

## Saturday, November 4, 3:30 P.M. to 5:00 P.M.

### **SESSION 573: Think Tank**

**Room: Lanai**

Sponsored by the Topical Interest Group on Needs Assessment

#### **Toward the Development of an Organizational Evaluation Capacity Index**

Facilitators: Michal Galin, Independent Consultant  
Sally R Durgan, The California Endowment  
William Trochim, Cornell University

Assessing the capacity for evaluation within an organization is an important precursor to conducting effective and efficient evaluation. This session will discuss the ongoing development of an index to measure the evaluation capacity of an organization. The development process will be highly interactive, drawing upon the experience of participants of the session. Concept mapping will be utilized to facilitate the process including: brainstorming evaluation capacities; and, individual sorting and rating of brainstormed items. A relational map of concepts will emerge that will then be used as the framework for development of an index of items to assess the evaluation capacity of an organization.

### **SESSION 574: Demonstration**

**Room: Akaka**

Sponsored by the Topical Interest Group on Computer Use in Evaluation

#### **Performance Based Prevention System (PBPS): From Needs Assessment, Program Planning, Service Tracking, Evaluation to Data Utilization**

Chair: Xiaoyan Zhang, Human Services Software Inc  
Presenter: Xiaoyan Zhang, Human Services Software Inc  
Holly P Wald, XYZ Research Associates

The aim of this demonstration is to highlight the essential elements of the Information Age and the fundamentals of Internet Technology in relation to Performance Based Prevention Data Collection and Program Monitoring. It will include a presentation of the primary components of knowledge based information technology systems including its applicability at a variety of different levels. A review of the milestones Pennsylvania has accomplished in developing and implementing a Performance Based Prevention System (PBPS) since its inception in 1996. It will also include a demonstration of the Web-based PBPS and its potential benefits. There will also be a presentation of a data utilization conceptual model and methodology that enables practitioners and researchers to make use of the prevention data for purposes of program planning, management, improvement, monitoring, evaluation and statewide benchmarking. Specific examples of data utilization at the state, county and service provider levels will be provided, including the logic and method of linking the various components of the PBPS, i.e. needs assessment, planning and service data to generate meaningful evaluation reports that can help practitioners, funders and policy makers gain insights into the process and impact of prevention services. PBPS is currently utilized by the state of Pennsylvania, its 49 Single County Authorities (SCA) for drug and alcohol programs, and 150 prevention service providers for community needs assessment, service planning, service activity tracking, and performance monitoring. Two annual statewide performance based prevention reports and many subreports were generated using the data. When presented at two national conferences on prevention, PBPS received high recognition as an innovation on performance and outcomes based prevention management.

### **SESSION 575: MultiPaper**

**Room: Koko**

Sponsored by the Topical Interest Group on Extension Education Evaluation

#### **Evaluation Strategies for Diverse Cultural Settings**

Chair: Nancy Ellen Kiernan, Pennsylvania State University

Presenters: *Development of a User Friendly Quality Evaluation for a Native American Population*, Ruth Carter & Howard Rogers, University of Arizona

Improving the lives of youth has been a goal of Cooperative Extension since it began. Demonstrating our success continues to be one of our greatest challenges. Many best-practice models such as the one discussed in this

presentation come with evaluations that are elaborate, difficult to score and analyze. The presenters will demonstrate an ethical decision making program that uses a storytelling model to enhance the moral reasoning of adolescents. This program is being implemented on an Apache Reservation. They will discuss the evaluation currently in use and then present steps they are taking to maintain quality evaluation while making it more user friendly. They will make the case that this approach is greatly needed if Extension staff are going to be able to conduct their own on-site evaluations as well as help community-based organizations gain expertise in measuring outcomes of their important work with youth.

*Evaluation Issues for the Emerging Extension System in Russia*, Andrew V Novotorov, Iowa State University

This paper is a case study which is devoted to evaluation issues as follows: a review of pressing issues in the project on the development of Extension services in Nizhni Novgorod Oblast, Russia and how the evaluation literature helps us consider them; and special section on evaluation capacity-building in countries and/or institutions that don't yet have it, with links to standards and how they might be important. Then, a section that describes and critiques existing evaluation approaches of FAO and World Bank via their web servers with attention to examples from each that seem relevant to Russia. Through the purpose this paper the possible application of capacity development standards efforts will be constructed. Regardless of the nature of the evaluation being conducted in American institutional organizations, those Russian evaluators who concerned with evaluation development instruments could use the suggested proposal and apply to specific situations, especially in extension education. The objectives of this case study are to: (a) raise awareness of the issues to be tackled when institutionalizing evaluation, by illustrating what has worked well or not so well in practice in Russia; and (b) allow the Russian evaluators to contextualize this information to best suit their own circumstances. Conclusions and recommendations are the evaluation products of this proposal which has implications not only for Nizhni Novgorod Oblast, but for developing a model that can be used to expand Extension services throughout the whole Russian Federation.

*Strategies for Evaluating Latino Outreach Programs, with Attention to Language, Literacy, and Cultural Factors*, Barbara J Sawyer, Oregon State University

Latinos are a fast-growing minority population in the U.S., and are often underserved in Extension programming. As outreach efforts to this audience increase, program evaluation methodology needs to be developed and adapted to accommodate language, literacy, and cultural considerations. Reported in this session will be examples of evaluation strategies effectively used in three county demonstration sites in Oregon as part of a statewide effort to target recently-arrived Latino youth and families, primarily of Mexican descent. An examination of four core value clusters contrasting Latino and dominant U.S. cultures (attitude toward change, perception of self, relating to others, and forms of activity), which strongly influenced the Oregon evaluation design, will be summarized. Selected evaluation techniques used for needs/assets assessment, program improvement, and outcome measurement will be outlined, including a visual form of needs assessment, "trolling", a variation of the "listening post" approach, and adaptations of more familiar techniques.

---

## **SESSION 576: Panel**

**Room: Waimea**

Sponsored by the Topical Interest Group on Evaluation Managers and Supervisors

### **Linking Program Evaluation with Performance Measurement at the Federal Level**

Chair: Mary Ann Scheirer, Independent Consultant

Federal managers have recently completed their first performance reports required under the Government Performance and Results Act (GPRA). Performance reporting is one of the key steps in using data for performance management, requiring the creation and use of performance measures. There has been much discussion around the federal government about the potential roles for program evaluation in supporting performance management, but little is known about actual ways in which evaluation is being used. The specific objectives for this panel are: a) to describe multiple ways in which program evaluation is currently being used to support performance management in the federal government's implementation of GPRA; b) to discuss issues and recommendations from current practice for improving links between program evaluation and the use of performance measurement; and c) to encourage audience discussion about other examples and approaches to addressing the issues identified.

Panelists: *An Overview of the Use of Program Evaluation in Support of Performance Management: Results from a Survey of Federal Agencies*, Mary Ann Scheirer, Independent Consultant & Kathryn E. Newcomer, George Washington University

This presentation will report results from a research project to examine the use of program evaluation in support of performance measurement, funded by the PricewaterhouseCoopers Endowment for the Business of Government. Topics in the open-ended interview survey of federal agencies include the variety of ways in which program evaluation can be integrated with performance management, processes used to link program evaluators with program managers and GPRA staff, and the current capacities of federal agencies to link evaluation and performance measurement. Issues and recommendations derived from this survey will be discussed.

