during training. Even they don't have a record of what they wrote. Because the ratings and scores are not easily interpreted without the original narrative, both documents must be shared. Trainees and supervisors are located throughout the state, so we prefer to use electronic communication whenever possible. One solution is to scan the originals, but this creates electronic files that often exceed the limits of trainer's email capacities. To counter this

problem, hard copies are sometimes made instead. The results are sent out as soon as the narratives are scored, which is usually within a week or two after the training.

APPENDIX A: SCORING RUBRIC EXCERPT

Trainee Name:

Training Unit: CMF 107 Accountability Through Documenting

CONTENT—WHO			
1. Worker's own full name	<input type="checkbox"/> No (0)	<input type="checkbox"/> Yes (2)	
2. Worker's job title: Protection and Safety Worker or Integrated Care Coordinator	<input type="checkbox"/> No (0)	<input type="checkbox"/> Yes (2)	
3. Who was present during visit: worker, Mrs. Samantha Swan, Mr. Steven Swan	<input type="checkbox"/> No (0)	<input type="checkbox"/> Yes (2)	
4. When mentioning anyone's name, describes who they are (Sammi, daughter; Scotty, son; Mr. White, foster parent)	<input type="checkbox"/> Never/ Rarely (0)	<input type="checkbox"/> Sometimes (1)	<input type="checkbox"/> Frequently/ Always (2)

Comments:

CONTENT—WHAT				
5. Sammy—placement	<input type="checkbox"/> N/A	<input type="checkbox"/> Little or No Info (0)	<input type="checkbox"/> Some Info (1)	<input type="checkbox"/> Complete Info (2)
<input type="checkbox"/> Was placed in foster care at the Whites' home <input type="checkbox"/> Ran away from Whites' and has not yet been located				

<input type="checkbox"/> Mr. Swan speculates that Mr. White sexually abused Sammi	
6. Visitation <input type="checkbox"/> Occurs every Saturday and one time during the week <input type="checkbox"/> Request for longer or more frequent visitation <input type="checkbox"/> Dislike supervised visitation <input type="checkbox"/> Request to have visits at home	<input type="checkbox"/> N/A <input type="checkbox"/> Little or No Info (0) <input type="checkbox"/> Some Info (1) <input type="checkbox"/> Complete Info (2)
7. Mrs. Swan—employment <input type="checkbox"/> Recently completed CNA program <input type="checkbox"/> Currently applying for jobs <input type="checkbox"/> Has no references	<input type="checkbox"/> N/A <input type="checkbox"/> Little or No Info (0) <input type="checkbox"/> Some Info (1) <input type="checkbox"/> Complete Info (2)
8. Mr. Swan—substance abuse <input type="checkbox"/> Submits to UA testing at Day Reporting Center <input type="checkbox"/> Has been clean for one month <input type="checkbox"/> Is willing to attend NA	<input type="checkbox"/> N/A <input type="checkbox"/> Little or No Info (0) <input type="checkbox"/> Some Info (1) <input type="checkbox"/> Complete Info (2)
9. Domestic Violence <input type="checkbox"/> Mrs. Swan has cut Mr. Swan with a knife <input type="checkbox"/> Both attend counseling <input type="checkbox"/> Swans perceive no relationship problems	<input type="checkbox"/> N/A <input type="checkbox"/> Little or No Info (0) <input type="checkbox"/> Some Info (1) <input type="checkbox"/> Complete Info (2)

10. Verbatim quotes when necessary or appropriate (e.g., for a child's words or emotionally charged comments)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Never/ Rarely (0)	Sometimes (1)	Frequently/ Always (2)
11. Avoids irrelevant or unnecessary information	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Sometimes (0)	Frequently (1)	Always (2)

Comments:

CONTENT—WHEN & WHERE	
12. Date of the visit	<input type="checkbox"/> <input type="checkbox"/> No (0) Yes (2)
13. Time of day	<input type="checkbox"/> <input type="checkbox"/> No (0) Yes (2)
14. Location of the visit: Swans' home, including address	<input type="checkbox"/> <input type="checkbox"/> No (0) Yes (2)

Comments:

STYLE—ORGANIZATION	
15. Synthesizes factual information and organizes topics through the use of paragraphs	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> Never/ Rarely (0) Sometimes (1) Frequently/ Always (2)

16. Clusters related ideas in a single paragraph; fully covers a topic before initiating the next topic	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Never/ Rarely (0)	Sometimes (1)	Frequently/ Always (2)

Comments:

STYLE—GRAMMAR			
17. Complete, grammatically correct sentences	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Sometimes (0)	Frequently (1)	Always (2)
18. Proper spelling and punctuation	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Sometimes (0)	Frequently (1)	Always (2)

Comments:

STYLE—CLARITY			
19. Acronyms or initialisms only after they have been defined in the same narrative; only commonly known abbreviations; avoids jargon/ department lingo	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Sometimes (0)	Frequently (1)	Always (2)
20. Short and concise sentences to facilitate understanding; (no extremely long or run-on sentences)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Never/ Rarely (0)	Sometimes (1)	Frequently/ Always (2)

Comments:

STYLE—NAMES			
21. <u>Both first and last names</u> to identify each person (may use only the first name in subsequent sentences, but only if there is not more than one person with the same first name)	<input type="checkbox"/> Never/ Rarely (0)	<input type="checkbox"/> Sometimes (1)	<input type="checkbox"/> Frequently/ Always (2)
22. Avoids informal titles such as “mom,” “dad,” etc.	<input type="checkbox"/> Never/ Rarely (0)	<input type="checkbox"/> Sometimes (1)	<input type="checkbox"/> Frequently/ Always (2)

Comments:

STYLE—PRONOUNS			
23. Reference of pronouns (e.g., we, he, she, him, her, it, they) are clear	<input type="checkbox"/> Sometimes (0)	<input type="checkbox"/> Frequently (1)	<input type="checkbox"/> Always (2)

Comments:

Total Score:

Group Average:

	Performance Based Contracting Aftercare Training and Evaluation— Lessons Learned in Aftercare

June Cairns, *M.S.W.*, Philadelphia Department of Human Services,
John H. Trudeau, *Ph.D.*, Temple University—CSPCD,
Amanda Schreiber, *M.S.W., L.S.W.*, Temple University—CSPCD,
and Maria G. Ramos, *M.S.W.*, Temple University—CSPCD

Abstract

As part of a transition to a performance-based contracting model for the provision of foster care services, Philadelphia's Department of Human Services trained foster care agency staff on the development of aftercare plans. The training was formulated after an assessment of existing plans. The training was evaluated using trainee satisfaction forms (including self report of increased knowledge and attitude change), knowledge testing, and systemic review of a random sample of Aftercare Plans. While trainees reported satisfaction with the training and gained knowledge, aftercare plans were still deemed incomplete three months post training. The authors discuss several hypotheses for why transfer of skills did not appear to occur.

Background

There is universal recognition in the child welfare system that children too often have extended stays in foster and kinship care homes before they achieve permanency. The Adoption and Safe Families Act of 1997 was the federal response to the large numbers of young people who remained in foster care too long. The Philadelphia Department of Human Services (DHS) recognized this concern in their own child welfare department, and under the direction of former Commissioner Alba Martinez, Philadelphia's Permanency Initiative began. Because approximately 85% of DHS's services are contracted out to local

nonprofit children and youth services, one branch of the initiative was the creation of a Performance-Based Contracting (PBC) process that began on March 3, 2003, for the nonprofits providing foster and kinship care for the Department. The implementation of PBC fundamentally changed the relationship between the Department and its provider agencies. Simply put, provider agencies received compensation for their services based upon their ability to safely but efficiently move children to permanency. This differed from the previous per diem basis by creating an incentive to move children more quickly toward permanency; thus, for the first time provider agencies could be financially rewarded or penalized based upon their ability to move children safely into permanency.

Concerns about PBC contracting reverberated throughout DHS, provider agencies, and other stakeholders involved in the child welfare system. Provider agencies especially wanted to ensure that they had the knowledge and tools necessary to achieve the permanency outcomes set forth by PBC. In response to this concern, DHS partnered with the Fels Institute of Government at the University of Pennsylvania, which has many years of expertise in helping government agencies redesign themselves, to examine the City's PBC process, in particular the component of aftercare services once permanency had been achieved.

Fels' analysis of and recommendations for aftercare planning were used to create a tailored training to fill-in the knowledge and skill gaps of the provider agencies. These trainings supplied them with the tools necessary to achieve their permanency outcomes under PBC. DHS maintained its long-established partnership with Temple University's Center for Social Policy and Community Development (CSPCD) to design and execute an Aftercare Training and Evaluation based upon the recommendations provided by the Fels Institute's needs assessment. Some of the Aftercare Training Learning Objectives were as follows:

- Define Aftercare as it pertains to children and youth leaving DHS foster care;
- Understand the core values, theory and philosophy guiding Aftercare planning;
- Understand the 6 domains to be addressed in each Aftercare Plan;

- Understand and be able to describe a broad spectrum of potential Aftercare approaches and services which can be included in Aftercare planning;
- Understand how a child and family's individual needs can be assessed and addressed in Aftercare planning;
- Develop child and family-specific Aftercare Plans in collaboration with the families and youth themselves;
- Understand the roles and responsibilities of all parties in the development and implementation of an Aftercare Plan.

Aftercare is defined as providing flexible goods and services to support family reunification and prevent child/ren from returning to care. The Aftercare Plan creates a “road map” that guides the family toward permanency, and is the primary vehicle for providing services once permanency is achieved. Therefore, it was important to DHS and the providers that a tool be designed to evaluate Aftercare Plans in a systematic manner.

Measurement Tools

Aftercare Plan Assessment

The Fels Institute created an Aftercare Plan Assessment Tool to identify the strengths and challenges of the Aftercare Plans. This information, while being used to develop the domains of training skill development, was then used in conjunction with the Aftercare Training to evaluate the quality of the Aftercare Plans written by provider agencies after the training. The Fels Institute assessed Aftercare Plans pre-training, while CPSCD evaluated the plans post-training; DHS used the data supplied to them to determine if there had been any meaningful change in Aftercare planning

The Fels Institute created 6 Aftercare domains with several subdomains which were used for evaluation based upon completeness and timeliness in the Aftercare Plan (*see Table 1 for details*).

Table 1. Aftercare Domains

Domain	Subdomain
Safety	Child Safety
	Parent Safety
Subsistence	Adequate living arrangement
	Adequate income
Emotional/Psychological	Family prepared for return
	Child emotional needs
	Parent emotional needs
Medical	Child medical issues
	Parent medical issues
	Health insurance
Educational/Vocational	Child enrollment
	Child special education needs
	Parent employment assistance
Family/Social Relationships	Parenting support
	Support network
	Link to community services
	Child recreation

Aftercare plans were categorized as: *incomplete*, *somewhat complete* and *fully complete* based upon how the provider social worker wrote the plan for the identified issues. Plans were considered *complete* if they contained the following criteria:

- Indicated why the child was placed into care;
- Described all the actions/steps child’s family has taken to move toward reunification (i.e., substance abuse treatment, finding suitable housing, etc.);
- Specifically stated the goals for Aftercare;
- Specifically stated whether the goals for the child while in care were met;
- Addressed all issues of the family from a strengths-based perspective;

Included all service providers by name, address, phone number and scheduled appointment time. Aftercare plans were considered timely if they were submitted to DHS prior to the day of reunification (i.e., between 1 and 40 days).

State-Mandated Evaluation Survey

Since the Aftercare Training was approved by the Pennsylvania Child Welfare Training Program, each training participant was asked to complete the State's standard survey form: *Child Welfare Competency-Based Training and Certification Program Feedback/Training Evaluation*. This standard form uses Likert Scale scoring to capture trainees' views on the usefulness of training content, trainers' presentation skills, and cultural sensitivity of the content area. It also gathers trainees' feedback on their increased understanding of the topic area.

CSPCD Aftercare Survey

CSPCD designed an additional evaluation survey to assess the trainees' knowledge gains about Aftercare training objectives through several Likert scale, True/False, and fill-in-the-blank questions.

Evaluation Results

Aftercare Plan Assessment

Three months after the completion of Aftercare Training, 49 provider cases were randomly chosen by DHS for review. Of the plans reviewed, 29% were considered to be completed in a timely fashion. Aftercare policy states that if an Aftercare Plan cannot be completed before reunification, the reason should be stated on the plan. Of the selected plans, none provided this information. Most plans identified goals in at least several areas of each domain. However, plans seemed to consistently provided vague information, such as "participating in mental health treatment," instead of detailed information about the type and location of the mental health treatment. Pertinent contact information was also left out despite training's instructions. Plans were fairly consistent in identifying the child's school, but did not consistently report the child's grades or other special education needs. Most of the plans reviewed, however, did specify the roles of the Aftercare Worker and the type and frequency of services provided. As was encouraged in the training, most services were "frontloaded" (i.e., more intense involvement that tapers off over time). Unfortunately, when plans were compared by agency, it was found that service outlined for each child were often very similar instead of tailored to meet the unique needs of the family, as training taught.

State-Mandated Evaluation Survey

The State-Mandated Evaluation Survey reported very high positive feedback about the training with 100% of trainees expressing that the Aftercare Training was either *excellent*, *very good* or *good*. Trainees also reported a 25% increase in their level of understanding before the training as compared to after the training. Eighty-five percent of trainees also reported that they had gained new knowledge from the training pertinent to their job and 40% reported that they had benefited from a change in attitude that would help them in their job.

CSPCD Aftercare Survey

CSPCD's Aftercare Survey found that 66% of trainees identified all six domains necessary to complete a comprehensive Aftercare Plan, while 94% of trainees were able to identify at least 4 parties who should be invited to an Aftercare Meeting. Over two-thirds (67%) of trainees also reported that after the training they clearly understood Aftercare's core values, theories, and philosophies, and over half (57%) of trainees reported that they now understood the importance of creating individualized and specific plans tailored to a child's unique needs. Finally, over 80% of trainees correctly answered five True/False questions addressing specific Aftercare planning procedures.

Lessons Learned

The major lesson learned through this training cycle was that, despite the overwhelmingly positive survey feedback about the value of the training and the new skills, information and perspectives learned, the overall quality of Aftercare Plans that were examined post-training were not well designed. In other words, there seemed to be limited "transfer of learning" to practice. Although trainees were provided specific guidelines in how to create *fully complete* Aftercare Plans, very few met those expectations in their actual field practice.

DHS/CSPCD created several hypotheses as to why this transfer of learning did not occur:

- 1) The training attempted to provide too much information in too short a time span. PBC was a drastic shift from previous practice and roles between DHS and provider agencies. Therefore, it may have been unrealistic to expect social workers to be able to apply these new guidelines without further feedback or training.

- 2) The training may have been too academic and not practice-oriented enough to support transfer of learning skills. During a small focus group session with provider social workers post-training, a recurring theme was that the training did not provide enough “how to” skills.
- 3) In general, provider agency social workers tend to be young, and inexperienced in the social service field and also change jobs frequently. These combined factors make it difficult to ensure that all new entries into the field are well versed on current standards, especially when they have limited knowledge of the field on which to base their practice.
- 4) The Aftercare Plan documentation is much like discharge planning in many ways. Discharge planning is not a typical activity for child welfare workers and was therefore a new skill to develop for both supervisors and provider line-staff. Supervisors were inexperienced in assisting line-staff with the development of such plans as they themselves had not yet mastered the process.
- 5) The start of PBC contracting created a general air of anxiety, confusion and apprehension throughout the child welfare field in Philadelphia. In the initial phases of PBC, provider agencies had many concerns about the new system and their roles and responsibilities that needed to be addressed with DHS. Therefore, it may have been unrealistic to expect provider agencies to immediately implement and enforce all of the guidelines provided in the Aftercare training with their staff.

An anecdotal review of Aftercare plans three years post-training does seem to suggest that there have been some improvements in the completeness and consistency of plans submitted to DHS currently. Presently, it seems as if provider agency staff are slowly gaining confidence in their ability to perform the many new roles demanded of them by PBC.

Based upon the evaluation knowledge gathered from the Aftercare Training, it is recommended that future trainings:

- 1) Be designed to be more hands-on and interactive to encourage transfer of learning to field practice;

Lessons Learned in Aftercare

- 2) Limit training material taught in one day to ensure that the amount of information provided is “digestible” to trainees;
- 3) Design support networks post-training to allow trainees to revisit and solidify new practice skills.

In conclusion, the Aftercare Training initiative reflected the difficulty in training a large population of child welfare staff on major policy changes. Although training was not immediately effective, it did create a foundation from which new practices could be built, slowly, as the new child welfare culture was indoctrinated by PBC agencies over time.

	Discussion: Aftercare: Planning for Post-Reunification Services

Theme: Methods and Measuring Skill

Title: Aftercare: Planning for Post-Reunification Services

Presenter: John Trudeau, *Ph.D.*

Facilitator: Jane Berdie, *M.S.W.*

Topics of Discussion

Dr. Trudeau explained how the use of trainee-designed aftercare plans was used for training evaluation. His presentation also uncovered the potential of quality assurance instruments for evaluating trainings. Unfortunately, as a result of recording difficulties, the following is not a complete summary of the discussion following Dr. Trudeau's presentation.

I. Using Quality Assurance as an Evaluation Tool

- One county designed exit surveys specific to training and development needs. An employee's exit from the field represents a key moment for evaluation. Thus, these surveys serve as a needs assessment or an indicator of evaluation of training. Additionally, they also have potential research purposes.
- In Nebraska, quality assurance and training evaluation are coincidentally linked. Nebraska Children's Services, because it does both, has had to determine how to connect them. Currently, the connection has been successful, although quality assurance is still new as an evaluation tool. Thus, even though it is proving to work, it is unknown if quality assurance can be applied to other fields though it does appear to have that potential.
- As of yet, other symposium participants have not used quality assurance instruments for training evaluation.
- Trainings that already use quality assurance instruments can be changed to incorporate an embedded evaluation designed specifically for the training. For example, trainees

- could complete part of an aftercare plan that would be used to promote learning as well as a measure of skill acquisition.