*Performance Measurement and Evaluation: Making Them Work Together in the Health Resources and Services Administration*, Roger B. Straw & Karen Pane, Health Resources and Services Administration

As with many Federal agencies, HRSA has been implementing the Government Performance and Results Act over the past several years. While it has been a difficult process for a number of reasons, excellent progress has been made both in both developing performance measures for reporting purposes and in getting program managers to understand and use those measures in managing their programs. On the other hand, historically HRSA has had a strong program evaluation program that depends largely on the Public Health Service 1% evaluation set-aside for funding support. The 1% evaluation funds have been instrumental in achieving the progress to date on developing and implementing performance measurement. HRSA is now in the process of using the 1% money as leverage in getting program managers to explicitly connect evaluations to their performance management strategies. Finally, the presentation will address external (to HRSA) pressures that support, and hinder, the ability to deliver on the promise of performance management.

*The Department of Education's Approach to Improving the Quality of Program Performance Data*, Lois Peak, US Department of Education

This presentation will describe the U.S. Department of Education's long and short term approaches to improving the quality of program performance data. The long-term approach involves working together with states to develop an Integrated Performance Benchmarking System (IPBS) for federal program data. The IPBS is a new Internet-based approach to the harvesting of information from states about federal program activities at the school and district level. The short-term approach involves having Assistant Secretaries and program managers formally review the quality of their data and attest to its quality. This short-term approach was implemented in the Department of Education's 1999 GPRA Performance Report, and will be expanded in the coming year. Both approaches are designed to better integrate routine performance reporting with periodic evaluation studies.

---

**SESSION 577: Panel**

**Room: Niihau**

Sponsored by the Topical Interest Group on Alcohol, Drug Abuse, and Mental Health

**Design and Analytic Issues in Multisite Studies Supported by the Center for Substance Abuse Treatment**

Chair: James M. Herrell, Center for Substance Abuse Treatment

CSAT conducts several multisite evaluations, some that approximate clinical trials, with manualized interventions, common data instruments, and consistent inclusion criteria, others with significant site variability in study populations, interventions, and measurements. Reduced cross-site standardization presents design and data-analytic problems. Panelists representing 3 continuing CSAT-sponsored, multisite studies, all departing from the ideal multisite model, will describe design and analytic issues faced in their studies, and will discuss the value added by prospective multisite studies of substance abuse treatment relative to after-the-fact, meta-analytic approaches. Evaluation findings will illustrate problems and solutions. The studies are: 1) The 4-site Cannabis Youth Treatment study, evaluating 5 treatment approaches for adolescents who abuse marijuana. 2) The Methamphetamine Treatment Program, a 7-site study comparing the effectiveness of a manualized intervention for methamphetamine abusers to 7 local "treatment as usual" approaches. 3) An evaluation of 50 residential treatment programs for substance-abusing,

pregnant or parenting women.

Panelists: *Overview of CSAT Multisite Studies*, James M Herrell, Center for Substance Abuse Treatment

Consistent with its mission of improving the availability and quality of substance abuse treatment services, CSAT supports several multisite studies of treatment interventions for specific populations. Currently active multisite evaluations focus on marijuana-dependent adults and adolescents, methamphetamine abusers, homeless families and others at risk for homelessness, adolescents who abuse alcohol, adolescents who abuse illicit drugs, pregnant and parenting women, adults and juveniles in criminal justice systems, persons with co-occurring disorders, women subjected to violence, and persons whose substance abuse behaviors put them at high risk for HIV infection or transmission. CSAT emphasizes evaluations of practical interventions conducted in real-world settings, and CSAT's multisite studies vary in the "tightness" of their designs. Current multisite studies will be contrasted along important design dimensions. Because coordinated, prospective multisite studies are expensive to conduct, it is important to maximize the value-added benefits of such studies, relative to after-the-fact meta-analyses of independently conducted studies.

*The Cannabis Youth Treatment Program: A 4-Site Study of Interventions for Adolescents Who Abuse Marijuana*, Michael L Dennis, Chestnut Health Systems Inc

The Cannabis Youth Treatment Program evaluates the effectiveness of 5 interventions designed to reduce marijuana abuse by adolescents. Treatment is provided at 4 sites: Farmington, CT, Madison County, IL, St. Petersburg, FL, and Philadelphia, PA. Chestnut Health Systems in Bloomington, IL, is the Coordinating Center. Adolescent are randomly assigned to one of 5 treatment conditions: 1) MET/CBT5 ù 2 individual Motivational Enhancement Therapy (MET) sessions and 3 group Cognitive-Behavioral Therapy (CBT) sessions; 2) MET/CBT12 ù 2 MET sessions and 10 CBT group sessions; 3) Family Support Network (FSN) ù MET/CBT12 group therapy plus additional case management, family support groups, and aftercare; 4) Adolescent Community Reinforcement approach (ACRA) ù 10 individual sessions with the adolescent and 4 sessions with care givers; and, 5) Multidimensional Family Therapy (MDFT) ù 12 to 15 individual family-focused sessions plus additional phone and case management contacts. The presenter will discuss the design and analytic and interpretive methods employed.

*The Methamphetamine Treatment Program: A 7-Site Study of the Matrix Model Approach to Treating Methamphetamine Abuse*, Richard A Rawson, University of California at Los Angeles

The 7-site Methamphetamine Treatment Program evaluates the effectiveness and potential for adaptation of the Matrix treatment model in community drug treatment programs. The Matrix model is a manualized outpatient treatment approach combining behavioral, educational, and 12-step counseling techniques. The study compares outcomes of the 16-week Matrix program with the usual treatment given at the programs. Each agency will treat 150 clients, selected using common inclusion/exclusion criteria. Data are collected for the following domains: Initial screening; Initial measures of drug/alcohol abuse; Demographics; Drug use history and severity; General/psychiatric symptomology; Drug craving; Psychiatric diagnosis; Exposure to domestic violence; Motivation or readiness-for-change; Quality of life; HIV high risk behaviors; Satisfaction with treatment; Cost effectiveness; Organizational change. This study, now in its second year, faces unique challenges in the use of 7 different "comparison" groups, and in the marked site variability -- including demographics and route of administration -- of treatment participants.

*A Cross-site Evaluation of 50 Residential Substance Abuse Treatment Programs for Pregnant and Parenting Women*, Kenneth Burgdorf, Caliber Associates

The 7-site Methamphetamine Treatment Program evaluates the effectiveness and potential for adaptation of the Matrix treatment model in community drug treatment programs. The Matrix model is a manualized outpatient treatment approach combining behavioral, educational, and 12-step counseling techniques. The study compares outcomes of the 16-week Matrix program with the usual treatment given at the programs. Each agency will treat 150 clients, selected using common inclusion/exclusion criteria. Data are collected for the following domains: Initial screening; Initial measures of drug/alcohol abuse; Demographics; Drug use history and severity; General/psychiatric

symptomology; Drug craving; Psychiatric diagnosis; Exposure to domestic violence; Motivation or readiness-for-change; Quality of life; HIV high risk behaviors; Satisfaction with treatment; Cost effectiveness; Organizational change. This study, now in its second year, faces unique challenges in the use of 7 different "comparison" groups, and in the marked site variability -- including demographics and route of administration -- of treatment participants.

**SESSION 578: Panel**

**Room: Hilo**

**Methods and Findings in CSAP's National Cross-site Evaluation of High Risk Youth Programs**

Chair: Soledad Sambrano, Center for Substance Abuse Prevention

This Symposium will summarize a) issues and approaches in research method; b) contributions to prevention theory, c) major prevention research findings; and d) lessons for prevention practice and policy that have emerged from CSAP's National Cross-Site Evaluation of High Risk Programs. The research base for these presentations is a 47-site evaluation that combines data from a self-report survey of over 10,000 youth, intensive site visits to participating programs, and individual records of program contact. Data were gathered at four points in time, with a retention rate of approximately 65 percent at the fourth time point. Data collection for the Cross-Site Evaluation was completed in Fall 1999, and analyses have produced important findings and lessons concerning prevention theory; program effectiveness across diverse age, gender, and risk configurations; program implementation, and the conduct of large, cross-site evaluations.

Panelists: *Conceptual, Methodological, and Interpretative Challenges in Cross-site Evaluation Research*, J Fred Springer, Evaluation Management and Training Associates Inc

In the context of CSAP's National Cross-Site Evaluation, examples of the challenges of cross-site research are identified and discussed. Many of the conceptual, data collection, and analytic procedures adopted by the study team provide potentially useful methods and techniques for other evaluators. The presentation will address: (a) the utilization of theoretical frameworks to encompass the diversity in program theories across sites and the challenges engendered by this need to be encompassing (b) the development of mediating and outcome measures, and a corresponding need for data reduction to simplify analyses, (c) the development of systematic field procedures to produce cross-site measures of program attributes, (d) the development of cross-site measures of program contact, (e) addressing comparison group contamination issues in the cross-site context, (f) presenting coherent findings of multiple studies using a phased model of effectiveness, and (g) consideration of different analysis approaches (e.g, pooled hierarchical modeling and meta-analytic approaches).

*Substance Use and Risk and Protection Outcomes Across 47 High Risk Youth Interventions*, Soledad Sambrano, Center for Substance Abuse Prevention

This presentation focuses on an analysis of outcomes from the cross-site study using both hierarchical linear modeling (HLM) and meta-analytic techniques. The CSAP cross-site evaluation was a longitudinal study of 47 programs, with youth ages 9-17 surveyed at four points in time; program entry, program exit, and 6 and 18 months after exit. The very large individual level sample (n=10,473) allows for pooled analysis of outcomes across age, gender, and racial / ethnic membership. The following topics will be discussed: (a) the identification of family bonding and academic self-efficacy as key mediating latent variables related to substance use using structural equation modeling, (b) immediate (baseline to exit) effects across the 47 sites and according to key demographic characteristics, and (c) longer term (3 and 4 point growth curve) intervention effects. Select results of HLM analyses will be used to identify promising explanations of differences in effects on substance use.

*Piecing Together the Prevention Puzzle*, Elizabeth Sale, Evaluation Management and Training Associates Inc

Because of the large number of programs and youth participating in the CSAP Cross-Site evaluation, this study offers a unique opportunity to assess the effects that differential program characteristics can have on program outcomes. This presentation focuses on findings concerning the relation between program design, program implementation, participant engagement, program contact (dosage) and the mediating and substance outcome measures for the CSAP cross-site study. Program design is measured with several constructs, including funding initiative type, planned

program strategy types, and characteristics of the setting in which the program is delivered. Program implementation is assessed through two multiple item measures: the degree of supportive management and the extent of results-oriented management. Participant engagement and dosage are measured through participant self-report and program records. Major findings are summarized concerning the relation between these variables and outcomes. Differing patterns of effects between risk and protection measures and substance use outcomes provide a basis for further analysis.

Discussant: Soledad Sambrano, Center for Substance Abuse Prevention

## **SESSION 579: Panel**

**Room: Puna**

Sponsored by the Topical Interest Group on Evaluation Use

### **Ethics and the Role of the Evaluator: What Are the Obligations and Limits?**

Chair: Leonard Bickman, Vanderbilt University

Basic experimental researchers typically identifying a situation, manipulate a variable within it, and then evaluate the outcome of that manipulation. It is assumed that the researcher is vested in discovering some truth about this situation. However, in program evaluation it is believed that stakeholders are too biased to truly evaluate a program objectively. Therefore an external evaluator is often used. This symposium will focus on the assumptions behind the use of an external evaluator and the characteristics of "objectivity". We will focus on the roles and moral obligations of the evaluator as an observer and participant and how these are characterized in formative and summative evaluation. These topics will be presented first as theory and then through a description of a current evaluation of a state project to integrate mental health services for children with serious emotional disturbances (SED).

Panelists: *Research Vs. Evaluation: Ethical issues*, Leonard Bickman, Vanderbilt University

Traditional research assumes that the researcher can be objective about his or her own research and that findings will be reported without bias. In contrast it is assumed by some perspectives on evaluation that very involved stakeholders should not evaluate their own programs. However, some evaluation approaches, e.g. empowerment evaluation, take the opposite perspective. When is it appropriate for heavily invested stakeholders to evaluate their own programs? Can we depend on the training and ethical standards of evaluators to guard against possible biases. Do basic research have the objectivity to conduct their own studies when they are also heavily invested. These and other questions will be considered during the symposium.

*Obligations of the Evaluator in Summative and Formative Evaluation*, Stephanie Reich, Vanderbilt University

'Objectivity' is often considered the gold standard of research and evaluation. Biases are viewed as hindrances to true discovery. But what happens when ethical obligations and objectivity clash? This talk will focus on the ethical issues of summative and formative evaluations. Specifically, how the evaluator balances objective research with moral responsibilities to stakeholders. The strengths and weaknesses of active participation and observation will be discussed and how these roles affect objectivity and the evaluators' responsibility toward the evaluation.

*A Model Intervention-oriented Evaluation of a Mental Health Service System for Youth*, Denine A Northrup, Vanderbilt University

Many evaluators strive for stakeholders to actively use the information gained through evaluation to make decisions. However, often stakeholders are unsure of how to use the findings to improve their program. In this presentation, an ongoing evaluation of a mental health service system for youth with serious emotional disturbance illustrates the use both formative and summative approaches to evaluation in action. Multiple sources of information were collected including youth and family semi-structured interviews, inter-organizational network analysis, and community-level surveys and focus groups. We will demonstrate how these data sources are useful in both formative and summative evaluation purposes. For the formative evaluation, multi-level feedback loops for performance monitoring provide 1) clinical information about individual youth in the program, 2) service outcome information, and 3) system performance information. From a summative standpoint, the youth-, family-, and provider-level data are used in longitudinal

analyses to address summative outcomes on youth and family, services and system level changes.

*Strengths and Challenges of an Intervention-oriented Evaluation*, Beverly B Mahan, Vanderbilt University

This next presentation will build on the evaluation process presented in the previous presentation and discuss the strengths and challenges experienced in implementation of this intervention-oriented evaluation. The strengths of this approach focus on increasing the communication between stakeholders, allowing efficient access to and results from standardized clinical measurements, ongoing status reports for all stakeholders, and identification of barriers in a timely manner that can be addressed in program implementation. Some of the challenges to be discussed include: Are we a member of the decision team or do we merely facilitate the process? Who should have access to the evaluation data? What type of information and in what form would be useful for stakeholders? What is the role of the evaluator in ensuring that the program responds and adapts to evaluation feedback or results? All of these strengths and challenges speak to the feasibility of concurrently conducting both formative and summative evaluations.