II. Aftercare Plans

- A pre/post aftercare assessment was created by the Fels Institute. Fels completed a pre aftercare assessment while CSPCD completed a post aftercare assessment using the Fels Instrument.. Pre-tests measured workers' pre-existing knowledge on how to complete Aftercare plans; the results directed curriculum development. Post-tests measured trainees' adherence to creating thorough aftercare plans after training. DHS used the data collected from these two agencies to compare the change in workers' practice internally.

	Using POLYDETECT in Assessing the Possible Number of Constructs in Social Work Surveys

Basil Qaqish, *Ph.D.*

Abstract

This article reviews the use of POLYDETECT, a data analysis software that may be used for cluster analysis. The software is not commercially available, but can be obtained by contacting its developer. A survey example is offered to illustrate the usefulness of POLYDETECT in evaluating the number of constructs the survey is trying to measure.

Background and Introduction

Cluster analysis can assist training evaluators to detect and analyze underlying constructs measured by test and survey data. This activity provides an overview of POLYDETECT and offers an example to show how to run POLYDETECT. POLYDETECT is, in a sense, a continuation of DETECT, a program designed to assess the possible number of constructs underlying psychological or educational tests or exams. The DETECT procedure is non-parametric. DETECT conducts a dimensionally sensitive cluster analysis, then uses a genetic algorithm to search through all possible item partitions to find the partition that maximizes the DETECT D statistic. The program provides results of the total amount of dimensionality in the test being analyzed and provides a value for “r”, a measure of the degree to which the test satisfies the assumption of approximate simple structure (unidimensionality). Simulation studies indicate that a D value less than 0.1 indicate that the test is unidimensional. That is, it possibly measures one construct. “R” values greater than 0.8 indicate an approximate simple structure. Stout et al. (1996) provides more information on DETECT.

POLYDETECT extends DETECT's functionality to polytomously scored data. The matrix that POLYDETECT analyzes may look like this:

12333342251124344545322.etc.
42312124424242423132452.etc.
..... etc.

Every column represents the responses of examinees to a certain item, and every row represents the responses of the same examinee to the different items on the test.

This type of analysis may also be used for tests that are polytomously scored by more than one reviewer. For instance, a writing examination questions may be scored to classify students' responses to one of four levels of competence. Surveys have a similar structure. These instruments are intended to measure certain constructs, and to ascertain how much of that construct is present for each person. POLYDETECT may assist the researcher in analyzing a Likert Scale-type survey to determine which constructs it measures.

Using POLYDETECT: An example

An example with some sample data will illustrate POLYDETECT's potential. There are three files that need to be available to run this program:

1. The Survey File: The survey files itself. It has to be Likert Scale type data and needs to be in a text format, (is has to have a .txt extension). In this example, the file name "susan.txt" will be used to illustrate survey data. This file contains data from a survey done by Professor Susan Dennison of the Department of Social Work at the University of North Carolina at Greensboro. The survey instrument itself is provided in appendix A. The original numbering of the items has been changed to correspond to the items as they appear in the file "susan.txt". The first few questions of the survey contained demographic information. This was deleted to obtain the text file suitable for use in running POLYDETECT. The resulting file resembles Figure 1.

Figure 1. Sample Text Data

```
33333333333321322112112122222223333
235343434233221111111111112111111
234334444444353332435455333434234333
324443423334133121222455442324445421
33343344334453322244444444455444343
23334243324433222222333234343223322
53434343555521151131111551154535333
33234344332352222222333222242333222
```

The survey file included more than 600 surveys (rows in the text file). The minimum number of surveys, or rows, needed to run POLYDETECT is 400.

In this survey, Professor Dennison theorized that the survey is measuring the following three constructs:

- a. Knowledge about the social work profession (items 1 through 12 in the text file “susan.txt”. This includes the first 12 columns in the survey file.)
- b. Interest in working in a social work setting (items 13 through 30 in the text file “susan.txt”. This includes columns 13 through 30 in the survey file.)
- c. Interest in job positions in social work (items 31 through 35 in the text file “susan.txt”. This includes columns 31 through 35 in the survey file.)

The last column in the survey was for a different question. Professor Dennison theorized that it did not belong to any of the three above mentioned constructs.

2. The Instruction File: An instruction file is also needed. This instruction file tells the POLYDETECT program which file to use in its analysis, in addition to other pieces of information that the program needs. Figure 2 shows what the instructions file looked like for this example:

Figure 2. Instruction file for susan.txt.

```
SPECIFICATION FILE FOR DETECT

----- DESCRIPTION ----- | -- NAME - | VALUE -----
Input response data file:   | fin      | susan.txt
Number of items in test:   | nn       | 36
Total perfect score:       | nT      | 180
Number of examinees:       | jj       | 616
Format of item responses:  | dfmt    | (36i1)
Min # examinees in a cell: | mincell  | 10
Number of items to mutate: | m        | 7
max. # of dimensions:      | mxdu    | 8
Random seed:                | iseed   | -3098

To do confirmatory analysis:
  (1=yes, 0=no)             | itruep   | 0
  If yes, enter your        | partition|

To do a cross validation:
  (1=yes, 0=no)             | icv      | 0
  If yes, enter a data file | fcv:     |
  and # of examinees       | jcv:     |
  item response format:    | cvfmt    |

Output brief file name:    | fbrf     | outbrf.txt
Output complete file name: | fcmp     | outcmp.txt
Output correlation file name: | fcor    | outcor.txt
```

The user is not allowed to change anything in this file except what comes under “VALUE”. Note that “susan.txt” is the survey file we want POLYDETECT to analyze. The number of items in this file is 36 and the data had a scale of 1 to 5. If one scores 5 for each question, then the total perfect score will be 180 (or 36*5). The file contained 616 rows of data, which corresponds to the number of people in the survey. The format of item responses (36i1) means that the survey has 36 columns or questions. Each row has responses to the 36 questions moving one column or cell at a time. This means each cell in that specific row corresponds to a response to one of the questions. One should change the number “36” here to correspond to the number of items in the survey file being analyzed. Professor Dennison theorized that the survey measured three constructs but the maximum number of dimensions exceeds that to give the program some leverage in its estimates. In this case, 8 dimensions were identified as a maximum. One could have reasonably used 6 or 7 instead of 8.

At the bottom of Figure 2 are the names of the output files that the program creates. This paper will discuss some of the output given in the file “outbrf.txt”. Once this file is created, the user should save it for future use.

3. The Execution File: The program “main.exe” is needed to run the application. This is the POLYDETECT execution program. Once the user runs it, he or she will get this message:

This is the beta version of Polydetect developed by Dr. Jinming Zhang. It is for research purposes only. Please enter your input specification file:

The input specification file is the instructions file in the example and specified the file “susan.txt” in it. In this case, “in2.txt” is the instructions file name. The user should enter “in2.txt” and hit enter on the keyboard. The program runs quickly and will produce output files, one of which is “outbrf.txt”.

The Output File “outbrf.txt”: Figure 3 shows some sections of the output file:

Figure 3. Output files in POLYDETECT, susan.txt. example

```

***** PART ONE: RESULTS BASED ON CONDITIONAL
COVARIANCES *****

      **** DETECT Results ****

*Preliminary Dimensional Diagnosis of Your Data*

The test data may be multidimensional.
The number of (essential) dimensions may be:  4

The number of clusters with maximum DETECT is:
4
The maximum DETECT value is:
6.3626
The Iss index value is:
0.5270
The ratio r is:
0.7567

The cluster memberships are as follows

*CLUSTER  1*
```

Using PolyDETECT in Social Work Surveys

1 4 6 11 12 15 31 32 33 34 35 36

CLUSTER 2

2 3 5 7 8 9 10

CLUSTER 3

13 16 17 18 20 25 29

CLUSTER 4

14 19 21 22 23 24 26 27 28 30

***** PART TWO: RESULTS BASED ON CONDITIONAL
CORRELATIONS *****

**** DETECT Results ****

Preliminary Dimensional Diagnosis of Your Data

The test data may be multidimensional.
The number of (essential) dimensions may be: 3

The number of clusters with maximum DETECT is:
3

The maximum DETECT value is:
0.0805

The Iss index value is:
0.5587

The ratio r is:
0.7536

The cluster memberships are as follows

CLUSTER 1

1 3 4 5 6 7 8 9 10 11 12 36

CLUSTER 2

2 14 19 21 22 23 24 26 27 28 30

CLUSTER 3

13 15 16 17 18 20 25 29 31 32 33 34 35

The program gives two solution clusterings of items: one is based on the item covariances, and the other is based on the items' correlations. From our example, the first solution indicates 4 clusters. Notice that the ratio "r" is not greater than 0.8., a value of 0.8 or above is indicative of approximate simple structure. Also, the DETECT value is not less than 0.1, a value below which essential unidimensionality is assumed. In the second solution the DETECT value is 0.0805, indicating possible unidimensionality, while the "r" value of 0.7536 indicating possible multidimensionality. It is the researcher's responsibility to decide whether the solution makes sense or not in my view. The role of the software is to offer a new way of seeing the items that may help the researcher "see" new aspects of his or her data. POLYDETECT is just another tool to help in the analysis of the research. The researcher must always decide on what item groupings make sense. This is no different than doing cluster analysis and looking at dendograms. It is up to the researcher to see where to draw the line and categorize the items.

Review of the Results

In the example, we see the items clusters (*Figure 3*). Professor Dennison theorized that items 1 through 12 are measuring one construct. Note that cluster 2 in the first solution and cluster 1 in the second solution contain most of those items. Now Professor Dennison must analyze why certain items did not show up in that cluster, or why some other items were included in that cluster. POLYDETECT solutions may shed light on an issue that the researcher did not see in their work. The solutions do not have to match what the researcher theorized, but they probably provide additional information to the researcher that went unnoticed previously. Researchers should search each item cluster for additional information regarding the available data. Also, one should remember that the software is only analyzing numbers, and real life data are not perfect. As a result, it is very possible to have specific items in the clusters that do not necessarily belong to that specific cluster. In this case, the researcher decides the appropriate action.

How to obtain POLYDETECT

The program is not available commercially at this time. One can ask its author, Dr. Jinming Zhang at ETS for a copy of the program at the following address:

Dr. Jinming Zhang
Educational Testing Service
MS 02-T
Rosedale Road
Princeton, NJ 08641

jzhang@ets.org

Dr. Zhang was gracious enough to give the author permission to use the software for research purposes. He may do the same for other researchers.

Reference

- Zhang, J. (2004). Conditional covariance theory and DETECT for polytomous items. ETS Research Report No. RR-04-50. Princeton, NJ: Educational Testing Service.
- Zhang, J., & Stout, W. (1999). The theoretical DETECT index of dimensionality and its application to approximate simple structure. *Psychometrika*, 64, 213-249.
- Stout, W., Habing, B., Douglas, J., Kim, H.R., Roussos, L., & Zhang, J. (1996). Conditional covariance-based non-parametric multidimensionality assessment. *Applied Psychological Measurement*, 20, 331-354.

APPENDIX

College Students' Attitudes Toward Social Work Survey

Shade the circle that corresponds to the response that most closely indicates your agreement with each of the items below.

- a. Major or Undecided: _____ Date: _____
- b. Year in College: Freshman Sophomore Junior Senior
- c. Gender: Male Female 4. Age: 20 or younger 21 to 24
 25 or older
- d. Ethnicity: Caucasian African American Hispanic
 Other _____
- e. Is anyone in your immediate family a social worker? Yes No
- f. Is anyone in your extended family a social worker? Yes No
- g. Do you have a close friend who is a social worker? Yes No
- h. Have you ever worked or volunteered in a social service agency?
 Yes No

Please give us your opinion of each of the following statements.

(Knowledge about Social Work—12 items)

	Strongly Agree	Agree	Somewhat Agree	Disagree	Strongly Disagree
1. I am very knowledgeable about the field of social work.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. The salary range for social workers is low compared to other helping professions.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3. Women are better suited for social work as a profession.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. The job market is good for someone with a social work degree.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5. A Bachelor's degree in social work is one of the easier degrees to earn.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6. The social work profession has a very positive image.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7. You can become a social worker without a social work degree.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

8. There are limited international job opportunities for social workers.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
9. Social work is mainly for those who want to work with families and their children.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
10. Social workers often have to take children away from their families due to abuse or neglect.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
11. There are a wide variety of job settings for social workers.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
12. There are a wide variety of roles for social workers within agency settings.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

How interested are you in working in the following settings? (18 items)

	Very Interested	Interested	Somewhat Interested	Not Very Interested	Not at All Interested
13. Public or private schools	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14. Medical setting (hospital, public health, etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
15. Department of Social Services	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
16. Employee Assistance program in a corporation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
17. Juvenile center/Training school	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
18. Prison or detention setting	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
19. Hospice or a facility working with the terminally ill	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Using PolyDETECT in Social Work Surveys

20. Outreach programs for immigrants	0	0	0	0	0
21. Veteran's Medical Center	0	0	0	0	0
22. Substance abuse treatment program	0	0	0	0	0
23. Community Mental Health Center	0	0	0	0	0
24. Psychiatric hospital	0	0	0	0	0
25. Youth residential group home	0	0	0	0	0
26. Adult residential group home	0	0	0	0	0
27. Nursing home or elderly care	0	0	0	0	0
28. Rehabilitation center	0	0	0	0	0
29. University/college setting	0	0	0	0	0
30. Medical outreach (i.e., with AIDS patients, diabetics, cancer survivors)	0	0	0	0	0

How likely would you be interested in the following job positions? (5 items)

	Very Interested	Interested	Somewhat Interested	Not Very Interested	Not at All Interested
31. Agency administrator	0	0	0	0	0
32. Trainer/Consultant	0	0	0	0	0
33. Counselor/therapist	0	0	0	0	0
34. Case Manager	0	0	0	0	0
35. Case worker	0	0	0	0	0

How interested are you:

36. In choosing social work as your major	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
---	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------

Thank you for taking the time to complete this survey!

	Discussion: How to POLYDETECT

Theme: Applying POLYDETECT to Training Evaluation

Title: Using POLYDETECT to Assist in Assessing the Number of Constructs in a Survey.

Presenter: Basil Qaqish, *Ph.D.*

Dr. Qaqish presented the POLYDETECT program and how it can be used for training evaluation. He described how it can serve as another research tool that can provide results that other tools may not uncover. Dr. Qaqish also stressed that the researcher's discretion ultimately decides how to appropriately apply any research to a study.

Topics of Discussion

I. How POLYDETECT Works

- Items that are highly correlated are grouped together based on Likert Scale responses. Example items include: "What is a student's perspective of their knowledge of social work?" and "What are their future plans going into the social work field?" These are attitude questions that one might hypothesize would be clustered together.
- The program determines which items can be grouped together from a quantitative perspective. It uses formulas to calculate certain values such as the ISS index value and the maximum DETECT value. These values indicate what the items are maximizing.
- The items' DETECT value determines how items will be clustered together. Values that are close to the critical D and r values may pose a problem in assessing the survey dimensionality.
- Items are not confined to one construct. They are rotated until they are matched with their highest detect value. The rotation forces one solution per cluster based on the highest detect values.

II. Using POLYDETECT as Another Measurement Tool

- Based on the theory that research is strengthened through use of multiple assessment tools, POLYDETECT represents another tool for this purpose. It is important to determine which software works best for which kinds of tests.
- Caution should be used in determining which assessment tools to use. All assessment tools make assumptions about the nature of the data you use. The tool one chooses depends on which assumptions one buys into. It is wrong to choose tools that promote our assumptions. This is a misuse of tools and can result in false data and distorted and biased research findings. However, POLYDETECT is non-parametric, which suggests that it does not operate from assumptions. Rather, it relies on the actual data.
- POLYDETECT can be used for test construction in testing the validity of items. It will help researchers realize if they are missing items by providing perspectives that other surveys did not detect. It is useful for social work surveys because they require more flexibility and may be less reliable than standardized tests.
- This program helps researchers see the internal workings of items.
- People are not very familiar with the program. It is therefore important to explain how it works when describing it in research.
- Ultimately, the researcher's discretion plays the largest role in appropriately applying POLYDETECT to research.

	Using Interpersonal Process Recall (IPR) to Improve Child Welfare Worker Effectiveness

Patricia G. Taylor, *Ph.D., LCSW*

Abstract

Interpersonal Process Recall (Kagan, 1980) has been underutilized to improve child welfare worker effectiveness. Video clips made during a unique training of Title IV-E graduate students demonstrated how *IPR* can be used in child welfare training to evaluate assumptions, attitudes, and values, particularly in the areas of cultural sensitivity and competency. These clips showed dramatic personal growth as a student met individually with a teacher/practitioner. The innovative use of *IPR*, not generally applied outside of medical and nursing training, is helpful in refining child welfare trainings and measuring the effectiveness in both a visual and quantitative fashion. Over a four-year period while this training was being offered, significant differences in burnout rate and intellectual/ethical development were found between students who participated in this unique training and a comparison group who did not. Both the use of *IPR* in the training and the *Scale of Intellectual Development (SID-IV)* (Erwin, 1981), a 115-item Likert-type scale instrument that measures four dimensions of cognitive development, were proposed as possible methods for use in specific aspects of Human Services Training. Specifically, the *SID-IV* was proposed as a predictor for retention among child welfare workers based upon the findings of this four-year study.

Interpersonal Process Recall (IPR)

Interpersonal Process Recall (Kagan, 1980) is an established model of supervision that utilizes video recording with the supervisee acting in the role of either social worker or client and

interacting either with a client or supervisor. Once the video recording has been done, the supervisee reviews the tape, and chooses what portion of the tape to discuss with the supervisor. This choice of which portion of the tape by the supervisee to discuss with the supervisor is a departure from other supervisory modalities that use videotaping, where the supervisor selects the tape excerpt to be discussed, since it puts the responsibility for selection and reflection in the hands of the supervisee, not the supervisor. Kagan has called this the development of the “Internal Supervisor” while others have coined the phrase “supervision without specific intervention” for *IPR*.