Discussant: Dianna L Newman, State University of New York at Albany

**SESSION 580: Panel**

**Room: Kohala**

Sponsored by the Topical Interest Group on Collaborative, Participatory & Empowerment Evaluation

**Collaborative Evaluation of Early Childhood Initiatives in North Carolina**

Chair: Rita G O'Sullivan, University of North Carolina at Chapel Hill

NC Smart Start is a comprehensive early childhood initiative that provides programs and services at the county level based on local needs and resources. The overall goal is to provide services that will prepare young children to be as successful in school as possible. Typical services provided by a local partnership include: programs to improve childcare homes and centers, scholarships for early childhood teachers, childcare subsidies so parents can continue to work, nutritional programs for families and children, hearing, vision and dental screenings, parent education classes and resources, and translation services for Spanish-speaking families. This approach has led to Smart Start's success and overall the program is widely popular. It is not a particular school or a particular curriculum or one particular service. This creates a tremendous challenge and need for evaluation of each individual program as well as their effects collectively within each county.

Panelists: *An Overview of How Collaborative Approaches Promote Evaluation with Community-based Programs for Young Children and Their Families*, Rita G O'Sullivan, University of North Carolina at Chapel Hill

Collaborative evaluation engages key program stakeholders actively in the evaluation process to the extent that they are able. Further, the collaborative approach promotes expanding the evaluation expertise of program staff, so that they become progressively better at evaluation. In this manner, overall evaluation findings are strengthened as stakeholders' openness to utilize evaluation findings increases. The collaborative approach assumes that project staffs and grantees share responsibility and decision making for the evaluation. This approach has yielded enhanced reporting that has presented progressively stronger evaluation information. Additionally, over time, projects assume greater responsibility for their evaluation activities.

*The Collaborative Process of Determining Community-wide Outcomes for Early Childhood Initiatives*, Michele Rivest & Tamara Dempsey-Tanner, Orange County Partnership for Young Children

The Orange County Partnership is a pioneer county in the 5-year-old North Carolina Smart Start Initiative and as such, it is time to begin examining the impact of the various early childhood programs on the county as a whole. The collaborative process of determining county-wide outcomes tied to the Partnership's goals and objectives in a relatively transient and increasingly diverse community will be discussed, as well as the process of the overall evaluation, from the perspective of the Executive Director of the Orange County Partnership for Young Children.

*Evaluation Planning as a Collaborative Process*, Tamara Dempsey-Tanner, Orange County Partnership for Young Children

The evaluation emphasis during the 1999/2000 year focused primarily on working with the Orange County Partnership Program Manager to develop and/or fine tune detailed evaluation plans for their programs that tie their program activities to Partnerships' objectives and specify ways to monitor activity and outcome accomplishments. Through individual site visits and contractor networking meetings, technical assistance was provided for agencies that needed assistance identifying and/or creating appropriate instruments (e.g., outcome measures, surveys, and/or summarizing information). With evaluation plans and data collection systems in place, the external evaluation team was able to work more closely with the Partnership to provide technical assistance to strengthen their evaluation expertise.

*Using Community-wide Outcomes To Create a Database of County Indicator Information*, Anne F D'Agostino, University of North Carolina at Chapel Hill

As part of their strategic planning process, evaluators have worked with the Orange County Partnership for Young Children to develop new systems for tracking county-level benchmarks for their Partnership. This includes development of a Partnership-friendly database that OCPYC staff can manage to track important county-wide indicators without elaborate and costly outside consultants. The database incorporates: 1) both primary and secondary statistical data, 2) the most recent available current data and comparison data for 1998 or prior, and 3) the definition of each data element and identification of specific sources for future collection.

*Evaluation of Early Childhood Projects Designed for Latino Families*, Marisol Jiminez, University of North Carolina at Chapel Hill

The rapid growth of the Latino population in Orange County has created a tremendous need among families of young children in the community who speak little or no English. The Partnership's goals in working with the Latino population are to 1) provide information to the community about funds for indigent medical care, 2) increase access and utilization of health social service agencies, increase the knowledge and understanding of health and social issues related to their Latino preschoolers, 3) increase knowledge and access to community resources, 4) increase parents' English language skills, and 5) increase literacy. The process of planning the evaluation of projects targeted to this population will be discussed, together with the challenges of gathering information from hard-to-reach participants.

Discussant: Rita G O'Sullivan, University of North Carolina at Chapel Hill

**SESSION 581: MultiPaper**

**Room: Kona**

Sponsored by the Topical Interest Group on Alcohol, Drug Abuse, and Mental Health

**Evaluating Programs for Persons with Severe Mental Illness**

Chair: Cheryl A Hosley, Wilder Research Center

Presenters: *Strategies for Measuring Accessibility to Mental Health Programs*, Cheryl A Hosley, Wilder Research Center

Accessibility is an important dimension of mental health program quality. However, this dimension is often difficult to accurately assess as most program evaluations are based on reports from clients who have successfully accessed program services. This paper will summarize the results of an examination of accessibility undertaken by the Wilder Research Center. Several strategies for measuring accessibility will be presented. Three categories of measures will be included: (a) information extracted from program records; (b) ratings of accessibility dimensions by program clients; and (c) use of population and market share estimates. Data will be presented summarizing the results of these measures as one component of the evaluation of a community-based mental health clinic. These data will be used to illustrate the nature of accessibility in mental health programming and the relationships between accessibility and other program evaluation elements, including client satisfaction and service effectiveness.

*Evaluation of the San Francisco Assertive Community Treatment Model (Three Single-point of Responsibility (SPR) Programs): Preliminary Cost Analysis and Findings*, Davis Y Ja & William Hargreaves, California School of Professional Psychology

During 1997, Community Mental Health Services (CMHS) of the Department of Public Health, City and County of San Francisco, initiated implementation of three pilot Single-Point-of-Responsibility (SPR) programs targeting high cost mental health service consumers with severe mental illness. The primary goals of these programs were to improve consumer outcomes and reduce overall costs through utilizing assertive community treatment (ACT) program approaches, capitated rates, and a unique partnership between CMHS and each program. This paper presents the methodology and initial findings of an evaluation of service utilization and cost. Initial cost findings indicate that consumers enrolled at least 100 weeks prior to enrollment in the SPR were using services at a rate of almost \$60,000 per year during the two weeks prior to enrollment. Following enrollment, this dropped to about \$40,000 and is now approaching \$22,000 towards the end of two years after enrollment. The six teams involved converged in reducing weekly capitated cost by month 20. Compared to non-SPR enrollees, consumers in the SPR's were among the highest cost users of the system prior to enrollment. However, after 20 months, SPR enrollees costs were twice as low as the non-SPR enrollees.

*The Measurement of Consumer Outcomes in Ohio's Mental Health System: Empirical Evidence From Cuyahoga County*, Richard E Njoku, Cuyahoga County Community Mental Health Board

Concerns regarding critical issues of access, quality, and cost, compelled the state of Ohio to implement a state-wide approach to assessing consumer outcomes in its publicly-funded mental health system. A state outcomes task force (OTF) recommended the measurement of consumer outcomes from the four domains of clinical status, quality of life, functional status, and health and safety. Two sets of home-grown instruments were developed to measure outcomes among the adult and the youth consumers. This paper discusses the implementation of Ohio Outcomes Initiative among 20 agencies in a large metropolitan county in the state. It provides empirical evidence on the four critical outcomes domains for adult consumers, and assesses changes in behavioral problems, level of functioning, hopefulness, among youth consumers. Finally, it discusses the use of outcomes data for making critical decisions in treatment planning and care management, quality improvement, and accountability for public resources.