Research has consistently supported the use of *IPR* as an effective supervision modality, particularly for beginning practitioners. Two separate studies, Kagan, Krathwohl, Goldberg, & Campbell (1967) and Kingdon (1975), found that clients of practitioners being supervised with *IPR* fared better than clients of those supervised by other methods. Though not recommended as the sole method of supervision (Bernard & Goodyear, 1992), the method has been demonstrated to be particularly useful in providing practitioners with the opportunity to confront their interpersonal issues of self-awareness and self-reflection in order to maximize the interpersonal encounter with their clients (Kagan, 1980). In fact, this is the purpose for which Kagan designed the *IPR* method.

Though *IPR* has been discussed in the literature since 1980, its use in social work education has been limited. The use of *IPR* in the fields of education and medicine has been widespread in recent years, and this paper proposes the use of this supervisory approach in the training of child welfare workers. As early as the 1950s, an eminent psychiatric caseworker in New York City, Hyman Grossbard (1954), stressed self-awareness as essential to social workers professional growth. Nearly 50 years later, Mary Ellen Kondrat (1999), another prominent social worker, reiterated that same theme, while both Gila Acker (1999) and Barbara Dane (2000) called attention to the need for social work curricula that minimizes the impact of clients’ conditions on social workers so that students will be adequately prepared to practice upon graduation. Nowhere is this more of an issue than for child welfare workers, who arguably see the most difficult human conditions outside of war zones. Harry Goldstein (2001), just prior to his

death, stressed the need for those entering the social work profession to become socialized to the profession by becoming self-aware and self-reflective. *Interpersonal Process Recall* helps the student to acquire such insights.

Use of LPR in *SELF* Course

A uniquely designed course to help socialize students to the profession and answer some of these needs of self-awareness was added to a large metropolitan university's graduate social work curriculum in 1999, especially for those preparing for child welfare practice. This course, entitled *Self-Examination of Life Foundations (SELF)*, incorporated eight reflective/experiential learning modalities with five aspects of self-discovery to help students develop an integrated sense of professional/personal self. These five areas of self-discovery are: (1) Self-awareness, (2) Personal Bias, (3) Self Disclosure, (4) Family of Origin Issues, and (5) Self Care. The unique structure of the course featured a teacher/practitioner model with one-on-one and small group learning. *SELF* was taught to a total of 24 students from 1999–2003, 18 of whom were participants in the Title IV-E (child welfare education) stipend program at that university. It was taught as an elective Independent study course with one on one learning and in a teacher/practitioner model.

The methods used in the *SELF* course were highlighted in a mixed method doctoral dissertation (hereafter, the *SELF* study) (Taylor, 2004) completed in 2004. The central aim of the *SELF* study was to focus on the unique contributions that the *SELF*'s structure, process, and content of the reflective/experiential learning modalities made in the students' growth as professionals upon graduation. The participants in the study included men and women who graduated from the MSW program in 1999–2003. The study group (Group A) consisted of 16 of 24 total *SELF* participants who agreed to complete the study questionnaires. The comparison group (Group B), consisted of 21 graduates from the same graduate social work program, during the same time frame who did not participate in *SELF*. The members of Group B were recruited through email solicitation and the alumni newsletter.

The ultimate aim of the *SELF* course was to help students develop a personal-professional integrative perspective. Literature suggests that if this is achieved, social work graduates will have:

(1) Improved relationship-building skills; (2) Reduced anxiety; (3) Enhanced cultural competence; (4) Enhanced ethical practice; (5) Reduced burnout; and (6) Improved practice competency. One of the reflective/experiential learning modalities used was Interpersonal Process Recall (IPR) (Kagan, 1980). Other modalities included: 1) Individualized learning process (one on one); 2) Video recording of learning; 3) Journaling; 4) Practice-oriented classroom learning; 5) Reading assignments; 6) Reflective learning through writing; and 7) Teacher-practitioner teaching model. Participants in the *SELF* study reported that the IPR, along with the other modalities, helped them to process, integrates their self-discovery, and socialized them into the social work profession.

The qualitative data from the students' written materials suggested that Interpersonal Recall Process (Kagan, 1980) used in *SELF* holds great potential for classroom and field use in the process of students' self-exploration. These findings support the recommendation of those authors (Acker, 1999; Dane, 2000; Goldstein, 2001; Grossbard, 1954; Kondrat, 1999) who have called for greater self-awareness, as well as curricula aimed at minimizing burnout and maximizing socialization for social work professionals. Educational methods based on self-examination of personal diverse experiences allow for students and practitioners to discuss and prevent burnout before it occurs. The *SELF* course was aimed at demonstrating how diverse learning methods could be integrated into social work students' self-examination processes to reduce burnout in the profession.

The quantitative data from the *SELF* study showed that the *SELF* students had a statistically significantly lower burnout rate than a comparison group of students who had graduated from the same university during the same time period but had not taken the *SELF* course. Groups A and B were compared on three factors of a widely used instrument used to measure burnout in a variety of helping professions, Maslach's Burnout Inventory for Human Services (MBI-HSS), (Maslach & Jackson, 1986). The MBI-HSS is a self-administered, 22-item, Likert-type scale inventory that measures three facets of burnout: emotional exhaustion, depersonalization, and lack of personal accomplishment. Group A scored significantly better than Group B in the area of personal accomplishment. Although Group A also scored lower burnout rates than Group B in both emotional exhaustion and

depersonalization, it was not of high enough statistical significance for inclusion in the study findings.

Application of IPR for Specific Supervision: A Case Example

A case example demonstrates the effective use of IPR. During spring 2003, a Title IV-E student, attending the same university where the *SELF* course was being given, was having challenges in her field placement that were jeopardizing her graduate education. Notes from the Field Instructor indicated that this student “did not respond well to constructive professional criticism” and failed to use supervision constructively.” This student was of advanced standing, earning A’s in all graduate classes and had passed the Basic Skills Training at the Department. However, she was failing her Master’s level field placement at Children’s Protective Services. Notes from the Field Instructor said that the student seemed “challenged by professional boundaries” and “needs to develop a professional communication style.” Probably the most telling statement from the field was that the student had “limited levels of self-awareness.”

In consultation with the student, the principal investigator (PI) of the Child Welfare Education Project constructed an individualized *SELF*-type independent study course aimed at self-awareness. Many of the same elements of the *SELF* course were included in the course outline. Based on research findings and anecdotal data from *SELF* students, the PI decided to use *IPR* as a supervision modality with this student.

Early in the semester, the student was asked by the PI to meet with a doctoral Graduate Associate (GA) in front of a video camera set up in a private area of the school. She was asked by the GA to demonstrate particular emotions in front of the camera as she remained seated and as the GA directed her by telling her what emotions to demonstrate. Although the student was asked to demonstrate a number of emotions, the only emotions shown in the video clips at the NHSTES conference were: “silly,” “anger,” and “sad.” The student was asked to demonstrate any other emotion that she could think of on her own, that had not been mentioned by the GA. In a very interesting and rather poignant clip shown at the NHSTES, she demonstrated “pain” and how she “experienced” that emotion. All of the initial video clips shown at the symposium

show a rather disheveled student who seemed to be in emotional turmoil from unknown causes. After the initial taping, and as a part of the *IPR* supervision, the student was given the videotape of the demonstrated emotions, and asked to take it home, view it, and pick out the portions of most interest to her to share with her supervisor (the PI). She was asked to bring the tape back, prepared to play those portions of the tape of the most interest to her, and to explain to the instructor the reasons for her selections.

This method of “supervision without intervention” appeared useful to the student based on the second set of video clips, which did not show the student’s final meeting one month later with the PI. These tapes show a remarkable change in the student’s demeanor, countenance, presentation, insight, and dress. The student reported that she developed “a whole new way of thinking.” Kagan might surmise that the student, through viewing her early videotapes, saw herself as others see her. She was then able to make the necessary changes that she could not make until she saw what others had been trying to verbally explain. The tangible results for this student were that she was able to resume her placement the next semester, and successfully complete the graduate social work program. She is currently enrolled in a doctoral social work program in another state. *Interpersonal Process Recall* was an integral part of the supervision that helped her successfully complete her graduate education.

Future Use of IPR in Training Child Welfare Workers

The dramatic changes observed in this student suggest that *IPR* may hold great promise for Human Services Trainers in the training of child welfare caseworkers. The relatively nonthreatening nature of the technique is particularly suitable for employees in the early stages of a protective services workplace environment. Video equipment is generally available in protective service settings for use with forensic cases of child abuse. This equipment could be made available for use by supervisors and caseworkers for *IPR* sessions when not in use for forensic purposes. *IPR* could be introduced in the early part of caseworker training, perhaps in peer training for more ready acceptance, then incorporated with supervisors at a later time. *IPR* could also be used for case consultation among peers in unit meetings. *IPR* would be most useful if the changes and growth that occur through

use of the supervision could be measured in order to show the worker's progress.

The changes in the worker might indeed be measurable and thus quantifiable. The *SELF* study (Taylor, 2004) findings showed that those who took *SELF* scored statistically significantly higher on a educational measure of moral/cognitive development—the *SID-IV* (Erwin, 1981)—than a comparison group who graduated from the same university during the same time period who did not take the *SELF* course. Thus, the *SID-IV* (Erwin) may be a useful instrument to measure some aspect of growth in caseworker development.

The *SID-IV* is a 115-item Likert-type scale instrument that has been validated to distinguish the four subscales from the nine original positions of Perry's (1970) Scheme of Intellectual and Ethical Development. In general, the high scores on the *SID-IV* indicate the high moral/cognitive development as measured by Perry's Scheme. Perry's Scheme of Intellectual Ethical Development (1970) is a model of adult cognitive development that describes individual growth through various meaningful stages based upon cognitive process development. Perry identified nine positions, and it is in the seventh through ninth positions that an individual makes lifelong moral and ethical decisions on the basis of their beliefs. Persons at this stage in their lives know that they need values, along with knowledge, to make choices in life. All positions on Perry's Scheme are testable on the *SID-IV*, and a student's/caseworker's progress along the continuum of growth can be followed by having that person take the *SID-IV* periodically. The scores can then be compared against previous scores to ascertain progress. Upon entry into CPS employment, a caseworker's growth could be measured by *SID-IV* at specific intervals. Although the *SID-IV* has been validated as an educational test for moral/cognitive development, it has not been tested for reliability or validity for any other purposes. If its ability to measure the progress of caseworker's personal/professional growth can be proved, the *SID-IV* would be a very useful tool in the training of child welfare caseworkers.

The results of *SELF* study (Taylor, 2004) determined that the highest levels of Perry's Scheme closely fit with the definition of Integrated Personal/Professional Self (IPPS[®]) as defined by the study. As a result, it is possible that the *SID-IV* will be validated at

a later date for use outside of the educational field in social work to test for IPPS[®]. The *SELF* study (Taylor), also found that the students who took the *SELF* course that scored high on the *SID-IV* tended to be former Title IV-E students. These workers also had high retention rates in child welfare work in their first five years post graduation. Therefore there may be correlation between a high score on Perry's Scheme (moral/cognitive development and high IPPS[®]) and remaining in child welfare work. This was verified by recent Texas statewide Title IV-E data (Title IV-E data, 2006). The sample size was small and the findings were beyond the scope of the study.

If this correlation could be established, the *SID-IV* could be used as a tool to screen child welfare caseworker applicants for retention potential. The *SID-IV* could establish a cut-off which would determine which applicants would be encouraged or discouraged to pursue CPS employment. The cut-off would be established over several years by workers' scores and retention rates. Such a screening instrument would prove invaluable by allowing states to spend resources for only those new employees most likely to be retained as long-term child welfare caseworkers. A great deal more study is needed, but such a promising potential is worth pursuit. A pilot study could be undertaken in which the *SID-IV* would be given to incoming Basic Skills Development classes. The *SID-IV* scores that new employees get could be kept in employee folders and correlated only upon the employee's departure from the department to see if there was any correlation between the score and longevity on the job. Such a pilot study would be inexpensive and nonthreatening since the instrument can be purchased for less than \$3.00 apiece and would only be used for purposes of correlation with retention rates.

Conclusion

Interpersonal Process Recall may be a useful supervisory tool for the training of human services workers, if it could be implemented early in the training in a nonthreatening manner. The *SID-IV* would be a very valuable tool to show worker growth if it could be validated to show growth in integrated personal/professional self. The *SID-IV* should be considered for use as a screening instrument for use with potential child welfare workers for purposes of retention.

References

- Acker, G. (1999). The impact of clients' mental illness on social workers' job satisfaction and burnout. *Health and Social Work, 24*(2), 112-119.
- Bernard, J.M., & Goodyear, R.K. (1992). *Fundamentals of clinical supervision*. Needham Heights, MA: Allyn and Bacon.
- Dane, B. (2000). Child welfare workers: An innovative approach for interacting with secondary trauma. *Journal of Social Work Education, 36*(1), 27-41.
- Erwin, T.D. (1981). *Scale of intellectual development*. Harrisonburg, VA: Developmental Analytics.
- Goldstein, H. (2001). *Experiential learning*. Alexandria, VA: Council on Social Work Education.
- Grossbard, H. (1954). Methodology for developing self-awareness. *Social Casework, 35*, 380-386.
- Kagan, S. (1992). *Cooperative learning*. San Juan Capistrano, CA: Kagan Cooperative Learning.
- Kagen, N., Krathwohl, D.R., Goldberg, A.D., & Campbell, R. (1967). *Studies in human interaction: Interpersonal process recall stimulated by videotape*. East Lansing, Michigan: Michigan State University.
- Kingdon, M.A. (1975). A cost/benefit analysis of the interpersonal process recall technique. *Journal of Counseling Psychology, 22*, 353-357.
- Kondrat, M.E. (1999). Who is the "self" in self-aware: Professional self-awareness from a critical theory perspective. *Social Service Review, December*, 451-475.
- Maslach, C., & Jackson, S.E. (1986). *MBI Human Services*. Palo Alto, CA.: Consulting Psychologists Press, Inc.
- Perry, W.G. (1970). *Forms of intellectual and ethncial development in the college years: A scheme*. New York: Holt, Rinehart & Winston.
- Taylor, P.G. (2004). *Enhancement of integrated personal/professional self (IPPS) through reflective/experiential learning in graduate social work s students (Doctoral dissertation, University of Houston, 2004)*. *Dissertation Abstracts International, 65*, 1123.
- Title IV-E statewide data base. (2006). Austin, Texas. Unpublished raw data.

	Discussion: Using Interpersonal Process Recall (IPR)

Theme: Can Child Welfare Workers Benefit from Using IPR?

Title: Using Interpersonal Response Program

Presenter: Patricia Taylor, *Ph.D., LCSW*

Facilitator: Marcia Sanderson, *LCSW*

Topics of Discussion

Dr. Taylor presented how she has used the Interpersonal Process Recall (IPR) with MSW students enrolled in a SELF course. Through the IPR, students gained insight into their biases and discomforts connected to their client population. The insight is credited with positively influencing how they worked with clients. She also explained how the Scale of Intellectual Development (SID IV) has been used to measure student growth and development while emphasizing that it is not connected to the IPR. It is theorized that student SID IV scores can predict worker retention rates, although more study is needed to establish this.

I. How Students Have Benefited from IPR?

- As a result of seeing themselves on video, students perceive themselves as the cause for the changes they make. New child welfare workers report that the videos provide them with insight into their biases, values, and discomforts. They highlight the experience of hearing it come from their own mouths as the catalyst for changes.
- Through this insight, IPR promotes skill development as students identify areas where they need to improve. The self-directed “working on the self” then influences how students work with clients.
- IPR has the potential to prepare new child welfare workers to face emotions and triggers that they will experience upon entering the field.
- Some students have not immediately embraced IPR for fear of breaches in confidentiality. However, these same

- students benefited from it after feeling secure that confidentiality would be maintained.
- IPR may potentially correct cultural miscommunications such as speech patterns and body movements. Although, students have not reported this, a supervisor who often detects burnout in her employees was unable to detect her own burnout symptoms until watching herself on video. This implies that non verbal communication can be addressed.
- There are no current examples of students failing to benefit from IPR. Out of a group of 16 students, four chose not to give the videos to the facilitators because they were uncomfortable being evaluated. However, these four students also reported to have benefited from the video by noticing areas where they could improve.
- A participant stated they she used IPR during the 1970s to develop skills for intervention with clients.

II. Concerns about IPR as a Tool for Skill Development

- The self-directed component, while providing the student with insight, may not provide essential information or skills that a child welfare worker will need upon entering the workforce.
- Students may be uncomfortable with supervisors watching their videos. Students may feel that they are being evaluated instead of using the videos for self-directed professional growth. Workers' perceptions of how their supervisors see them influence their self-efficacy. Thus, workers/students may be uncomfortable with their supervisors watching videos of them at vulnerable moments. It was proposed that peer evaluation, during the initial stages, may reduce discomfort level and prepare students to use IPR with supervisors in the future.
- In Texas, the family-based safety workers complete a basic skills development training consisting of six weeks of classroom training, including a group component that explores cultural biases. Breaches in confidentiality and lack of skill in addressing sensitive topics created the potential for harm being done to clients. Workers entering the field without addressing these topics may do more harm than good when working with clients.

- For the IPR to serve as an indicator of change, more subjective and systematic evaluation is necessary.
- Use of the IPR needs to focus on supporting workers in serving clients better through improving practice-based outcomes. Videotaping simulations between workers and clients can also help workers identify interpersonal values and skills. A videotape followup can measure change in the worker.
- IPR appears to be similar to another model, SPIN USA, which has been used with trainers involved in Family to Family and TDM. Trainers use the videos to target certain behaviors/characteristics that they want improve for supervisory development. They give the videos to facilitators who provide strengths-based feedback which reduces anxiety connected to evaluation. However, SPIN USA does not use a measurement tool to evaluate the impact of the intervention

III. Using the SID IV as a Measurement Instrument

- The SID IV is the instrument to measure growth and development while the IPR is a supervisory technique. They are not connected. Students' pre- and post-SID IV test scores have recently been used to measure student growth.
- Currently, SID IV has been used successfully to measure the cognitive and moral development of workers in medical and educational settings. It is hypothesized that it can be effectively used to predict retention among child welfare workers, but current information is unavailable. Currently, students with high SID IV scores demonstrate high retention; however, the connection has not been proven. The presenter also asked if other measurement tools existed that have been used in child welfare.
- SID IV appears susceptible to social desirability in the student responses, which negatively influences validity. Because a component of the changes students make applies to their reaction to serving clients living in poverty, the "Attitudes towards poverty" scale may apply because it measures an attitudinal shift as workers encounter biases working with poor clients.

Discussion

- Before deciding on a measurement tool, it must be decided what the intervention is supposed to measure (i.e., attitude, personality, or behavior shift). Once that is decided, a measurement tool can be chosen.

	Evaluating Values and Attitudes: Exploring the Impact of Training on Disproportionality and Disparity

Barrett Johnson, *LCSW*, Leslie W. Zeitler, *LCSW*, and Chris Mathias, *MSW*

Introduction

The advent of transparent, publicly available data on the composition and outcomes of the child welfare system has focused attention on differences related to race and ethnicity. Two characteristics become clear in reviewing the national data. First, Black and Native American children are overrepresented in the system relative to their proportion of the population—this is commonly referred to as “disproportionality.” Second, outcomes differ (and are worse) for Black and Native American children—this is referred to as “disparity.” When looking at state and local data, the relative level of disproportionality and disparity varies. In some jurisdictions, other groups, such as Latinos, experience disproportionality and disparity.

As the child welfare system grapples with this issue, the training community frequently becomes involved. The California Social Work Education Center (CalSWEC) at the University of California Berkeley School of Social Welfare has sponsored several statewide symposia to analyze different approaches to addressing the problem, as well as to distinguish the role of the in-service training system in supporting a solution. The training community has also responded by increasing efforts to infuse discussion of racism, ethnocentrism, and cultural difference in curricula throughout the state. Much of the training involves examination of conscious and unconscious bias by trainees in the way that they assess and interact with families. Problems arise, however, when attempting to evaluate the effectiveness of the training interventions. Since race and ethnicity are highly

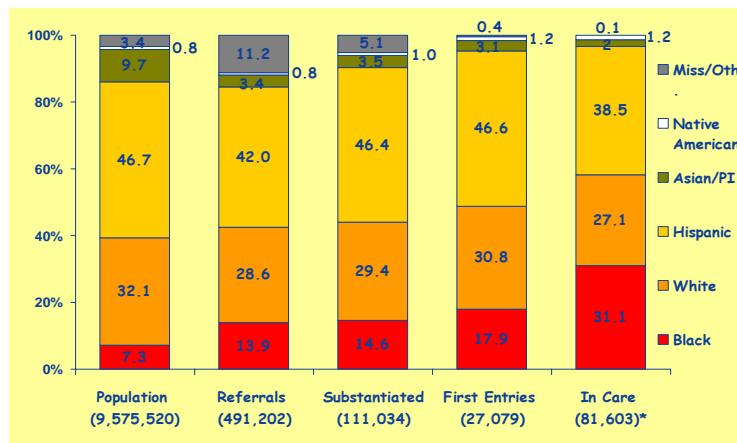
value-laden and emotional topics, standard training evaluation methods become complicated. How, for instance, does one determine whether the training made an impact on biases within the trainees? For that matter, how does one measure value- and attitude-laden constructs such as bias and ethnocentrism in the first place? This paper will outline CalSWEC’s preliminary efforts to improve the way we evaluate training in such value-laden content.

Context

Disproportionality in California

Before discussing training on disproportionality and disparity, one must first become more familiar with the demographic data. In California, the problem of disproportionality is best illustrated by Figure 1, which compares the composition of California’s child welfare population in 2004 at different stages of intervention with the composition of the state’s child population.

Figure 1. Ethnicity and the Path Through the California Child Welfare System—2004



* Data includes 18 year olds

Chart taken from presentation by Barbara Needell, CSSR UC Berkeley, *Race/Ethnic Disproportionality and Disparity in Child Welfare: New Views, New Measures* at May 2006 Family to Family Conference.

In this histogram format, one can see that Black children make up 7.3% of the child population in California. White children make up 32.1%, Hispanic/Latino children 46.7%, Asian/Pacific Islander children 9.7%, and Native American children 0.8%. However,

Black children represent 13% of all child welfare referrals, 15% of substantiated referrals, 18% of first entries into care, and 31% of children in care. National incidence data show no difference in the incidence of abuse and neglect between Black families and other groups (Hill, 2006). Native American children are also disproportionately represented when compared with their percentage of California's total population at some stages, though the effects are harder to see graphically because of their smaller total numbers. They make up 0.8% of referrals, 1.0% of substantiations, 1.2% of first entries into foster care, and 1.2% of children in care (all data and graphic are from Needell et al., 2006).

Literature on Disproportionality

A substantial body of research (both within California and nationwide) supports the differential impact of race and ethnicity at different intervention points in the child welfare system. This is summarized thoroughly in Robert Hill's report for the Casey Center for the Study of Social Policy Alliance for Racial Equity in the Child Welfare System (Hill, 2006). The highlights of the research are noted here to contextualize the data in Figure 1.

Data on reporting to the child welfare agency is often difficult to analyze, since reporting relates a great deal to factors associated with poverty, such as substance abuse, domestic violence, etc. (Hill, 2006). Moreover, reporting is the point in the system where the child welfare agency has arguably the least impact, since it only indirectly impacts those who make reports to the child abuse hotline or screening center. Nevertheless, a variety of studies have shown that Black families are more likely to be reported for abuse than white families with similar sets of injuries to their children. (Hampton & Newberger, 1985; Lane et al., 2002; Jenny, Hymel, Ritzen et al., 1999). While some studies suggest race is not related to referral, Hill concludes that "most research studies . . . have found race to be an important factor in submitting reports to CPS hotlines" (Hill, 2006).

Hill comes to a similar conclusion on the rates of substantiation of abuse once allegations are assessed by the child welfare system. He indicates that almost all of the known studies found racial differences in substantiation of reports. In several studies, race continued to be a factor even when controlling for other factors (Ards et al., 2003; Fluke et al., 2003). Decision-making and bias is more obviously related to substantiation of these reports because

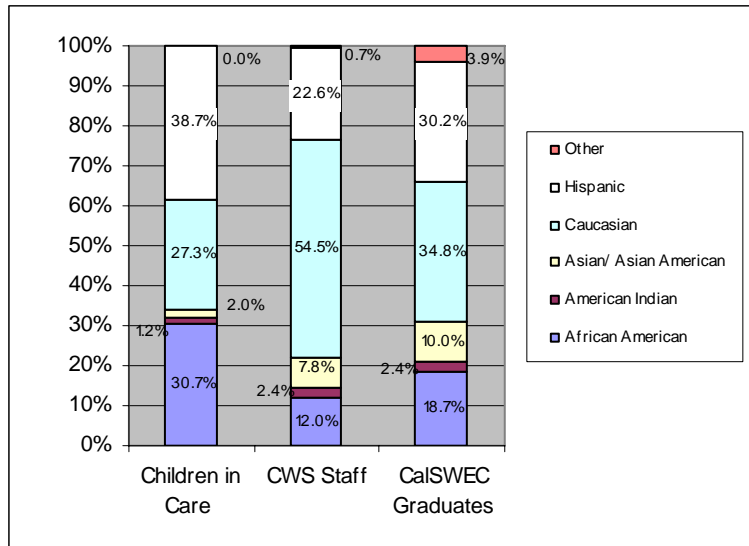
the child welfare agency makes the decision whether a report is substantiated or not.

The decision to place in foster care or not is similarly vulnerable to biases in decision-making, since the child welfare agency completes the assessment and makes recommendations to the court to initiate placement. Black children are also more likely to be placed in foster care, even when compared with white children with similar family characteristics (U.S. Children's Bureau, 1997). Again, Hill concludes that—despite some conflicting studies—race is an important factor in deciding to place children in care. Black children are also more likely to remain in care longer, though this may be impacted by higher levels of kinship care by extended Black family members.

The Child Welfare Workforce

The composition of the child welfare workforce has generally been seen as an indicator of the cultural sensitivity of the system, under the logic that a workforce that reflects the population will display less bias in their decision-making. Little research exists in this area for child welfare services; the literature on ethnic matching emerges largely from the mental health field and therefore may have limited applicability. The findings in mental health are also mixed (Clark & Jacquet, 2002). It is also assumed more broadly that a more educated workforce will be able to examine families' situation in a more empirical, less biased manner, though this assumption is also not tested extensively. CalSWEC conducts a study of California's child welfare workforce approximately every three years as part of the evaluation of the Title IV-E stipend program, which provides financial assistance and specialized curricula to masters- and bachelors-level students preparing to work in public child welfare. Figure 2 shows the racial/ethnic composition of the workforce in California according to the 2004 Workforce Study (Clark et al., 2005), compared with the percentage of children in care and the percentage of Title IV-E recipients since 1992.

Figure 2. Ethnicity of Children in Care, CWW Staff, and CalSWEC Graduates (Title IV-E Stipend Recipients)



Again in the histogram format, an exact match of the workforce with the children in care would make the bars identical. CalSWEC’s stipend program appears to reflect the population more closely—though Blacks and Latino/Hispanics remain under-represented, and Whites and Asian/Pacific Islanders are over-represented. Other data from earlier workforce studies show that the percentage of African Americans working in child welfare who have an MSW has remained constant, while the percentage for Latinos has increased. During the same time period, the percentage of child welfare workers who possess a MSW has increased from 21% to 36%. (Clark & Fulcher, 2005)

Responses to the Data

Historically, it has been difficult to address the problem of disproportionality and disparity within the child welfare system, since Black and Native American children also experience disproportional representation and disparate outcomes in a variety of other systems, such as the welfare system, the health/mental health system, and the juvenile justice system. The tendency has been to blame the failures of these external systems, while characterizing the child welfare system as a reflection of these other failures. To counter this tendency, California has focused

attention on analysis of the data that demonstrates the role of the child welfare system. The histogram in Figure 1, for example, is used widely in common core curriculum for new line workers and supervisors (CalSWEC, 2005) in order to promote the concept that biased decision-making within the system is an issue. Barbara Needell of UC Berkeley's Center for Social Work Research has also presented the data to a variety of audiences, making the point that a non-biased system might take in more referrals due to outside influences, but should be able to assess families adequately enough to bring the proportion of children in care down to a level more congruent to the proportion in the population. Once children are in care, then the system surely has a responsibility to move them to permanency in an equal manner. Such an analysis promotes action rather than paralysis: while the child welfare system is not responsible for eliminating disparities in other systems, it *can* focus on its own biases and disparities.

The Role of Training Evaluation

Given the likely role of conscious and unconscious bias in producing some of the outcomes outlined above, and the mismatch of the workforce with the population served, California's training system strives to promote more understanding of the problem by infusing curricula with activities and text aimed at promoting fairness and equity. (The terms *fairness and equity* were initially delineated by the California Stakeholders Group as the common language with which to discuss all issues of disproportionality and disparity (California Department of Social Services, 2002).) CalSWEC, the Child and Family Policy Institute of California, and the Regional Training Academies sponsored and participated in a series of annual symposia on the topic, with separate events aimed at county and state leadership as well as the training community. Some success was achieved in infusing curricula with knowledge, skill, and values learning objectives that reflected fairness and equity issues. Training focuses on identifying and examining possible biases in decision-making. Individual child welfare workers are encouraged to identify and acknowledge unconscious biases that they (and anyone else) may have when making decisions about families. It is hoped that the focus on the unconsciousness and universality of the bias will make the trainees more open to examining their own role in the system's outcomes.

Since California also attempts to evaluate the effectiveness of training interventions, several issues emerged: if the goal of the curriculum infusion was to promote attitude change, how exactly could such a change be measured? More basically, what are the key values and attitudes that underlie culturally sensitive, unbiased practice, and how might one measure whether a particular training had an impact? Answering these questions would help define the role of training in changing attitudes.

Evaluation Issues at the Levels of Learning

Measurable outcomes are needed to effectively determine if the information a trainee receives during training promotes behavior changes and results in positive results for families and children. Since many confounding variables can influence outcomes in child welfare, directly connecting training to outcomes is not usually possible. California's training evaluation framework is therefore based on the notion of levels of training evaluation that build on one another to form a "chain of evidence" for the effectiveness of the program as a whole, as well as evidence of the effectiveness of particular courses (Parry, Berdie, & Johnson, 2004). The chain of evidence refers to establishing a linkage between training and desired outcomes for the participant, the agency, and the client such that a reasonable person would agree that training played a part in producing that outcome (Parry & Johnson, 2005). In this case, the outcome would be decreased disproportionality and disparity for Black children and families. California adopted a *Framework for Training Evaluation* in 2005 as a basic structure and strategic plan for evaluating training. It also forms the basis for the chain of evidence. (CalSWEC, 2005). By reviewing the issues that emerge at these levels, we can begin to strategize about how to measure the impact of training on attitudes, biases and behaviors.

Level One—Tracking Training. This most basic level of evaluation means keeping track of who attends training in an organization, often to determine the rate at which staff have been exposed to the desired behaviors, attitudes, and values contained in the training. While this seems a straightforward prospect, several issues emerge specific to evaluating values and attitudes. Most basically, the organizational leadership needs to determine whether the training is optional or mandated. In values-laden training, this is particularly important due to the controversial nature of the typical content. Since the training may challenge conscious and unconscious core beliefs, those who hold undesirable attitudes and

beliefs are likely to avoid such a challenge. Unrepentant homophobes, for example, are unlikely to attend a training that promotes equality and respect for gay, lesbian, bisexual, and transgender people. More significantly, the charged nature of race relations makes many well-meaning people uncomfortable discussing their potential biases. An ability to significantly penetrate the organization with this type of training is therefore of paramount importance; without assuring that those who need the training get it, there is unlikely to be change at any subsequent level.

Level Two—Course Evaluation. This level includes assuring that content on disproportionality and disparity (as well as conscious and unconscious bias) is included in the content of training. This type of content has already been integrated into California line worker core training curricula, so all new line workers receive at least an introduction to the concepts. At this level alone, however, one cannot evaluate whether or not a change in course content or methodology is actually correlated with a change in behavior, or with a change in outcomes for children and families. The main issue here is inclusion of content that is consistent across different curricula, with examples of potential biases in decision-making in different situations and cases.

Level Three—Trainee Satisfaction and Opinion. Trainee satisfaction/opinion provides a snapshot of how the trainee perceived the training, and whether they liked it or not. There are several limitations specific to evaluating attitudes and behaviors at this level. First, trainees' positive responses to a satisfaction survey may decrease as a result of being challenged on issues like racism and ethnocentrism. Uncovering personal biases can be uncomfortable and unpleasant work for trainees. Trainees' responses on an opinion survey of attitude change (e.g., Likert-type scale, possibly pre- and post-) may reflect social desirability factors. Finally, one cannot evaluate whether or not a change in opinion is actually correlated with a change in knowledge or behavior, so the link to outcomes for children and families is tenuous at best. While one could make the case that trainees that were satisfied and comfortable in the training room were more amenable to changing their pre-conceived biases, one could equally assert the opposite. Sometimes people change when they are most exposed and most uncomfortable.

Level Four—Knowledge Acquisition. Depending on the course competencies and knowledge learning objectives, a change in knowledge should provide the foundation and context for the need for further skill development. Knowledge of the problems of disproportionality and disparity, for example, may lead to more critical thinking about a caseworker's own decision-making processes. In California, knowledge tests in the common core could, for example, ask simple questions about the definitions of key terms, and assess whether trainees knew which groups were disproportionately represented. Such content and testing requires clear delineation of knowledge that is necessary to promote attitude and value change. Given the differences between the California workforce and the population, key pieces of knowledge might conceivably include some perspective on the history of race relations in America, and some information on the day-to-day experience of racism and discrimination. The relationship between this knowledge and behavior or attitude change is not clear.

Level Five—Skill Acquisition (as measured in the classroom). The skill level presents some possibilities in terms of evaluation, since changes in attitudes would need to manifest themselves in some sort of change in behaviors or skills in order to affect outcomes. At this level alone (or even combined with knowledge and satisfaction), one cannot evaluate whether or not a change in skill is actually correlated with a change in outcomes for children and families—but one can see whether the trainees are capable of desired behavior when they leave the classroom. Trainees may be able to demonstrate, for example, methods of non-biased communication in the classroom. Again, delineating the specific behaviors related to attitudes and values is critical—if behaviors are described and observed clearly, then measuring them becomes much easier.

Level Six—Transfer of Learning (as measured in the workplace). Evaluating the transfer of skills to the workforce is closest to evaluating the impact on outcomes. If behaviors are delineated well, and linked to specific outcomes related to disproportionality and disparity, then observing behavior change in the workplace with actual families and children will show most clearly the impact of training on the outcomes. Again, specific behaviors must be determined, and these behaviors must be shown to affect decision-making such that the outcomes of cases change.

For example, if one can observe a trainee using non-biased communication skills with one of their families, then we establish that they are applying the skill to their everyday work life. It is still unknown, however, whether this communication changes the path of that child or family through the child welfare system.

At this level, the issue of authenticity in the helping relationship also emerges. Training the workforce to change their behavior alone may not lead to more engagement with clients. Discussions of cultural difference can be fraught with emotion, and assuring that they have an impact may depend on an authentic, empathic relationship between the social worker and the clients. Behaviors and techniques that are rehearsed and practiced in the training room may not appear authentic to clients if there is dissonance between these behaviors and the underlying attitude of the social worker. For example, an overtly racist person can mimic behavior change, but their racism is likely to still be perceived in the context of a conversation with a member of the group that they discriminate against.