*Factors Affecting Successful Long-term Care to Community Placement for Individuals with Serious Mental Illness*, Saumitra Sengupta & Kevin McGirr, San Francisco Community Mental Health Services

There have been a wide variety of attempts to investigate institutional recidivism among the seriously mentally ill. The results appear to be inconclusive with respect to clearly delineating consumer success in the community. Most of the literature is focused upon re-hospitalization prediction. This paper will attempt to link individual factors, length of stay and community service variables with length of tenure in the community. Community tenure for this study is defined as the length of time elapsed between discharge from long-term care facilities and re admission to one of these facilities. Any acute care or psychiatric emergency episodes are tracked but not considered as terminating events. The study examines the service utilization pattern of a cohort of 340 consumers of mental health services in San Francisco.

---

**SESSION 582: Panel****Room: Ewa**

Sponsored by the Topical Interest Group on Pre-K-12 Educational Evaluation

**Building Capacity Among State- and Local-level Partners: Experiences from the Evaluation of North Carolina's Smart Start Early Childhood Initiative**

Chair: Kathleen Y Bernier, University of North Carolina at Chapel Hill

North Carolina's Smart Start Early Childhood Initiative is a statewide, community based, public-private, collaborative effort to provide early childhood education and services for children birth through five years and their families. Over the past six years, the Frank Porter Graham Center (FPG) at the University of North Carolina at Chapel Hill (UNC-CH) has conducted the evaluation of Smart Start. Since evaluation is needed at the state-level, county-level, and program-level, FPG has partnered with Smart Start and program staff, not only to help conduct evaluation, but also to help build capacity among the partners.

Panelists: *Building Capacity Among State-level Partners: Experiences from the Evaluation of North Carolina's Smart Start*

*Early Childhood Initiative*, Donna M Bryant, University of North Carolina at Chapel Hill

Our state partners include Smart Start administrators, state agency staff and legislators. Building capacity among these groups has meant advocating for the importance of evaluation, partnering to formulate state-level evaluation plans, and helping partners understand appropriate uses of results. Dr. Donna Bryant, who leads the FPG/UNC-CH Smart Start Evaluation Team, will discuss the challenges and strategies used in building evaluation capacity at the state-level in a political environment.

*Building Capacity Among County-level Partners: Experiences from the Evaluation of North Carolina's Smart Start Early Childhood Initiative*, Kathleen Y Bernier, Joy Sotolongo & Karen Taylor, University of North Carolina at Chapel Hill

Our county-level partners include county Smart Start administrators, staff evaluators and external contract evaluators. Building capacity among these partners has been a two-way street -- FPG has been able to contribute evaluation expertise and resources and county partners have helped with the development and implementation of evaluation plans and specific measures. Dr. Kathleen Bernier, who leads the FPG Smart Start Evaluation Assistance Team, will discuss the challenges of meshing expertise and implementation. The many successful strategies that have been used include formally utilizing the evaluation expertise and experience of county evaluators; one-on-one contacts with county partners through visits, telephone and email; quarterly statewide evaluation meetings; annual seminars and conference; and web site and web data collection.

*Building Capacity Among Smart Start Boards of Directors and Service Program Staff: Experiences from the Evaluation of North Carolina's Smart Start Early Childhood Initiative*, Eva Hansen, Rhode Bicknell & Linda Blanton, Cumberland County Smart Start Partnership for Children

Our county-level partners have been primarily responsible for building evaluation capacity among other county Smart Start and early childhood education and service staff. Eva Hansen from the Cumberland County Smart Start Partnership for Children will discuss the challenges of advocating for evaluation with the county's Smart Start Board of Directors and county service staff. Ms. Hansen will detail the numerous successes her partnership has experienced in making evaluation a primary focus of Smart Start and of the larger county service system, and the primary role that has evolved for her Smart Start partnership in county-wide evaluation.

**SESSION 583: MultiPaper**

**Room: Honolulu**

Sponsored by the Topical Interest Group on Program Theory and Theory-driven Evaluation

**Using Social Science Theories To Improve Evaluation Practice**

Chair: Laura E Gooler, Claremont Graduate University

Presenters: *Emerging Issues of the Ecological Model: Utilizing Health Behavior Theory To Provide a Theoretical Framework for Evaluation Design*, Debra J Holden & Leah M Ranney, North Carolina State University; and Rena L Large, National Education Association Health Information Network

A national breast and cervical cancer education project was designed and piloted by the Health Information Network (HIN), a not-for-profit organization operating within the National Education Association (NEA). The project was based on the theory of the ecological model and provided training to state affiliates and local school employees on developing health interventions at each level of the model. Follow-up evaluation measures were collected from a random sample of NEA members in both an intervention and comparison county. Qualitative and quantitative measures were collected to measure the perceived influence of different levels of the ecological model on a participant's decision to get a mammogram. Emerging issues relevant to the ecological model will be discussed, as well as lessons learned as derived from the evaluation findings. Implications for utilizing health behavior theory as a strategy for providing a theoretical framework for evaluation study will also be addressed.

*Enhancing Learning with Interactive Activity: Using Evaluation Data to Support Education Theory*, Anthony C Gallagher, American Red Cross, and Treva A Stack, Jacob France Center

We used archival evaluation data to test Knowles (1988) assertion that relevant, applicable, interactive, experiential education enhances adult learner outcomes. All agency employees received two hours of Workplace HIV/AIDS education, covering HIV transmission and workplace legal issues. Supervisors had an additional hour for an interactive role-play exercise to explore social/legal aspects of AIDS at work. 180 supervisors attended the 3-hour sessions; another 130 attended the two-hour sessions intended only for line-staff, creating an unplanned opportunity to test Knowles' theory. Supervisors did benefit from the interactive exercise. After controlling for pre-scores, post-scores for role-players were significantly higher for Knowledge of Legal Rights/Responsibilities ( $p < .0005$ ), Attitudes about AIDS in the Workplace ( $p = .019$ ), and Behavioral Intentions ( $p = .054$ ). There were no differences for topics covered in equal depth by both types of session. Although this exercise was non-experimental, it demonstrated that evaluation data can support and advance education theory. Specifically, it provides evidence that interactive exercise can positively benefit the learner.

*Community Health Improvement Initiatives: Assessing the Degree to Which Community Organizations Integrate and Institutionalize Program Approaches within Their Own Organizations – What Can Organizational Theory Tell Us?*  
Ann P Zukoski, University of California at Berkley

Community health coalitions represent a powerful means of addressing local health problems. It has been argued that the extent to which coalition efforts lead to sustainable and long term improvements in community health status will be largely dependent upon the adoption and integration of new programs, activities and system level changes within its member organizations. Assessing and understanding the factors, which influence a coalition member to adopt and integrate changes within their own organization, is very challenging. This paper will explore these issues by presenting findings from an evaluation of the National Community Care Network Demonstration Program funded by the W.K. Kellogg Foundation. The paper will propose a model for assessing institutional changes within coalition member organizations based on interorganizational theory and present key findings of factors influencing integration based on quantitative (cross sectional survey data) and qualitative analyses using a sample of 54 health care organizations.