Level Seven—Outcomes. One can argue that linking the changes in knowledge and behavior to changes in outcomes is the aim of the entire enterprise of training evaluation. In order to do this, the training community and the evaluators must assure that they are engaged with the system sufficiently that everyone knows the outcomes that are being measured. In this case, a particular jurisdiction could aim to make their caseload at different points look more like the population they serve. Some caution is warranted, however, in that the data is sometimes difficult to decipher—for any given jurisdiction, demographic factors (such as extreme concentrations of poverty in one population, or an epidemic of methamphetamine that affects one community more than another) must be considered.

Literature on Assessing Attitude/Value Change

Little research is available on the impact of child welfare training on attitudes and values, though some limited work has been done in the area of mental health, education, and medicine. The mental health/counseling field has perhaps the most relevant set of tools and research, particularly with regard to measurement instruments. Several cultural competency instruments stand out: the Multicultural Awareness Knowledge Skills Survey (MAKSS)

by D'Andrea et al. (1991); the Multicultural Counseling Awareness Scale (MCAS-B) by Ponterotto, et al. (1991, revised in 2003 to the Multicultural Counseling Knowledge and Awareness Scale); the Multicultural Counseling Inventory (MCI) by Sadowsky, et al., (1994); the Multicultural Competency and Training Survey (MCCTS) by Holcomb-McCoy (1999); and the Cross Cultural Counseling Inventory, Revised (CCCI-R) by LaFramboise, et al. (1991). Most of these measures use the cross-cultural competencies framework set out by Sue et al. in 1982, which posits four dimensions to cultural competence: beliefs/attitudes, knowledge and skills.

While some studies of these measures have attempted to assess their validity and reliability (Ponterotto, Rieger, Barrett, & Sparks, 1994; Pope-Davis, 1994), much more research needs to be done to ascertain whether they are rigorous enough to assess such competencies accurately. Gamst et al. (2004) used a series of sub-studies to create the California Brief Multi-Cultural Competency Scale (CBMCS), which synthesized items from four instruments (CCCI-R, MAKSS, MCAS-B and MCCTS) in an attempt to create one brief instrument that could be used in subsequent research. The CBMCS attempts to address the issue of social desirability by evaluating whether particular items are associated with an established social desirability measure, and eliminating them if they do. While such research is welcomed to move toward more empirical study of attitudes and beliefs about culture, numerous limiting factors are acknowledged by the authors. These include the use of convenience samples, and research that suggests a limited relationship between self-reported cultural competence and observed cultural competence (Constantine & Ladany, 2002; Worthington, Mobley, Franks, & Tan, 2000). All the measures, with the exception of the CCCR, use self-report Likert-scale questions rather than observation of skills or knowledge testing. They also do not necessarily apply to the public child welfare workforce, since they were primarily designed to assess mental health practitioners.

Operationalizing Competencies

One strength of the above instruments is their reliance on a single theoretical framework of competencies, developed as a position paper by Sue et al. in 1982. By using one framework,

these instruments were attempting to operationalize and measure similar competencies. This meant that they could be compared, and even synthesized, as they developed and more research was completed. Such an approach is advisable as specific measures (and resulting instruments) are developed related to quantify change in attitudes or values for child welfare practice. In California's child welfare system, this is made easier by the fact that standardized competencies have already been implemented statewide—the CalSWEC Title IV-E and In-Service Competencies. By operationalizing these competencies, we can take the first steps to systematically measuring changes in attitudes and values. Since the training curriculum is already based on the competencies, we also have an opportunity to test different measures with new child welfare worker trainees. Looking at how (and whether) particular competencies can be operationalized will also assist in revising the competencies to make them more measurable and precise.

In order to begin the work of operationalizing the CalSWEC competencies, CalSWEC conducted an introductory exercise as part of the 2006 Symposium on Fairness and Equity Issues in Child Welfare Training. Participants comprise trainers and staff development personnel involved with the Child Welfare System from around the state.

Using a focus-group method, participants were asked to operationalize knowledge and skills for curriculum competencies, by delineating the specific pieces of knowledge and skill that would underlie such a competency. Questions used for each competency related to culture included: What specific pieces of knowledge undergird this competency? (*Prompt: Think about items that have right and wrong answers on the paper test.*); and What behaviors would demonstrate integration of this competency? (*Prompt: Think about behaviors that you could observe and document.*) Participants were also invited to formulate some specific, measurable learning objectives related to knowledge and skill that were identified. Responses are summarized for one competency in Figure 3.

Figure 3. Operationalizing Competencies; Focus Group Responses

Competency 1.2: *Student demonstrates the ability to conduct an ethnically and culturally sensitive assessment of a child and family and to develop an appropriate intervention plan.*

What specific pieces of knowledge undergird this competency?:

- Definitions of culture, ethnicity, race, bias, prejudice, discrimination, stereotype.
- Knowledge of models of identity development—e.g., minority identity development and white identity development;
- Knowledge of the impact of culture, immigration, and history of different groups on help-seeking behaviors;
- Knowledge of differences in communication styles—verbal & nonverbal—of five major ethnic/racial groups in U.S.

What behaviors would demonstrate integration of this competency?

- Behaviors that demonstrated respect, sensitivity, open-mindedness, willingness to learn, non-defensive communication, tolerance for ambiguity, and diminished defensiveness.

Despite prompts to be as specific as possible, the identified behaviors tended to be quite general and non-measurable. It is possible, however, to conceptualize some of the specific behaviors. With regard to open-mindedness, for example, one could measure the ability of the trainee to ask follow-up questions that were related to culture, or to ask specific questions about how a family's cultural beliefs impacted their decisions in parenting. As a follow-up activity, CalSWEC will conduct more time-intensive focus groups across the state in order to more fully explicate the desired behaviors. Data from these focus groups will be used to develop measures and instruments to analyze attitude and value change at different levels of the learning process.

Conclusion

Obviously, the work of empirically measuring the impact of training on changes in attitudes and values requires more

development. There is little literature that is directly applicable to child welfare practice, particularly in the area of training evaluation. An obvious next step would be developing and testing key definitions. To do this, competencies must be fully delineated, measures developed, and instruments piloted and validated. A look at the different levels of training evaluation sheds light on many of the key areas that must be addressed in order to move toward a better understanding of the role of training. There is work to do at each level: we must consider everything from who attends training to the authenticity of the trainees when they practice the skills that they attain in the classroom. This may appear daunting. But all of the necessary behaviors and skills that we are attempting to measure are fundamental to basic social work practice. Many of them (unbiased communication, expression of empathy) form the building blocks of the relationships between social workers and those that they aim to help. Measuring them effectively is therefore a high priority for training and program evaluators throughout the human services.

References

- Ards, S., Myers, S., Malkis, A., Sugrue, E., & Zhou, L. (2003). Racial disproportionality in reported and substantiated child abuse and neglect: An examination of systemic bias. *Children and Youth Services Review, 25*(5/6): 375-392.
- California Social Work Education Center. (2005). The California Common Core Curriculum for Child Welfare Workers and Supervisors. Berkeley, CA. See <http://calswec.berkeley.edu/CalSWEC/CommonCoreCurricCA.html>
- California Welfare Services Stakeholders Group. (2002). Building a Culture of Fairness and Equity in California's Child Welfare System. Sacramento, CA.
- Clark, S., & Jacquet, S. (2002). Affecting recruitment and retention of students of color preparing to work in public child welfare. Berkeley, CA: California Social Work Education Center.
- Clark, S., & Fulcher, G. (2005). The 2004 California public child welfare workforce study. Berkeley, CA: California Social Work Education Center. See http://calswec.berkeley.edu/CalSWEC/WorkforceStudy_2004.pdf
- Constantine, M.G., & Ladany, N. (2000). Self-report multicultural counseling competence scales: Their relation to social desirability attitudes and multicultural case conceptualization ability. *Journal of Counseling Psychology, 47*(2), 155-164.
- D'Andrea, M., Daniels, J., & Heck, R. (1991). Evaluating the impact of multicultural counseling training. *Journal of Counseling & Development 70*, 143-150.
- Fluke, J., Yuan, Y., Hedderson, J., & Curtis, P. (2003). Disproportionate representation of race and ethnicity in child maltreatment: Investigation and victimization. *Children and Youth Services Review, 25*(5/6), 359-373.
- Gamst, G., Dana, R.H., Der-Karabetian, A., Aragon, M., Arellano, L., Morrow, G., & Martenson, L. (2004). Cultural competency revised: The California Brief Multicultural Competence Scale. *Measurement and Evaluation in Counseling and Development, 37*, 163-183.
- Hampton, R.L., & Newberger, E. (1985). Child abuse incidence and reporting by hospitals: Significance of severity, class and race. *American Journal of Public Health, 75*(1), 56-60.

- Hill, R.B. (2006). Synthesis of research on disproportionality in child welfare: An update. Washington D.C.: The Case-CSSP Alliance for Racial Equity. See http://www.cssp.org/major_initiatives/racialEquity
- Holcomb-McCoy, C.C. (2000). Multicultural counseling competencies: An exploratory factor analysis. *Journal of Multicultural Counseling and Development, 28*, 83-97.
- Holcomb-McCoy, C.C., & Myers, J.E. (1999). Multicultural competence and counselor training: A national survey. *Journal of Counseling & Development, 77*, 294-302.
- Jenny, C., Hymel, K., Ritzen, A., Reinert, S., & Hay, T. (1999). Analysis of missed cases of abusive head trauma. *Journal of the American Medical Association, 281*(7), 621-626.
- LaFromboise, T., Coleman, L.K., & Hernandez, A. (1991). Development and factor structure of the cross-cultural counseling inventory—Revised. *Professional Psychology, Research and Practice, 22*(5), 380.
- Lane, W., Rubin, D., Monteith, R., & Christian, C. (2002). Racial differences in the evaluation of pediatric fractures for physical abuse. *Journal of the American Medical Association, 288*(13), 1603-1609.
- Needell, B. (2006, April). Race/ethnic disproportionality and disparity in child welfare: New views, new measures. Symposium on Fairness and Equity Issues in Child Welfare Training. Berkeley, CA.
- Needell, B., Webster, D., Armijo, M., Lee, S., Cuccaro-Alamin, S., Shaw, T., Dawson, W., Piccus, W., Magruder, J., Exel, M., Conley, A., Smith, J., Dunn, A., Frerer, K., & Putnam Hornstein, E. (2006). Child Welfare Services Reports for California. University of California at Berkeley Center for Social Services Research. See: <http://cssr.berkeley.edu/CWSCMSreports/>
- Parry, C., & Berdie, J. (2004). Training Evaluation Framework Report. Berkeley, CA: California Social Work Education Center.
- Parry, C., & Johnson, B. (2005). Training evaluation in a large system: California's framework for child welfare training evaluation. In B. Johnson, M. Henderson, & M. Thibedeau (Eds.), *Proceedings of the 8th Annual Human Services Training*

- Evaluation Symposium (pp. 15-31). Berkeley, California: California Social Work Education Center.
- Ponterotto, J.G., Gretchen, D., Utsey, S.O., Rieger, B.P., & Austin, R. (2002). A revision of the multicultural counseling awareness scale. *Journal of Multicultural Counseling and Development, 30*(3), 153.
- Ponterotto, J.G., Rieger, B.P., Barrett, A., & Sparks, R. (1994). Assessing multicultural counseling competence: A review of instrumentation. *Journal of Counseling & Development, 72*(3), 316-322.
- Pope-Davis, D.B. (1994). An empirical comparison of two self-report multicultural counseling competency inventories. *Measurement and Evaluation in Counseling and Development, 27*, 93.
- Sodowsky, G., Taffe, R.C., Gutkin, T.B., & Wise, S L. (1994). Development of the multicultural counseling inventory: A self-report measure of multicultural competencies. *Journal of Counseling Psychology, 41*(2), 137-148.
- Sue, D.W., Bernier, J.B., Durran, M., Feinberg, L., Pedersen, P., Smith, E., & Vasquez-Nuttall, E. (1982). Position paper: Cross-cultural counseling competencies. *The Counseling Psychologist, 10*, 45-52.
- U.S. Department of Health and Human Services, Children's Bureau. (1997). National study of protective, preventive and reunification services delivered to children and their families. Washington, D.C: U.S. Government Printing Office.
- Worthington, R.L., Mobley, M., Franks, R.P., & Tan, J.A. (2000). Multicultural counseling competencies: Verbal content, counselor attributions, and social desirability. *Journal of Counseling Psychology, 47*(4), 460-468.

	Discussion: How Do We Evaluate Assumptions, Attitudes, and Values?

Theme: Evaluating Assumptions, Attitudes, and Values

Title: How Do We Evaluate Assumptions, Attitudes, and Values?

Presenters: Leslie W. Zeitler, *LCSW*, and Chris Mathias, *M.S.W.*

Facilitator: Barrett Johnson, *LCSW*

Topics of Discussion

Ms. Zeitler and Ms. Mathias presented statistics demonstrating racial disproportionality in California child welfare. The statistics also suggest that worker bias plays some role, even if it is unconscious bias. Thus, influencing child welfare workers' attitudes, beliefs, and values through training curriculum is targeted as a way of reducing disproportionality. The presenters introduced the literature pertaining to ethnic matching of clients to workers and the challenges of operationalizing learning objectives to verify that workers had adopted the attitudes, beliefs, and values necessary to improve practice methods.

I. Is Disproportionality a Problem?

- How data is presented and understood determines whether it is viewed as a problem or not. Therefore, data on disproportionality research must be interpreted cautiously before determining whether it is a symptom of institutional racism. Was socioeconomic status controlled for? Is the data reflective of the condition of our communities and families?
- When SES is controlled for, the role of child welfare in perpetuating disproportionality is apparent. If the child welfare system was not biased, the rates of African American and Native American children in care would decrease to levels proportional to their actual population makeup. This is supported by research demonstrating that

these population groups do not abuse their children at higher rates than other ethnicities.

- A helpful statistic would be the proportion of excess children in care. Quantifiable methods could be used to know the family circumstance of children who enter care and those who do not enter care. This method could look at children of families from similar conditions and why some groups of children enter care and others do not. This would provide a measure of excess children who do not belong in care.
- Client class should also be assessed in addition to ethnicity. Workers can come from the same ethnic background as clients and still be prejudiced toward them because of their lower class standing. This topic is not currently addressed and needs to be.

II. Can People Learn Attitudes, Values, and Beliefs?

- Nebraska child welfare is implementing family-centered practice, a new system which emphasizes attitudes, values, and beliefs. A job analysis rates when these attitudes, values, and beliefs are needed as well as how and when they can be learned in training. However, according to the analysis, employees should already have these qualities instead of learning them through training.
 - Nebraska child welfare is using this information to hire new employees and must figure out ways to measure these competencies (cultural sensitivity, demonstrate respect).
 - The challenge is to determine how to measure these competencies as well as to determine which ones can be learned and which ones an employee must already have.
- Attitudes, values, and beliefs should be instilled during social work education. However, it is unknown if MSW students graduate with them. Schools of social work need to evaluate if graduates are entering the workforce with these attitudes, values, and beliefs. If it can be determined that social work education is preparing workers, then child welfare agencies will have reason to hire MSWs in order to combat disproportionality. The case can also be made for

agencies to support their workers' continued education in order to acquire these attitudes, values, and beliefs.

- Northern California Training Academy discusses the topic of institutional racism during trainings for supervisors.
 - One finding has been that workers may unconsciously perpetuate disproportionality despite good intentions to help others.
 - Another finding has been the presence of a strong social desirability bias amongst trainees not to appear racist. To counter the bias, trainees are asked to discuss topics that are clearly racist or not. Without the face validity, trainees discuss these topics with less fear of appearing racist. Trainees provide their initial (and possibly racist) opinions with the possibility of changing them after discussion. The training attitudes are measured in both pre- and post-tests of the course material and misconceptions of it.
 - During the skill components of trainings, supervisors develop Action Plans to establish concrete strategies that facilitate attitude change to impact behavior. Currently, Action Plans have not been effective because trainings have not covered the skill component sufficiently. One reason is that trainers have needed more time to cover the attitude awareness component to increase trainee receptiveness. Thus, expanding the classes to allow more time for discussion and the skill component has been identified as a solution.
- One concern is how to train authenticity. Even though trainees may be learning new attitudes, their attitudes, values, and beliefs must be authentic in order to have successful outcomes with clients. It should not be the goal to train “multicultural robots”.

III. Solving Disproportionality

- When presenting disproportionality in training settings, caution must be used to not blame the social workers as the cause of it. Workers who feel they don't harbor racist biases may feel blamed during cultural competency trainings.
- Laws and agency policies must be investigated to determine how they may push workers into making

unconsciously racist decisions that perpetuate disproportionality. All members of the child welfare system—workers, supervisors, and trainers—must analyze their role in disproportionality.

- Basic social work skills are identified as part of the solution; however, it is unknown if MSW students are entering the workforce with the appropriate attitudes, values and beliefs. Evaluation efforts must focus on this component.

IV. Findings from Utah

- Utah has similar disproportionality rates and has employed family-centered wrap-around services to increase cultural responsiveness. One finding is that more flexibility in basic social work skills is needed to provide culturally competent service. Workers also confront knowledge-based problems when generalizable data specific to certain ethnic groups is not generalizable.
- Evaluations must start from the outcomes. Receiving feedback from families who have received services can be valuable in determining which interventions were effective and ineffective. This feedback can be incorporated into trainings that serve to guide workers' actions in certain situations.
- Decision making at the assessment, planning, and court levels need to be evaluated for breakdowns.
- Organizational culture must also be investigated because our organizational culture reflects society's inability to address disproportionality.
- Workforce disproportionality is another problem in Utah pertaining to its increasing Latino population. Child welfare must be committed to promoting the professional development of Latino workers, which is a current challenge.
- Utah's current success in changing worker attitudes towards engaging families results from training current employees. Workers have learned and innovated new skills because the new organizational culture supports them. This environment has allowed old employees to learn new attitudes, values, and beliefs through training. Thus,

workers must not always have certain attitudes and values before entering the workforce.