## **SESSION 584: Roundtables**

**Room: Kahuku**

### **Exploring Issues of Theory, Use and Methodology**

*(This session includes two 45-minute rotations of roundtables. The Host will ask the tables to rotate at 4:15.)*

Host: To Be Announced

### **Roundtable A (First Rotation): Improving the Dissemination of Findings from Foundations**

Presenters: *If a Tree Falls...: How Can a Foundation Best Share the Lessons It Learns from Evaluations?* Victor Kuo, The James Irvine Foundation

Does an evaluation make any noise if no one is there to hear it? Too often evaluators go unheard -- their reports unread, their recommendations ignored. Realizing that an evaluation is only as useful as it is used, The James Irvine Foundation made it a 2000 priority to improve the way it communicates the lessons learned from its evaluations to both external and internal audiences. In this roundtable, staff members of Irvine's Evaluation and Communication offices will share their plans for improving the dissemination of program evaluation findings, and invite other evaluators to share the lessons they've learned about sharing lessons. The discussion will focus on a number of questions: How can foundations incorporate evaluation and communications plans into the design of grantmaking programs? How to make sure audiences are brought to the "forest" to hear the "tree" fall? How can foundations connect the traditionally distant functions of evaluations and communications?

*Expanding the Learning Loop: Dissemination of Foundation Results and Insight*, Sharon L Sturtevant, Pew Charitable Trusts

Interviews with foundation staff, grantees and evaluators reveal that despite an eagerness to share and learn from each other, and an institutional emphasis on strategic grantmaking and accountability, foundations are doing relatively little to share the knowledge they gain from evaluation work. Whether 'open books' (professing to share everything),

'black boxes' (not pretending to share anything), 'transformers' ('sanitizing' internal reports to make them acceptable for external publication) or working on 'parallel tracks' (designing documents as internal or external from the outset), foundations, like many sectors, are struggling to find a balance between public accountability and internal responsibilities. Beginning with the belief that dissemination of evaluation is an important responsibility, participants will discuss viable opportunities and alternatives for dissemination. In particular, how can we use evaluation findings to improve program design and build capacity beyond our own walls without having the institution (or evaluation unit) feel like it is breaking internal confidences?

### **Roundtable B (First Rotation): Conflict in the Evaluation Context**

Presenters: *Conflicting Demands of Evaluation Users: A Bureaucratic Case Study*, Todd M Franke, Christina A Christie, Michelle T Parra & Walter Furman, University of California at Los Angeles

This paper will discuss the conflicting demands of evaluation users. The California Office of Child Abuse Prevention (OCAP) commissioned an evaluation for two purposes: 1- as a progress report to the federal government and 2- to inform internal agency staff of program improvements. The evaluation, then, has two primary user groups, the federal government and the state government office. Within the state office we have specified primary users however the federal government is a generalized user group. The informational purposes and needs are different for each group and, as a result, the evaluation team had to negotiate the conflicting demands of both within one evaluation. To promote use among each group it was critical to design an evaluation that determined, within each sampled county, the unique impact of OCAP programs on child abuse prevention amid several other programs contributing to this same goal. The user conflicts that arose and the resolution of these conflicts will be discussed.

*The Role of the Evaluator in a Conflicted Evaluation Setting*, Liliana Rodriguez-Campos & Rigoberto Rincones-Gomez, Western Michigan University

This paper reviews an evaluation of a neighborhood financial assistance program for two targeted economically distressed areas of Kalamazoo, Michigan. In order to address multiple perspectives about the program a participatory evaluation approach was used. This approach was useful to capture differences among stakeholders in the program, as a mechanism for addressing conflict that was known to exist. The evaluation illustrates steps needed to gain stakeholder ownership and support for evaluation findings by helping them gain a sense of appreciation for the evaluation process. Moreover, this paper provides a systemic way for organizing, applying, and reporting evaluation findings that move participants toward an effective use of the results in a conflicted environment. This paper illustrates how communication among evaluators and participants, and continuous evaluation helps ensure program development and eventual success.

### **Roundtable C (First Rotation): The Many Faces of Evaluation Use and Influence**

Presenters: *Exploring "Constitutive Effects" of Evaluation*, Peter Dahler-Larsen, Odense University and Hanne Krogstrup, Institute for Sociale Forhold og Organisation

Although our typology describing various forms of utilization of evaluation has been expanding over the years, even a broad typology of "instrumental", "enlightenment", "strategic", "tactical" and "symbolic" use does not fully exhaust how evaluation sometimes influences social reality in subtle ways. This motivates us to explore the concept of "constitutive effects" of evaluation. These include how certain standards, criteria, and world views become taken for granted. Goal displacement is one such effect, but there are many more. Institutionalization processes help these effects travel through time and space and extend beyond the control of a user of evaluation. This is illustrated with empirical examples of the use of performance indicators in Scandinavia. We find constitutive effects on the content of work, on the timing of program activities and on social relations and identities of people involved in programs. Finally, the paper offers speculations about factors which may influence the amount and form of constitutive effects of evaluation.

*Utilization Assumptions Implicit in Theoretical Approaches: A Comparative Examination*, Christina A Christie, Janet Cadiz, Anne E Marshall, Eric Barela & Stephen Zuniga, University of California at Los Angeles

It is a commonly accepted notion that evaluations are conducted for an audience that is particularly interested in its

results. Patton (1997) states that "in any evaluation there are many potential stakeholders and an array of possible uses" (pg. 20). Thus, what distinguishes one theoretical approach from another is how potential stakeholders are defined, the importance placed on the role of the stakeholder, and the significance placed on use. This paper will identify the primary utilization assumptions of five noted evaluation theories; House, Patton, Stake, Weiss, and Guba and Lincoln, all whom have written about utilization in some context. It will compare the primary utilization assumptions of these evaluation theories, and a comparative matrix will be presented. Hypothetical case examples will be used to illustrate each utilization assumption.

**Roundtable D (First Rotation): Success Measures Project: Tools for Building Evaluation Capacity at the Community Level,**  
Susan B Naimark & James C Alexander, Development Leadership Network

We propose to walk participants through a participatory evaluation process developed to increase the capacity of community-based organizations and the residents they serve. The Success Measures Project engaged hundreds of practitioners throughout the U.S. in a two-year participatory process of identifying benefits (impacts) of their work in community development, and indicators for measuring these benefits. Out of this process the Success Measures Guidebook was developed, containing 44 indicators for measuring impact of housing, economic development and community building initiatives. SMP's primary objectives are: to increase evaluation capacity at the community level; and to enable community-based development organizations to more effectively communicate the impacts of their work to a broad range of audiences. It is unique as a model for impact evaluation in that:

- Indicators were developed by the community-based practitioners who will use them;
- The Guidebook is an effective, hands-on tool for bringing impact evaluation to community-based organizations.

Format/Logistics:

Presenters will describe project objectives and background which led to the development of the Success Measures Guidebook, then utilize an interactive format to introduce sample indicators. Finally, an interactive exercise will be used to demonstrate ways in which professional evaluators can support this approach in communities.

**Roundtable A (Second Rotation): The Larger Role of the Evaluator: Choosing an Approach and Increasing Use**

Presenters: *The Impact of Evaluation Approach on Knowledge Gain, Credibility, and Utility of Program Evaluation Findings,*  
Shelley A Potts, Arizona State University

*Completing the Equation: Building the Capacity of Decision-makers To Use Evaluation Findings,* Tenzing Donyo, The California Endowment

**Roundtable B (Second Rotation): Components to a Longitudinal Evaluation of the Lockheed Martin/University of Central Florida Academy**

Presenters: *Focus Group Results: Part of a Longitudinal Evaluation of the Lockheed Martin/University of Central Florida Academy,*  
Debra Daniels, Michael Hynes, Patrick J Moskal, Nancy Lewis, Gayle Sitter, Mariann Schmutte, Judith Johnson & Carine Strebels Halpern, University of Central Florida

Quantitative methods traditionally have dominated an evaluator's tool kit. Standardized tests, randomized experiments, forced-choice surveys, and structured interviews have been the methods of choice for evaluating the impact of social programs. Researchers have argued that methodological rigor, achieved through quantitative methodology, produces valid data. Why are results and recommendations from such evaluations, however, often not used to improve service delivery or to impact policy?

This study investigated the relationship between evaluation approach and the strength of its knowledge claims, its credibility, and its usefulness. Specifically, the researcher investigated whether or not quantitative, qualitative, and mixed-method evaluations produced different kinds and amounts of knowledge gain, different levels of credibility, or suggested different types of use. This presentation will address the study's methodology, findings, and significant implications for evaluation research and practice. Findings extend current views regarding evaluation use and suggest a new approach for the training and development of evaluation professionals.

*Strengthening the Capacity of Observations: Part of a Longitudinal Evaluation of the Lockheed Martin/University of Central Florida Academy*, Nancy Lewis, Debra Daniels, Gayle Sitter, Mariann Schumde, Pat Moskal & Michael Hynes, University of Central Florida

Capacity building as it relates to evaluation has received heightened attention in recent years. However, the focus of resulting discussions has been limited towards building capacity to conduct evaluation. Remarkable progress is being made with respect to building capacity of individuals and communities impacted by programs to engage in evaluation activities. Yet, little or no attention is paid around building capacity of decision-makers to use evaluation findings. This paper argues that building capacity to conduct evaluation is necessary, but not sufficient for achieving successful evaluation. This shortcoming explains one of the reasons why notable works in evaluation go unused by decision-makers. The paper contributes to our understanding of the pathway from building evaluation capacity to achieving successful evaluation. Furthermore, it urges evaluation professionals to work on building capacity of decision-makers (funders and program managers) to understand, absorb and use evaluation findings.

### **Roundtable C (Second Rotation): Evaluation Use in International Contexts**

Presenters: *How Much Does Evaluation Matter? Some Examples of the Utilization of the Evaluation of the World Bank's Anti-corruption Activities*, Mita Marra, The World Bank

This paper offers empirical evidence of the utilization of evaluation findings of the World Bank Institute's (WBI) efforts to help reduce corruption in Tanzania and Uganda. The initiatives are part of the World Bank program to curb corruption in developing countries. This analysis focuses on the midterm evaluation of WBI's anticorruption activities in those countries, entitled 'WBI's Anticorruption Initiatives in Uganda and Tanzania.' The study shows, through a series of examples, how evaluation has been used in both an instrumental and an enlightenment fashion by program designers and implementers. It also shows that links between knowledge generation and utilization are seldom clear and direct, and that specific information cannot be always isolated as the basis for a particular decision. Nevertheless, the examples highlight that utilization brings about change in program design and implementation.

*Evaluation Utilization Issues in Cross-border Education: A Case Study of North-South Educational Links in Ireland*, Roger Austin, University of Ulster

Since 1997, University of Ulster staff has been evaluating cross-border K-16 links. Two projects have been funded by the governments of Northern Ireland and Republic of Ireland as well as two commercial sponsors. The projects are tasked to develop teachers' ability to integrate instructional technology (ICT) into the curriculum, sustain ICT practice and examine the role ICT can play in extending pupils' understanding of different cultures and perspectives.

Program evaluation efforts have generated many issues, four which will be discussed in the session:

1. the interplay of internal and external evaluation where the sharp lines which normally separate the outsiders and the insiders have been blurred;
2. finding 'suitable' ways of evaluating "improved understanding of other perspectives" amongst project participants is a highly charged issue in the faltering peace process in Northern Ireland;
3. searching for a negotiated evaluation model has challenged partners' use of evaluation language and methods;
4. the public-private funding for the project has generated a creative tension about the purposes and audiences of evaluation.

### **Roundtable D (Second Rotation): Theory and Paradigm: Examining the Foundations that Drive Practice**

Presenters: *From Program Theory to Constructivism: On Tragic, Magic, and Competing Programs*, Peter Dahler-Larsen, Odense University, Hanne Krogstrup, iborg University

Strong ideas about programs persist in many policy fields regardless of an evaluator's formalistic "test of the program theory". A more realistic and sensitive approach to program theory beyond simple verification/falsification may begin with what Pawson and Tilly (1997) call context - mechanism - output (CMO) configurations; but the involved contexts are under ongoing social (re-)construction. An elaboration of program theory terminology (especially "moderator variables" in "contexts") actually leads to a constructivist idea: the ongoing social construction of reality - represented

by "moderators" - determines not only the fate of programs, but also the truth value of a program theory in a given context. Tragic, magic and competing programs affect their own moderators in different ways. In real life, people connected to programs may care for either programs, program theories, or moderators for different reasons. Implications for evaluation practice are discussed.

*When Worlds Collide: The Use of Multiple Methods and Cross Paradigms in Evaluation*, Cynthia M Hewitt-Gervais,  
Florida Gulf Coast University

This presentation will provide participants with the advantages and limitations of an evaluation design that includes both interpretivist (more specifically, constructivism) and post positivist paradigms' beliefs and practices. An example of this design is an evaluation of a staff development grant (5 STAR) that utilized a team whose members espouse different research paradigms. This illustrates our support for the use of multiple paradigms together in scientific research or evaluation, IF the team includes multiple collaborators, who each have specific assignments which originate from a unique paradigm. Because each collaborator approaches the research with his/her own paradigm bias, neither could individually do justice to utilizing both paradigms. This implies that the use of a team of researchers from differing perspectives presents the opportunity to capitalize on the advantages of each perspective providing a broader picture of the evaluation's situational context without violating underlying assumptions.

**Roundtable E (Second Rotation): Program Evaluation Partnerships: Grantor and Grantee**, Ann L McCracken, Health Foundation of Greater Cincinnati and Diane M Wright, Cincinnati Restoration Inc

Evaluation can and should be a collaborative process between grantor and grantee. For both the grantor and the grantee the evaluation answers the question "Do we make a difference?" Both entities are vested in the sustainability of ongoing projects beyond the period of grant funding. Both debate the efficacy of allotting resources for services versus evaluation. This paper presents program evaluation from the perspective of the grantor, the Health Foundation of Greater Cincinnati (HFGC), and the grantee, CRI, an agency that provides services for persons with severe mental illness. This paper outlines the Foundation's evaluation efforts in community betterment, maximization of learning and appraisal of the Foundation's grantmaking process. Maximization of learning addresses the question "Does this particular grant make a difference in the lives of those served by the project?" CRI responds from the agency perspective sharing the possibilities and perils of the mandated HFGC grant evaluation.