- Attitudes, values, and beliefs of the staff who take the referrals need to be examined more closely. Are these staff members reflecting values of dominant culture toward racism (in the form of disproportionate referrals), or are they adding to racism through their own attitudes, values, and beliefs? Through education about disproportionality, workers must be trained to be aware of their biases and not to mirror larger society. Accomplishing this goal requires teaching workers skills.

	Wrap-up and Evaluation

The Responses to Questions on the General Evaluation are listed below.

1. What were the most useful parts of this Symposium?

Keynote:

- Dr. Kraiger was great. Thanks for being open to having him. (x8)
- Keynote speaker challenged our assumptions about Kirkpatrick's model, which has been the cornerstone of our evaluations.
- Some ideas from Kurt Kraiger that are very technical/skill based (e.g., rolling group design, internal referencing).
- Kraiger's keynote, although I wish he had gotten right to his slides without all of the digressions. As a result, we missed the good content at the end, the application example of interactions.

Format of presentations:

- Discussions—lets us learn more about practices and concepts others are using (x4)
- Limiting presentations to 15 minutes (x2)
- Both presentation and discussion (Q&A) from other states' perspectives was good, but not thorough enough for first-time participants.
- Opportunity to hear about the topic, then Q and A.
- Hearing about others' struggles and successes.
- Fewer presentations than in the past
- Honest questions and exchange
- Safe atmosphere to admit limitations and mistakes and to ask questions. (e.g., Henry's, Cindy and Susan's, Allen and Joan and Patricia's presentations.)
- Learning about what people are doing at different places in the country and about the challenges they encounter.

Wrap-up and Evaluation

Specific presentations:

- The presenters/presentations (x2)
- Really liked presentation Online vs. Classroom training (x2)
- Broad range of topics covered on given theme.
- Focus on “how to” evaluate training.
- Discussion of emerging issues.
- Megan’s procedure list
- Concept of aftercare focus for all families.
- Logic model and related curriculum and interventions.
- How logic models can serve as a framework for developing an evaluation plan—this was very helpful.
- Seeing ways in which other states have applied principles to their training data and at least tried to analyze them (e.g., evaluating documentation skills).
- Specific tools i.e., Documentation scoring rubric, Aftercare planning format
- The broad range of topics
- The statistics of instrumentation and measurement to evaluation of the relationship between personal values and professional values. We demonstrated clearer use of levels of evaluation in the work we’d been doing in the past year.

Atmosphere:

- The networking
- Meeting with colleagues
- Speaking with colleagues informally about training issues.

Logistical:

- Mailing the chapter from Kraiger’s book ahead of time was a great idea.
- Pre-symposium reading: 2005 *Proceedings* and Dr. Kraiger’s chapter
- Packet very good—including different color paper for different PowerPoint presentations.
- Logistics—tech support, food handouts,
- Location is good but is missing intimacy of Faculty Club.
- Staff work to coordinate and support symposium were excellent.
- Size of group

2. What could be done differently to make this Symposium even more effective?

Keynote:

- Open symposium with a speaker who frames the discussion: i.e., who articulates the rationale for the chosen theme—e.g., evidence-based practice
- I think it would be helpful to have had a Q&A session after the keynote that focused specifically on Kraiger’s talk.
- Another good speaker on evaluation—maybe more fitted to child welfare.
- Could have used more time on Kraiger’s presentation.
- Bring a measurement specialist/evaluation specialist (i.e., Terry Ackerman or Debra Bartz)

Programming:

- I also think we should try to do some sort of “Retrospective” on the growth of the symposium and where we’ve all come from where we were 10 years ago. Example— for the 1st symposium I was on a panel discussion talking a bout “how to develop written knowledge test item”—isn’t that amazing that we actually had a *session* on this.
- Fewer presentations and maybe ½ day with a carousel format to allow more in-depth small group discussion on topics previously presented earlier in the symposium.
- Keep format of short time for presentation, most time for questions, comments.
- Small group discussions—maybe having the presenters’ handouts available prior to meetings.
- To get more people involved perhaps have roundtable discussions or divide people into interest groups.
- Break-out discussion groups
- Interactive exercises
- Use a point-counterpoint format for one topic, differing eval results presented and defended.
- Have presenters highlight their learning points earlier in their presentations. Earlier learning points would focus and kick off the discussion.
- This is an excellent symposium—great discussions and presentations—I’d have preferred to have more formal

Wrap-up and Evaluation

introductions at beginning to put names with faces and places. Some folks I never really even learned their names.

- Most of the participants knew each other. New folks (like myself) did not find out what systems in other states looked liked and how they got there.
- More chances to hear what people are doing.
- It might be good to have a list with names of participants and the type of current program/projects they are working on so people can network with individuals working on like efforts during the breaks. (Note: CalSWEC sent the link to this in early July to all participants.)

Specific of presentations:

- Suggest/require that presenters distribute copies of papers/drafts of their presentations.
- Presenters could take a small amount of time to explain how their information relates to Evaluation and Training (for benefit of new people).
- Having participants submit questions that would be disseminated, such as, “I need help with...” “Does anybody know about...”
- Provide suggestions to presenters for ways to enhance presentation skills. Individualize their feedback but provide feedback to each presenter so no one is made to feel less capable than others. Perhaps a small team could be selected to do this each year.

Possible topics:

- More of the information on training efficacy
- Bring back topics and presenters with follow-up to last year
- RASCH modeling
- Debate political and ethical issues related to training evaluation

Logistical:

- Go back to the Faculty Club (upstairs room) for the 10th Anniversary. I realize this may mean limiting enrollment, but it really adds to the experience. (x2)
- After 9 years, I think all the kinks are out and the planning has been well honed.

- Please reconsider the necessity of recording the discussion, it interrupts the flow and stunts the discussion. Is it really essential to capture all of this anyway?
- Don't like room and use of microphone—takes energy and spontaneity away from discussion, adds formality.

3. How will you apply what you've learned at this Symposium to your job? Please provide at least two specific examples.

From Keynote:

- Apply some of Kraiger's concepts in evaluations
- Probably will use Kraiger's model as a foundation for N.C.'s new strategic evaluation plan we will be developing
- Discuss the idea of rolling designs as a way to get at changing effectiveness of trainings
- I can directly use the two items described from Kraiger's presentation and may use the POLYDETECT system.
- Using Kraiger's model to re-think evaluation construct.
- Rolling group design
- Dr. Kraiger's presentation and chapter will inform my next literature review.

Strategic efforts:

- Focus more on strategic purpose of training evaluations
- Advocate for more comprehensive evaluation of statewide pre-service training
- Provide more leadership on evaluation of current trainings for CW workers within state
- I will inquire about how CW training in my state is evaluated
- Research on our training moving beyond Level 1, I was not familiar with Kirkpatrick but will learn more about levels and how to measure in our new training system
- Explain the use of QA instruments as training evaluation measures and the development of attitudinal assessments
- Looking at different ways to deliver training—I will share with our training director ideas about IPR and web-based training.
- Will consider the effectiveness of current evaluations being used

Wrap-up and Evaluation

- Will be weaving specific activities into pre-service evaluations
- Will definitely use the info on online training to develop processes
- Using the examples/handouts to enhance training methods

Technical:

- RASCH modeling (x2)
- Cut scores
- I'd like to try Internal Referencing strategy in some of our work. May be look at correlating item difficulty or errors or readability index with.

Other:

- Use of logic models for training evaluations, do item analysis, e.g., Parry's approach, look beyond "levelese" and think/act less linearly about training evaluations
- Create a logic model to use as a framework for our larger evaluation plan
- Logic modeling
- I didn't personally get much that was new. I will use the methodology of questions on point and those that were not taught to make comparisons stronger on pre- and post-tests.
- Discuss aftercare and administration
- I will take online trainings to learn more about web-based vs. traditional training modalities.
- Evaluation learning process
- Evaluation activities
- Be in touch with specific presenters who are working on things that are similar to what I am working on
- Contact people who have mentioned particular literature to get citation, instruments, etc.
- Megan Paul's handouts
- CalSWEC info

4. What topics would you like to see covered in the next National Training Evaluation Symposium?

Evaluation development and methodology:

- Details on developing knowledge assessment
- Evaluation design, methodology and analysis
- Evaluation methods

- Qualitative methods
- Qualitative methods in training evaluation
- Mentoring as an evaluation medium (Midge should present)
- Pull out studies with mixed methods, counseling educators and training, especially as SW educators and Child Welfare workers and increasingly intersecting
- Relationship of changes in practice to the way evaluation is done
- More of decision-based evaluation
- More examples of instruments that others have successfully used
- Could we have more examples of behaviorally specific indicators for practice approaches such as engagement on teamwork, risk assessment or removing a child?

Outcomes/Impact:

- Keep looking for ways to measure impact—or “organizational payoff” using Kraiger model.
- Defining outcomes—I think a lot of the presentations revolved around strategies with doing this well. This also implies building the causal chain.
- Impact of training and child outcomes
- Impact of training in organizational culture and climate
- More on performance outcomes—how to measure improved practice

High stakes testing:

- I would like to do a session on high-stakes testing w/Cindy—like: pros and cons, or a point/counterpoint debate. We could draw on our experiences with certification testing, development issues, working w/the agency to sort through the pros/cons issues inherent in high stakes testing. Realize this isn’t exactly “training evaluation” but it seems that the group gets to this topic every year and we never have time to fully explore the issues. (x2)
- Highs stakes testing, implications, barriers, successes

Measuring cultural competence/biases:

- Maybe it's time to talk about how to evaluate cultural competence in this forum.
- Training and evaluation as Ethnicity Bias (A/V/B)

Other:

- Issues of cut-scores
- More presentation from the local level as opposed to university research
- Ways to overcome/address challenges and barriers that we all encounter as training evaluators
- Please renew the focus on evaluation, perhaps by increasing or changing the requests for proposals.
- Continuation of this year's theme
- I am interested in how SW Educators and CW professionals can inform each other to further professionalize agencies and make educators more relevant.
- How has all of the change that has flattened the world the 10 years (The World is Flat by T. Friedman) improved training evaluation?
- Where are the states represented in terms of training of CW staff? How are state or county systems different from privatization models (Florida, Kansas, Texas)?

5. What could Symposium planners do differently to enhance the Symposium?

Overall:

- Great job as usual!! You guys rock! (x3)
- It was nicely planned and executed.
- Excellent job with the handouts and coordination of presenter's materials
- The planners did a fantastic job.
- Could the planners do an eval-based survey, needs assessment in the fall of 2006 to assist planning?
- Thanks for everything; my first time was a great experience.

Programming:

- Accept proposals that pose works in progress and questions for participant discussion and feedback.
- Look at an overview of range of topics.

- Maybe if there was some type of mechanization to find out what others are doing in other states. The resource table is OK, but maybe devote one slot in the symposium for everyone to share to the entire group.
- Avoid presentations that focus on the content of a curriculum and the specific evaluation findings.
- Require presenters to generate discussion questions as part of submission (of proposal).
- Increase the time allotted for presenters to speak (from 15 to 20-25 minutes).
- A few of the presentations lacked focus—guidelines for them would help.
- If the symposium continues to serve a small group of experienced CW training evaluators, then current process/topics is OK. If the symposium wants to reach out and help other CW training systems, not as sophisticated as current participants, then the old way of doing business won't work. The average age of the room is about 45 years old. Where is the next generation of evaluators?
- Not really sure, thought the symposium was well done. Would like more specific data on retention in systems and what presenters/participants believe have affected this.
- Some presentations were not what I expected based on the proposal. Maybe people should have to provide more detail about what they plan to do?

Venue:

- I really like being in the Faculty Club, even though it is often a bit crowded—the auditorium creates a distance between speaker and audience that does not fit the ambiance/ethos of the symposium. Also, food is much better at the Faculty Club. (x2)
- Please keep it in a setting like the auditorium which provides plenty of space.
- The auditorium is perfect—especially the big screen.
- Auditorium was a little too big to encourage full interaction/participation, otherwise, planning/implementation was great.
- Smaller room (cozier)
- Use the same venue and place tables closer together.

Wrap-up and Evaluation

- Audio quality is much better this year.
- Room setting that supports discussion—Having to use a microphone was a challenge.

Meals:

- Thanks for the hot breakfasts.
- Breakfast was very heavy. I'd prefer lighter choices. It really would make me happy to have fruit available and then I would be less grumpy.
- Fruit/veggies with food!

	Program

Ninth Annual
**National Human Services
Training Evaluation Symposium
2006**

Wednesday, May 24

2:00-3:00 p.m.

Registration begins

3:00-3:15 p.m.

Convene, Welcome, and Introduction to the Symposium

- Barrett Johnson, LCSW, *Regional Training Academy Coordinator, California Social Work Education Center (CaISWEC), University of California, Berkeley*
-

3:15-4:00 p.m.

Interactive Exercise

Theme: Building Evidence

Title: Icebreaker Exercise

Facilitators:

- Jane Berdie, M.S.W., *Jane Berdie & Associates*
 - Dale Curry, Ph.D., LSW, *Kent State University*
-

4:00-5:00 p.m.

Theme: How Evaluation Brings Attention to Gaps in Training

Title: Training Evaluation as Quality Assurance: Moving from Theory to Reality

Presenter:

- Henry Ilian, D.S.W., *Administration for Children's Services, James Satterwhite Academy for Child Welfare Training*

Program

Facilitator:

- Midge Delavan, Ph.D., *Utah Division of Child & Family Services*

5:00-6:00 p.m.—Reception

6:00-7:30 p.m.—Dinner

Thursday, May 25

8:00-9:00 a.m.—Breakfast

9:00-9:15 a.m.

Reconvene, Introduction of Keynote Speaker

- Chris Mathias, M.S.W., *Director, California Social Work Education Center (CalSWEC), University of California, Berkeley*

9:15-10:15 a.m.

Keynote

- Kurt Kraiger, Ph.D., *Colorado State University—Fort Collins*

10:15-10:25 a.m.—Break

10:25-12:00 noon

***Theme:* Evaluating Teaching/Training Methods that Contribute to Knowledge Retention**

***Title:* Comparing Trainee Learning in Online and Classroom Deliveries of MEPA Training**

Co-Presenters:

- Cindy Parry, Ph.D., *CF Parry Associates*
- Susan Brooks, M.S.W., *Northern California Regional Training Academy*

Discussant:

- Kurt Kraiger, Ph.D., *Colorado State University—Fort Collins*

Facilitator:

- Todd Franke, Ph.D., *University of California, Los Angeles*

12:00 noon-1:00 p.m.—Lunch

1:00 -2:00 p.m.

Theme: How Logic Models Can Apply to Training Evaluation

Title: Using the Logic Model to Conceptualize and Enhance Training Evaluation Design

Co-Presenters:

- Alan Dettlaff, Ph.D., LMSW, *Texas Christian University*
- Joan Rycraft, Ph.D., ACSW, *University of Texas—Arlington*

Facilitator:

- Norma Harris, Ph.D., *University of Utah*

2:00-2:10 p.m.—Break

2:10 -3:10 p.m.

Theme: Methods of Measuring Knowledge/Skill

Title: *Evaluating Trainees' Documentation Skills*

Presenter:

- Megan Paul, Ph.D., *University of Nebraska—Lincoln*

Facilitator:

- Charlene Urwin, Ph.D., *University of Texas—Austin*

3:10-3:20 p.m.—Break

3:20-4:20 p.m.

Theme: Methods of Measuring Knowledge/Skill

Title: Aftercare: Planning for Post-Reunification Services

Presenter:

- John H. Trudeau, Ph.D., *Temple University Center for Social Policy & Community Development*

Facilitator:

- Jane Berdie, M.S.W., *Jane Berdie & Associates*

4:20-4:30 p.m.

Logistics for evening and morning

- Leslie W. Zeitler, LCSW, *California Social Work Education Center (CalSWEC), University of California, Berkeley*

4:30 p.m.—Break for the evening

Friday, May 26

8:00-9:00 a.m.—Breakfast

9:00-9:10 a.m.

Reconvene

- Barrett Johnson, LCSW, *Regional Training Academy Coordinator, California Social Work Education Center (CaISWEC)*

9:10 -9:40 p.m.

Technical Assistance Forum

Title: Using POLYDETECT to Assist in Assessing the Number of Constructs in a Survey

Presenter:

- Basil Qaqish, Ph.D., *University of North Carolina—Greensboro*

9:40-10:40 a.m.

Theme: Evaluating Assumptions, Attitudes, and Values

Title: Using Interpersonal Process Recall (IPR) to Measure Worker and Supervisor Effectiveness

Presenter:

- Patricia Taylor, Ph.D., LCSW, *University of Houston*

Facilitator:

- Marcia Sanderson, LCSW, *University of Texas—Austin*

10:40-10:50 a.m.—Break

10:50-11:50 a.m.

Theme: Evaluating Assumptions, Attitudes, and Values

Title: How Do We Evaluate Assumptions, Attitudes, and Values?

Presenters:

- Leslie W. Zeitler, LCSW
- Chris Mathias, M.S.W.
California Social Work Education Center (CaISWEC), University of California, Berkeley

Facilitator:

- Barrett Johnson, LCSW, *Regional Training Academy Coordinator, California Social Work Education Center (CaISWEC), University of California, Berkeley*

11:50-12:00 noon

Closing Remarks

- Barrett Johnson, LCSW, *Regional Training Academy Coordinator, California Social Work Education Center (CalSWEC), University of California, Berkeley*

12:00 noon-12:10 p.m.

Pick up box lunches: Working lunch

12:10-1:00 p.m.

Wrap up and Strategize for 2007

- Barrett Johnson, LCSW
- Leslie W. Zeitler, LCSW
California Social Work Education Center (CalSWEC), University of California, Berkeley

Acknowledgments

CalSWEC extends its gratitude to the members of the Steering Committee of the 9th Annual National Human Services Training Evaluation Symposium, who made this event possible:

Becky Antle	Kyle Hoffman
Jane Berdie	Henry Ilian
Sherrill Clark	Barrett Johnson
Dale Curry	Chris Mathias
Midge Delavan	Cynthia Parry
Bill Donnelly	E. Douglas Pratt
Michelle I. Graef	Leslie W. Zeitler

Program

The *California Social Work Education Center (CalSWEC)* is a partnership between the state's schools of social work, public human service agencies, and other related professional organizations that facilitate the integration of education and practice to assure effective, culturally competent service delivery and leadership to the people of California. CalSWEC is the nation's largest coalition of social work educators and practitioners.

California Social Work Education Center (CalSWEC)
University of California, Berkeley
School of Social Welfare
Marchant Building, Suite 420
6701 San Pablo
Berkeley, CA 94720-7420
Phone: 510-642-9272
FAX: 510-642-8573
<http://calswec.berkeley.edu>

For information about training evaluation in child welfare:
<http://calswec.berkeley.edu/CalSWEC/CWTraining.html>



Directory of Presenters and Participants

<p>Keynote Kurt Kraiger, Ph.D. Professor Colorado State University Department of Psychology Campus Delivery 1876 Fort Collins, CO 80523-1876 970-491-6821 970-491-1032 Kurt.Kraiger@colostate.edu</p>	<p>Kurt Kraiger is a Professor of Psychology and Co-Director of the Center for Organizational Excellence at Colorado State University. With a background in Industrial/Organizational Psychology, he is a noted expert on training and training evaluation. Dr. Kraiger has co-edited two books on training and published or presented over 120 academic papers, most on topics related to training and performance appraisal.</p>
<p>Jane Berdie, M.S.W. Child Welfare Consultant 435 South Gaylord St. Denver, CO 80209 303-733-9532 303-733-9532 (call first) jberdie@msn.com</p>	<p>Jane Berdie currently provides strategic planning and technical assistance to the CalSWEC Statewide Training Evaluation Project/Macro Evaluation Team, as well as to several training evaluation projects for Pennsylvania's Child Welfare Core Training and for independent living training in Colorado and North Carolina that involve embedded evaluation of knowledge and/or skills.</p>

Presenters and Participants

<p>Susan Brooks, M.S.W. Director Northern California Training Academy UC Davis Extension 1632 Da Vinci Court Davis, CA 95616-4860 530-757-8643 530-752-6910 sbrooks@unexmail.ucdavis.edu</p>	<p>Susan Brooks has nearly 20 years of experience in social services, with expertise in substance abuse, collaboration, team-building, and supervision. For seven years, Ms. Brooks supervised the multi-disciplinary team for children's services in San Mateo County. She was also the founder and Executive Director of the San Mateo Perinatal Council, a nonprofit community collaborative.</p>
<p>Soledad Caldera-Gammage, M.S.W. Evaluation & Curriculum Specialist Central California Child Welfare Training Academy 112 Ron Court Vallejo, CA 94591 707-647-1655 (phone & fax) scaldera@csufresno.edu</p>	<p>As the Curriculum, Program, and Evaluation Specialist for the Central California Training Academy at California State University, Fresno, Soledad Caldera-Gammage is responsible for the coordination and oversight of curriculum development. Her duties include the coordination of the training and mentoring evaluation projects at the Central Academy. She is also responsible for the recruitment, development, and evaluation of all Academy trainers. Additionally, she is a Trainer and Curriculum Developer in child welfare for the Specialized Foster Parent Training Project at CSU, Fresno.</p>

<p>Sherrill Clark, Ph.D., LCSW Research Specialist CalSWEC School of Social Welfare University of California, Berkeley 6701 San Pablo Ave. Suite 420 Berkeley, CA 94720-7420 510-642-4480 510-642-8573 sjclark@berkeley.edu</p>	<p>Sherrill Clark, formerly Curriculum Specialist and Executive Director at CalSWEC, has taught policy, practice, and research methods to first- and second-year master's students in the School of Social Welfare at the University of California, Berkeley. Her primary research areas include workforce preparation and retention of specially educated MSW child welfare workers and child welfare education. From 1998–2001, Dr. Clark was the Project Coordinator for the Child Welfare Fellows Project, a grant to enhance applied child welfare research and curriculum development awarded by the Children's Bureau. Dr. Clark was a member of the California Department of Social Services Stakeholders' Group for three years participating on the workforce preparation/support, fairness and equity, and the early intervention and differential response groups. She recently completed the 2004 California Public Child Welfare Workforce Study.</p>
<p>James A. Coloma, M.S.W. Research & Evaluation Specialist Public Child Welfare Training</p>	<p>James A. Coloma is the Evaluation Specialist of each of the seven programs which the Public Child Welfare</p>

Presenters and Participants

<p>Academy Academy for Professional Excellence 6505 Alvarado Rd., Suite 107 San Diego, CA 92120 619-594-3219 619-594-1118 jcoloma@projects.sdsu.edu</p>	<p>Training Academy administers. One of the seven programs is the Leaders in Action for Executive Participants program using web-based surveys.</p>
<p>Dale Curry, <i>Ph.D., LSW</i> Assistant Professor Kent State University School of Family & Consumer Studies Nixson Hall POB 5190 Kent, OH 44242-0001 330-672-2998 330-672-2194 dcurry@kent.edu</p>	<p>Dale Curry coordinates evaluation for a trainer development certificate program in collaboration with the Northeast Ohio Regional Training Center. He is the Principal Investigator for two statewide training evaluation projects in Ohio, and served as the consultant to the American Humane Associ- ation on its comprehensive evaluation of the Pennsyl- vania Competency-Based Child Welfare Training and Certification Program. Dr. Curry co-chairs both the National Staff Development and Training Association's Evaluation Committee and the Assessment Committee of the North American Certification Project for child and youth workers.</p>
<p>Midge Delavan, <i>Ph.D.</i> Training Coordinator Utah Division of Child & Family Services 120 North 200 West, #225 Salt Lake City, UT 84103</p>	<p>Midge Delavan, an instructional psychologist, is the Training Coordinator of the Training Team at Utah's Division of Child & Family Services, which is planning an</p>

Presenters and Participants

<p>801-538-4404 801-538-3993 mdelavan@utah.gov</p>	<p>evaluation system that includes mentoring as an evaluation medium.</p>
<p>Diane DeMark, LHRM Senior Vice President, Administration Children’s Home Society of Florida 1485 S. Semoran Blvd. Suite 1448 Winter Park, FL 32792 321-397-3000 321-397-3018 diane.demark@chsfl.org</p>	<p>Diane DeMark is responsible for the oversight of the human resources, information systems, training and curricula, and the development of research-based best-practice models within the Children’s Home Society of Florida. She is also responsible for coordinating the development of strategies to respond to community child welfare privatization initiatives throughout Florida.</p>
<p>James C. DeMark, M. Ed. Vice President, Quality Management Children’s Home Society of Florida 1485 S. Semoran Blvd. Suite 1448 Winter Park, FL 32792 321-397-3000 321-397-3018 jim.demark@chsfl.org</p>	<p>James C. DeMark is responsible for the overall administration of Children’s Home Society of Florida’s quality improvement system. He supervises a staff of three Quality Directors, a Corporate Director of Accreditation, and a Compliance Coordinator, and manages a department of 22 staff with a budget of over \$1.8 million. Mr. DeMark also serves as the agency’s HIPAA privacy officer.</p>
<p>Christy Derrick, M.P.H. Research Associate Center for Child & Family Studies University of South Carolina College of SW Benson Bldg.</p>	<p>Christy Derrick manages the evaluation of child welfare training projects, including the evaluation of child welfare new worker basic training. She is assisting in developing</p>

Presenters and Participants

<p>Columbia, SC 29208 803-777-8494 803-777-1366 cderrick@sc.edu</p>	<p>and conducting the evaluation for a child welfare supervisory training series. Ms. Derrick manages several smaller evaluations of child welfare training and conducts training needs assessments.</p>
<p>Alan J. Dettlaff, Ph.D., MSW Jane Addams College of Social Work University of Illinois at Chicago 1040 W. Harrison (MC 309) Chicago, IL 60607 312-996-4629 aland@uic.edu</p>	<p>Alan J. Dettlaff is an evaluator of a federal grant from the Administration of Children, Youth, and Families that is providing training to child welfare staff on effective practices with Hispanic children and families. His research interests focus on supervision, retention, and effective practice in child welfare. He is currently involved in research that is exploring the relationship between personality type and retention in Child Protective Services, and effective strategies to address racial disproportionality in child welfare.</p>
<p>Nancy S. Dickinson, Ph.D. Executive Director Jordan Institute for Families UNC Chapel Hill School of Social Work 301 Pittsboro St., CB #3550 Chapel Hill, NC 27599-3550 919-962-6407 919-843-9827 ndickins@email.unc.edu</p>	<p>Nancy S. Dickinson is a Clinical Professor and the Executive Director of the Jordan Institute for Families at the UNC Chapel Hill School of Social Work. The Institute focuses on strengthening families through research and educational and technical assistance projects. Dr.</p>

Presenters and Participants

	<p>Dickinson is currently Principal Investigator and Director of a five-year Children’s Bureau-funded project entitled “Child Welfare Staff Recruitment and Retention: An Evidence-Based Training Model” in which a resources and curriculum model will be developed and delivered to supervisors, managers, and directors in 17 North Carolina counties. A rigorous experimental evaluation will compare outcomes within each of the trained counties before and after training and with outcomes in 17 comparison counties.</p>
<p>Bill Donnelly, LCSW, M.P.A. Director Inter-University Consortium Center for Child & Family Policy 3250 Public Policy Bldg. PO Box 951656 Los Angeles, CA 90095-1656 310-825-2811 310-206-2716 donnelly@spps.ucla.edu</p>	<p>The Inter-University (IUC) Consortium and Los Angeles County Department of Children and Family Services (DCFS) Training Project is a collaborative endeavor between the DCFS and the graduate programs of social work at California State University, Long Beach; California State University, Los Angeles; University of California, Los Angeles; and the University of Southern California. The IUC/DCFS Training Project is currently</p>

Presenters and Participants

	<p>in its 11th year. In 2000–2001 the Training Project trained 6,971 staff and generated 155,174 training hours.</p>
<p>David Foster, LCSW Director Central California Child Welfare Training Academy California State University, Fresno 5310 N. Campus Dr., M/S PH102 Fresno, CA 93740-8019 559-294-9760 559-292-7371 david_j_foster@csufresno.edu</p>	<p>David Foster has 19 years' experience in public child welfare as a social worker, supervisor, and manager. He was Title IV-E Coordinator for the California State University, Fresno Social Work Program from 1993-1998, was founding member of the Central California Regional Training Academy Development and Implementation Project, and has served as Academy Director since 1998. He is married to Barbara; they have five children, (Adrian, Amanda, Nicole, Matt, Stephen) and one grandchild (Ryan). They are avid Ice Hockey fans (in Fresno, of all places...go figure!).</p>
<p>Todd M. Franke, Ph.D. Associate Professor UCLA Department of Social Welfare 405 Hilgard Ave., Box 165606 Los Angeles, CA 90095-1656 310-206-6102 310-362-8673 tfranke@ucla.edu</p>	<p>Todd M. Franke has numerous years of experience in research, teaching, and evaluation in the fields of education and child welfare. He has considerable experience in designing and conducting cross-sectional and longitudinal research and evaluation in the areas of child abuse and adolescent violence. He is currently</p>

Presenters and Participants

	<p>the Principal Investigator for the evaluation of Partnership for Families, a community-based child abuse and neglect prevention initiative in Los Angeles County. In addition, he is currently working with the county's Department of Children and Family Services on the redesign of the Core Academy Assessment and on evaluating the implementation and impact of training numerous modules including strength-based family-centered case planning and concurrent planning.</p>
<p>Walter Furman, M.Phil. Academic Coordinator UCLA Department of Social Welfare 3250 Public Policy Bldg. POB 951656 Los Angeles, CA 90095-1652 310-825-0852 310-206-2227 wfurman@ucla.edu</p>	<p>Walter Furman's work focuses on research and evaluation of child welfare services, including training evaluation for the Inter-University Consortium that recently included studies of Strength-Based and Concurrent Planning Training Evaluation of Prevention Initiative in California's small counties.</p>
<p>Elizabeth Gilman, M.A., J.D. Curriculum Specialist CalSWEC School of Social Welfare University of California, Berkeley 6701 San Pablo Ave., Suite 420 Berkeley, CA 94720</p>	<p>Elizabeth Gilman became a member of the CalSWEC staff in March 2002. Her responsibilities include curriculum development, evaluation, and planning associated with the statewide Title IV-E MSW Program and the newly launched BSW</p>

Presenters and Participants

<p>510-642-9273 510-642-8573 egilman@berkeley.edu</p>	<p>Program. She also serves as staff to the CalSWEC Board Curriculum Committee. Prior to joining CalSWEC, Ms. Gilman was an Associate Research Scientist in the Psychology Department at Yale University, where she served as an instructor and policy analyst specializing in child development and social policy issues. Her direct service experience includes several years as a group facilitator with troubled adolescents and as a protective services social worker.</p>
<p>Shaaron Gilson, LCSW Title IV-E Project Coordinator School of Social Welfare University of California, Berkeley 120 Haviland Hall, MC 7400 Berkeley, CA 94720 510-642-2424 510-643-6126 shaaron@berkeley.edu</p>	<p>Shaaron Gilson has served as the Title IV-E Project Coordinator for the University of California, Berkeley School of Social Welfare for over 10 years. Her prior experience includes a career in public and private agencies in mental health and child welfare services, as well as academic appointments. Her most recent evaluation projects include the ongoing assessments of the Title IV-E training curriculum.</p>
<p>Erika Gonzalez Training Specialist Bay Area Academy 2201 Broadway, Suite 100 Oakland, CA 94612</p>	

Presenters and Participants

<p>510-419-3605 510-663-5532 erika815@sfsu.edu</p>	
<p>Michelle I. Graef, Ph.D. Research Associate Professor Center on Children, Families & the Law University of Nebraska— Lincoln 121 South 13th St., Suite 302 Lincoln, NE 68588-0227 402-472-3741 402-472-8412 mgraef1@unl.edu</p>	<p>Under a contract with the Nebraska Health and Human Services System (NHHSS), Michelle L. Graef is involved in designing and implementing all the training evaluation procedures for the pre-service training for CPS workers in Nebraska. Her current work also includes: the design and evaluation of new supervisor training; research and consultation with human services agencies in Nebraska and other states on staff recruitment, selection, and retention issues; and providing research and statistical support to the NHHSS Quality Assurance Unit.</p>
<p>Carol J. Harper, M.P.A. Consultant Northwest Institute for Children and Families University of Washington School of Social Work 4101 15th Ave. NE Seattle, WA 98105 719-635-7942 charper@u.washington.edu</p>	<p>Carol J. Harper is on the faculty at the University of Washington School of Social Work and serves as a research manager for the Northwest Institute for Children and Families. She has over 17 years of experience in child welfare and human service evaluation. Ms. Harper is a former Director of Systems and Community Programs at the American Humane Association. She is currently</p>

Presenters and Participants

	<p>evaluating the State of Washington’s IV-E education and professional development program, developing a knowledge assessment tool for new social worker training (Academy), and conducting a follow-up evaluation of Academy effectiveness. Ms. Harper’s M.P.A. from Troy State University includes an emphasis on research and planning. She has co-authored several implementation and research guides on family group decision making, conducted social worker trainings nationally and internationally, and supported the development of safety and risk assessment tools as well the development of a parental skills assessment.</p>
<p>Norma Harris, Ph.D. Director Social Research Institute University of Utah College of Social Work 1500 E 395 S. Rm. 130 Salt Lake City, UT 84112 801-581-3822 801-585-6865 normaharris@socwk.utah.edu</p>	<p>Norma Harris is the Director of the Social Research Institute and Research Professor/Lecturer in the College of Social Work at the University of Utah. Her degrees are from Montana State University (B.S.), University of Utah (M.S.W.), and Florida International University (Ph.D.). She is currently responsible for the management of grants at SRI that exceed \$8 million. Formerly Director of Community Services in</p>

Presenters and Participants

	<p>Montana, Dr. Harris has extensive experience in child welfare services. She also has experience in evaluation of child protective and child welfare programs. She was a member of the committee that assisted in the development of the U.S. Department of Health and Human Services Resource Guide: Rethinking Child Welfare Practice Under The Adoption and Safe Families Act of 1997. She is a contributing author of the first edition and the principal author of the revised Guidelines for a Model System of Protective Services for Abused and Neglected Children and Their Families.</p>
<p>Claudia “Kyle” Hoffman Project Manager Ohio Child Welfare Training Program Evaluation Project Institute for Human Services 1706 E. Broad St. Columbus, OH 43203-2039 614.251.6000 614-251-6005 khoffman@ihs-trainet.com</p>	<p>As a training coordinator with the Institute for Human Services, Kyle Hoffman chairs the Ohio Child Welfare Training Program (OCWTP) Evaluation Work Teams. Team members include: County child welfare staff, Regional Training Center coordinators, the Ohio Department of Job and Family services staff, and research consultants. The Work Team is involved in the effort to develop a comprehensive system to evaluate knowledge acquisition, knowledge comprehension, skill</p>

Presenters and Participants

	demonstration, and skill transfer that occurred as a result of OCWTP Training.
<p>Teresa Hubley, M.P.A., Ph.D. Research Associate Institute for Public Sector Innovation Edmund S. Muskie School of Public Service University of Southern Maine 295 Water St. Augusta, ME 04960 207-626-5292 207-626-5210 teresa.a.hubley@maine.gov</p>	<p>Teresa Hubley specializes in child welfare projects and general training evaluation. Dr. Hubley’s doctoral degree from Syracuse University is in cultural anthropology, specializing in medical anthropology, and her M.P.A, also from Syracuse, focuses on health policy.</p>
<p>Henry Ilian, D.S.W. Training Evaluator James Satterwhite Academy for Child Welfare Training New York City Administration for Children’s Services 492 First Ave., 5th Floor New York, NY 10016 646-935-1410 646-935-1604 henry.ilian@dfa.state.ny.us</p>	<p>Henry Ilian has been involved since 1987 in evaluation and testing at the James Satterwhite Academy for Child Welfare Training. He has been developing competency-based measures to accompany the New York City adaptation of the New York State Common Core training system for child welfare workers. He also teaches research at the Columbia University School of Social Work.</p>
<p>Stephanie Jackson Training Assistant Bay Area Academy 2201 Broadway Suite 100 Oakland, CA 94612</p>	

Presenters and Participants

<p>831-583-9419 831-583-9419 jacksonsbaa@sbcglobal.net</p>	
<p>Susan E. Jacquet, Ph.D. Research Specialist CalSWEC School of Social Welfare University of California, Berkeley 6701 San Pablo Ave. Suite 420 Berkeley, CA 94720-7420 510-643-9846 510-642-8573 sjacquet@berkeley.edu</p>	<p>Susan E. Jacquet works with Dr. Sherrill Clark on CalSWEC’s research component, including ongoing surveys of California’s MSW students and CalSWEC Title IV-E MSW graduates who have completed payback work in Child Welfare, and the development of new research initiatives on outcomes for child welfare and the efficacy of the Title IV-E program. Dr. Jacquet is also responsible for coordinating CalSWEC’s funded research process from RFP through review of proposals.</p>
<p>Phyllis Jeroslow, M.F.A., MFT Training and Curriculum Specialist CalSWEC School of Social Welfare University California, Berkeley 6701 San Pablo Ave. Suite 420 Berkeley, CA 94720-7420 510-643-5440 510-642-8573 pjero@berkeley.edu</p>	<p>Phyllis Jeroslow assists in the development and dissemination of statewide standards and curriculum for core and ongoing child welfare training. She also coordinates the annual Symposium on Fairness & Equity Issues in Child Welfare Training.</p>

Presenters and Participants

<p>Barrett Johnson, LCSW Regional Training Academy Coordinator CalSWEC School of Social Welfare University of California, Berkeley 6701 San Pablo Ave. Suite 420 Berkeley, CA 94720-7420 510-643-5484 510-642-8573 barrettj@berkeley.edu</p>	<p>Barrett Johnson oversees CalSWEC’s coordination of statewide training efforts, including development and implementation of a common core curriculum, and strategic planning for a statewide training evaluation system. He is involved in planning the training efforts for California’s Outcomes and Accountability System. Mr. Johnson has worked for many years with urban children and families, with an emphasis on intervention in cases of child sexual abuse.</p>
<p>Dulcey G. Laberge, M.S.W., LSW Director Division of Public Service Management Department of Health and Human Services Office of Child & Family Services 221 State St., Station 11 Augusta, ME 04333 207-287-6462 207-287-5282 dulcey.laberge@maine.gov</p>	<p>Dulcey G. Laberge oversees program funding, expenditures, grants contracts with service providers, SACWIS system, and personnel matters as the Director of the Division of Public Service Management for the Office of Child and Family Services with the Maine Department of Health and Human Services.</p>
<p>Judith Lefler, PHN, CLNC Assistant Director Bay Area Academy 2201 Broadway, Suite 100</p>	<p>As Assistant Director of the Bay Area Academy, Judith Lefler has oversight of all deliverables and capacity</p>

Presenters and Participants

<p>Oakland, CA 94612 707-428-0133 707-428-0137 judithlefler@aol.com</p>	<p>building activities in 12 Bay Area Counties</p>
<p>Elizabeth “Betsy” Lindsey, Ph.D. Professor Department of Social Work University of North Carolina— Greensboro POB 26170 Greensboro, NC 27402 336-334-5225 betsy_lindsey@uncg.edu</p>	<p>Elizabeth Lindsey teaches in the Joint Master of Social Work program administered by the University of North Carolina—Greensboro and North Carolina A&T State University. Since 1998, she has been the Principal Investigator for the North Carolina Child Welfare Training Evaluation Project. Along with Project Co-Director Dr. Fasih Ahmed of North Carolina A&T State University, Dr. Lindsey has developed and implemented a strategic evaluation plan for the child welfare training system in North Carolina.</p>
<p>Susan Mason, Ph.D., LCSW Professor Social Work Education Consortium Yeshiva University Wurzweiler School of Social Work 2495 Amsterdam Ave. New York, NY 10033 212-960-0806 masonse@yu.edu</p>	<p>Susan Mason is currently working on several projects, including: (1) Year 4 of the New York State Social Welfare Child Welfare Consortium’s evaluation of incentive course work at the state’s schools of social work to gauge the transfer of learning; (2) Year 3 of a supervisor study to learn how supervisors value and/or observe workers’ transfer of learning; and (3) Year 3 of a student unit evaluation of a</p>

Presenters and Participants

	designated child welfare unit at a children’s sex abuse clinic.
<p>Chris Mathias, M.S.W. Director CalSWEC School of Social Welfare University of California, Berkeley 6701 San Pablo Ave. Suite 420 Berkeley, CA 94720-7420 510-642-7490 510-642-8573 cmathias@berkeley.edu</p>	<p>Chris Mathias began her work with CalSWEC in March 2000, when she initially headed the Regional Training Academy Coordination Project. As Director, she leads the development and evaluation of the Title IV-E Stipend Program for Public Child Welfare and the Regional Training Academy Coordination Project. She heads a consortium that includes 18 universities, the County Welfare Directors Association, the California Mental Health Directors Association, the four Regional Training Academies, the Inter-University Consortium in Los Angeles, and the California Department of Social Services. For 14 years prior to joining CalSWEC, Ms. Mathias worked primarily in the private non-profit sector with children in out-of-home care. During that period, she developed curriculum, training, and quality assurance methods for practice for direct care workers, clinicians, and administrators.</p>

Presenters and Participants

<p>Tim McCarragher, Ph.D., LISW Associate Professor School of Social Work University of Akron 80 No. Portage Path 8B11 Akron, OH 44325 330-972-5976 330-972-5739 TimMcC100a@aol.com</p>	<p>Tim McCarragher teaches graduate level research courses. For the past three years, he has collaborated with the Institute for Human Services, and the Ohio Child Welfare Training Program (OCWTP) Evaluation Work Team. Team members include: county child welfare staff, Regional Training Center Coordinators, Ohio Department of Job and Family Services staff, and research consultants. The Work Team is involved in the effort to develop a comprehensive system to evaluate knowledge acquisition, knowledge comprehension, skill demonstration, and skill transfer that occur as a result of OCWTP Training.</p>
<p>Michael Nunno, D.S.W. Senior Extension Associate College of Human Ecology Family Life Development Center Cornell University Beebe Hall Ithaca, NY 14853-4401 607-254-5127 607-255-4837 man2@cornell.edu</p>	<p>Michael Nunno is a Senior Extension faculty member of the College of Human Ecology at Cornell University and Principal Investigator of the Residential Child Care Project at the Family Life Development Center. His primary work is in the evaluation of training, participant performance in, and organizational assessment in residential treatment, hospital, and corrections facilities for children.</p>

Presenters and Participants

<p>Kirstin O'Dell, M.S.W. Training Evaluator Portland State University Child Welfare Partnership 4061 Winema Place NE Salem, OR 97305 503-399-5152 503-399-6439 kodell3@chemeketa.edu</p>	<p>Kirstin O'Dell has eight years of research and evaluation experience in child welfare and early childhood. Her current projects include evaluation of the training for Excellence in Child Welfare Practice in rural Oregon and Alaska, and grant evaluation of Oregon DHS-funded trainings provided by CWP.</p>
<p>Cynthia Parry, Ph.D. Consultant C.F. Parry Associates 520 Monroe Ave. Helena, MT 59601 406-449-3263 406-449-3263 cfparry@msn.com</p>	<p>Cynthia Parry specializes in program evaluation in human services. She has over 20 years of experience in evaluation in child welfare, juvenile justice, and education. Dr. Parry is a former Director of Program Analysis and Research at the American Humane Association. She is currently involved in evaluations of training effectiveness in Colorado, California, and Washington, and has been co-facilitator for a strategic planning process for training evaluation in California. Her doctorate is in educational psychology from the State University of New York at Albany, with a strong emphasis on curriculum design and assessment, and she holds a permanent Teacher Certification from New York State. She is a co-author of the National Staff</p>

Presenters and Participants

	<p>Development and Training Association (NSDTA) publication <i>Training Evaluation in the Human Services</i> and co-chair of the NSDTA Evaluation Committee. Dr. Parry has also served for several years on the Planning Committee for the National Human Services Training Evaluation Symposium, and has made numerous presentations on the evaluation of training at this meeting and nationally.</p>
<p>Megan E. Paul, Ph.D. Research Assistant Professor of Psychology Center on Children, Families & the Law University of Nebraska, Lincoln 121 S. 13th St., Suite 302 Lincoln, NE 68588 402-472-9812 402-472-8412 Mpotter2@unl.edu</p>	<p>Megan E. Paul is jointly responsible for conducting training evaluation and competency assessment projects designed to improve the selection, training, performance, and retention of protective service workers in Nebraska. Her recent training evaluation projects include development of annual reports to identify trends in trainee feedback within and across training units and trainers; continual development and update of written knowledge tests; and development of skills evaluations, such as for computer documentation and home visits.</p>
<p>Robin Perry, Ph.D. Professor Florida State University</p>	<p>Robin Perry is the Associate Director for the Institute for Health and Human Services</p>

Presenters and Participants

<p>School of Social Work 2500 University Center, Bldg. C Tallahassee, FL 32306-2702 850-322-1901 reperry@fsu.edu</p>	<p>Research and Principal Investigator for Florida State University’s Title IV-E Stipend Program and associated evaluation efforts. He is a former research associate at CalSWEC and has published on training and training evaluation topics. Dr. Perry’s recent research within child welfare has focused on the following areas: the examination of the influence of educational background upon CPI and CPS workers’ job performance in Florida; the development of a task-analysis tool/model for child protective service workers; a cost analysis of integrating select information technology into child protection investigations in Florida; and the development of a statistical/actuarial model for the equitable distribution of child welfare service money in Florida.</p>
<p>E. Douglas Pratt, D.S.W., LCSW President Policy–Practice Resources Inc. 3786 Evans Rd. Atlanta, GA 30340 770-723-9105 770-723-9105 ppr-edp@juno.com</p>	<p>As a Quality Service Review (QSR) Trainer, E. Douglas Pratt has contributed to refinements of that practice evaluation methodology and has explored Level III and IV training evaluation. With QSR proprietors and colleagues, Dr. Pratt is developing ways to modify the methodology, including instrument</p>

Presenters and Participants

	<p>recalibration, and adjustments in data collection and analysis processes, for training evaluation. He will be training scientists at The Centers for Disease Control and Prevention on community empowerment benefits of QSR for child abuse and domestic violence prevention this spring.</p>
<p>Basil Qaqish, Ph.D. Research Scientist University of North Carolina at Greensboro 262 Stone Bldg. Greensboro, NC 27402 bfqaqish@uncg.edu</p>	<p>Basil Qaqish works on several projects, including the development of five knowledge assessment standardized tests for different social work trainings in North Carolina. Dr. Qaqish also works on survey developments and analysis.</p>
<p>Leslie Rozeff, LCSW Director University of Southern Maine Child Welfare Training Institute 295 Water St. Augusta, ME 04330 207-626-5218 207-626-5088 leslie.rozeff@maine.gov</p>	<p>Leslie Rozeff oversees the administration and broad evaluation activities of child welfare training within the program. Recent activities include providing resources and support to priority tasks identified by human services state government offices related to refining and reorganizing existing programs into a comprehensive client-centered system. Additional activities include technical assistance, policy analysis, and facilitation.</p>

Presenters and Participants

<p>Joan R. Rycraft, Ph.D. Associate Dean Director , Ph.D. Program University of Texas at Arlington Box 19129 Arlington, TX 76019 817-272-5225 817-272-2028 rycraft@uta.edu</p>	<p>Joan R. Rycraft is Associate Dean and Associate Professor at the University of Texas at Arlington. Dr. Rycraft is the Director of the Ph.D. Program, and a Faculty Associate of both the UTA School of Social Work's Center for Child Welfare and its Center for Research, Evaluation and Technology. Her social work practice spans 20 years of public child welfare services in the state of California, as an investigator, supervisor, and administrator. Dr. Rycraft teaches courses in child and youth policy, advanced administration, and research methods and evaluation in the baccalaureate, master's, and doctoral programs.</p>
<p>Marcia Sanderson, LCSW Director Protective Services Training Institute School of Social Work University of Texas at Austin 1925 San Jacinto Austin, TX 78712-0358 512-471-0521 512-232-9585 MSanderson@mail.utexas.edu</p>	<p>Marcia Sanderson has been the Director of the Protective Services Training Institute (PSTI) since 1993. From 1999 to 2002, she was also the Director of the Child Welfare Education Project, a Title IV-E stipend program at the University of Houston Graduate School of Social Work. In 2002, she became full-time PSTI Director and relocated to the University of Texas at Austin. While at the University of Houston, she taught grant writing in the</p>

Presenters and Participants

	<p>social work master's program and program evaluation as part of a continuing education workshop in program planning and proposal writing.</p>
<p>Bernadette Sangalang, Ph.D., M.S.W. Director of Evaluation and Research Center for Training, Evaluation & Research of the Pacific University of Hawaii, Manoa School of Social Work 1800 East West Rd., Henke Hall Honolulu, HI 96822 808-956-9364 808-956-5964 bbs@hawaii.edu</p>	<p>Bernadette Sangalang is currently conducting an evaluation of Hawaii's Department of Human Services child welfare training for supervisors, line workers, and foster parents.</p>
<p>T. Maureen Sinclair, LMSW Project Staff Associate Professional Development Program/OCFS Rockefeller College University at Albany 52 Washington St., Rm. 321 North Rensselaer, NY 12144 518-473-8182 518-474-1842 Maureen.Sinclair@ocfs.state.ny.us</p>	<p>Maureen Sinclair is affiliated with the Professional Development Program—Rockefeller College, and with the Social Work Education Consortium, University at Albany—SUNY. In addition to working on child welfare curriculum development for New York State child welfare supervisors, Ms. Sinclair is engaged in doctoral studies at the University at Albany. Her research interests include the impact of trainer characteristics on skill acquisition (training</p>

Presenters and Participants

	pedagogy) as well as transfer of training issues.
<p>Jeri Smith Staff Development Specialist Contra Costa County Employment & Human Services Department 1465-D Enea Circle #188 Concord, CA 94520 925-808-2526 925-808-2539 smithjo@ehsd.cccounty.us</p>	<p>Jeri Smith’s primary responsibility is training new child welfare workers for Contra Costa County. Her most recent evaluation project was measuring change in trainee attitudes after viewing the movie “Crash.” Ms. Smith and a co-worker created pre- and post-tests to measure change in trainee attitudes regarding race, and elicited from trainees how they could affect social work practice after viewing the movie.</p>
<p>Patricia G. Taylor, Ph.D., LCSW Curriculum Coordinator Child Welfare Education Project Graduate College of Social Work University of Houston 237 Social Work Bldg. Houston, TX 77204-4013 713-743-8081 713-743-8888 ptaylor2@uh.edu</p>	
<p>John H. Trudeau, Ph.D. Director Center for Social Policy & Community Development Temple University 1601 N. Broad St., USB #100 Philadelphia, PA 19122</p>	<p>John H. Trudeau is the Director of the Center for Social Policy and Community Development in the School Administration at Temple University. The Center conducts a wide range of</p>

Presenters and Participants

<p>215-204-7491 215-204-3424 jtrudeau@temple.edu</p>	<p>demonstration and applied research projects in youth development, community capacity building, workforce development, and continuing education programs. Recent evaluation efforts include program evaluations of a Compassion Capital Fund grantee, and of a new technology program of Philadelphia's Office of Emergency Shelter Services, as well as continuing training evaluation of the city's Department of Human Services staff development program.</p>
<p>Charlene Urwin, Ph.D. Curriculum Director Protective Services Training Institute School of Social Work University of Texas at Austin 1 University Station D3500 Austin, TX 78712 512-471-0560 512-232-9585 curwin@mail.utexas.edu</p>	<p>Charlene Urwin is site manager for the Protective Services Training Institute of Texas at the University of Texas at Austin. She has directed training needs assessments and competency projects in child and adult protection, and recently completed a performance dimension approach in Child Care Licensing. Dr. Urwin has extensive experience in teaching and directing social work education programs.</p>
<p>PJ Verrecchia, Ph.D. Associate Professor of Research University of Pittsburgh Pennsylvania Child Welfare</p>	<p>PJ Verrecchia is involved in three main evaluation areas: (1) providing quantitative analysis for traditional CORE training and the new Charting</p>

Presenters and Participants

<p>Training Program 403 East Winding Hill Rd. Mechanicsburg, PA 17055 717-795-9048 x293 717-795-8013 pjv6@pitt.edu</p>	<p>the Course (CTC); (2) conducting case studies to further analyze the CTC training; and (3) validating the embedded evaluation questions that are a part of three CTC training modules.</p>
<p>Leslie W. Zeitler, LCSW Training and Evaluation Specialist CalSWEC School of Social Welfare University of California, Berkeley 6701 San Pablo Ave., Suite 420 Berkeley, CA 94720 510-643-6400 lzeitler@berkeley.edu</p>	<p>Leslie W. Zeitler coordinates and implements CalSWEC's statewide training and evaluation efforts. She also provides technical assistance to the California Regional Training Academies and all California counties. Ms. Zeitler also coordinates CalSWEC's annual National Human Services Training Evaluation Symposium.</p>