**SESSION 585: Panel**

**Room: Oahu**

Sponsored by the Topical Interest Group on State and Local Government

**Evaluation in the Federal Government: Funding and Employment Opportunities at NSF with Special Emphasis on Asian Pacific Islanders**

Chair: Jane T Stutsman, National Science Foundation

This session is focused on funding and employment opportunities in science, mathematics, engineering and technological (SMET) at the National Science Foundation (NSF). The emphasis by the first speaker will be on the need for more evaluative research studies in SMET education among the sub-groups within the Asian American Pacific Islanders (AAPI) communities. This is in response to a special White House Initiative on AAPI's. The second speaker will describe studies funded by NSF on Asian Americans and planned future studies. The third speaker, an awardee of NSF, will describe some of her findings in the education field regarding AAPI's. The Chair will conclude the session with a description of NSF's goals and missions, employment opportunities both at NSF and in NSF-funded projects, with a focused outreach to Asian Americans, Native Hawaiians and Pacific Islanders.

Panelists: *Bridging the Information Gap on Asian American Pacific Islanders' Science and Math Achievement*, Deh-I Hsiung,  
National Science Foundation

Deh-I Hsiung will highlight issues and opportunities for Ph.D students and faculty concerning dissertation topics, with a focus on research regarding Asian American Pacific Islanders' achievements and/or barriers to achievement in SMET. Most of the data collected by researchers and the Federal Government have been centered around women and minorities who are under-represented in science and mathematics. However, there have been few studies on

S.E. Asian Americans and Pacific Islanders who are under-represented. President Clinton signed Executive Order 131125 in June 1999, to improve the quality of life of Asian Americans and Pacific Islanders (AAPIs) through Federal programs where they are underserved. One goal is to "foster research and data collection for AAPI ... subpopulations". To this end, there will be a brief discussion of some of evaluative research on AAPI we know, some of the research areas we don't know and pointers to some of the areas we need to know.

*Asian American Pacific Islanders in the Science and Engineering Workforce*, Nirmala Kannankutty, National Science Foundation

Nirmala Kannankutty will highlight some of the data that NSF currently collects on Asian Americans Pacific Islanders in the S&E workforce. The presentation would include a profile of Asian American Pacific Islanders who are scientists and engineers, including their educational attainment, occupation and employer characteristics, and demographic characteristics.

*Student Participation in Science and Mathematics by Racial-ethnic Groups*, Frances K Stage, Indiana University at Bloomington

Frances Stage, currently, Professor former Senior Fellow at NSF, has directed several NSF-funded projects. Her research focuses on student participation in SMET majors using the High School and Beyond and National Education Longitudinal Study data bases as well as qualitative studies. Stage will be reporting on work that analyzes causal path models for explaining students' participation in SMET majors by racial-ethnic groups by gender. She discusses suitability or lack of suitability of models for particular groups. Additionally, Stage will discuss her findings regarding students' high school coursetaking patterns and their mathematics achievement scores by ethnic group.

*Opportunities for Employment at the National Science Foundation (NSF) with Special Emphasis in Education, Evaluation and Social Sciences*, Jane T Stutsman, National Science Foundation

Jane Stutsman will describe the funding and research opportunities in science, mathematics, engineering and technological (SMET) education at the National Science Foundation with special attention to programs of the Directorate for Education and Human Resources (EHR), Division of Research, Evaluation and Communication. This Division aims to support evaluation studies and evaluative research studies in the whole spectrum of education, from K through 12, and all levels of higher education including professional development. She will also describe other areas of the Foundation where the audience might have particular interest, including the Social, Behavioral and Economic Sciences Directorate.

## **SESSION 586: MultiPaper**

**Room: Waiialua**

Sponsored by the Topical Interest Group on Pre-K - 12 Educational Evaluation

### **Evaluation in the Service of Children Placed at Risk**

Chair: Sam Stringfield, Johns Hopkins University

Presenters: *Initiative Management and Evaluation of Programs for Students at Risk in Schools*, Carol Cardno & Janet M Clinton, UNITEC Institute of Technology

This paper reports an initial investigation of the impact on specified evaluation outcomes of organisational capacity to manage an innovation. The management of an externally funded innovation (such as the introduction of a programme for students at-risk) assumes capability in managing a project and requires a set of skills such as proposal writing, implementation planning, implementation monitoring and sustaining the initiative. The context for the study is an evaluation being conducted by UNITEC Institute of Technology on behalf of the New Zealand Ministry of Education. Seventeen schools obtained resources (from a \$1m funding pool) for a variety of initiatives related to supporting students at risk. The evaluation aims to determine the merit and worth of such programmes as well as to provide opportunity for schools to network and share 'best practice'. The development of a set of standards to measure initiative management capability was informed by educational management theory and the initiation of the evaluation process. Data will be collected and analysed to compare the preliminary evaluation outcomes against these

standards. This evidence will be used to test the hypothesis that the inability to manage an innovation capably will influence the outcomes.

*Evaluating the Effectiveness of a Technical Assistance System To Build the Organizational Capacity of a High-need School District: One Piece of the Puzzle*, Kimberly S Cowley, AEL Inc

This evaluation supported school change facilitators as they developed and refined a technical assistance system to build the organizational capacity of a rural, high-poverty, low-performing school district to sustain continuous school improvement. Three surveys, focusing on schools' capacity to engage in and sustain improvement initiatives, were first administered to school staff in 1997 to assess their pre-intervention capacity. Interventions included staff development, curriculum alignment for mathematics and language arts, and electronic curriculum databases. The surveys were re-administered again in 1999 as one measure of change. Findings show significant improvement in teacher empowerment at the high school level; no progress in creating sustainable professional learning communities at the elementary, middle, or high school levels; and a decrease in the perception of organizational effectiveness at the middle school level. Recommendations were shared informally with district staff throughout the three-year project, as well as in formal pretest and posttest reports. This study offers further insight into the issue of school reform, especially in the area of organizational capacity.

*Evaluation of California's Immediate Intervention/Underperforming Schools Program (II/USP) of the Public Schools Accountability Act of 1999: Year 1 Findings*, Anne E Just, California Department of Education

The California Department of Education (CDE) is implementing a three-year internal evaluation study of the II/USP component of the PSAA. This legislation provides for determining the academic performance of individual public schools, intervening with low-performing schools, and rewarding schools attaining the target state performance level or making substantial progress toward this goal. II/USP includes two types of schools. The 78 "implementation" schools were funded with federal Comprehensive School Reform Demonstration (CSRD) monies to immediately undertake comprehensive reform to improve student achievement. The 353 state-funded "planning schools" have one year to plan and two years to implement their reforms to raise student achievement. Focused on implementation and impact the CDE evaluation has the following components:

- 1) creating a database with a composite statewide profile of II/USP schools and individual school profiles, for annual updating to track progress
- 2) conducting surveys to enhance database information and provide overall implementation and impact data
- 3) conducting fieldwork to enhance survey findings via comprehensive site visits to 30 purposively selected II/USP implementation and planning schools, using a site visit perception scale to ensure interrater reliability and to highlight school differences for analysis of effective and less effective II/USP implementation and impact
- 4) documenting CDE assistance and management activities to inform CDE about effective ways to assist II/USP schools and their districts.

Preliminary study indications suggest wide-ranging school capabilities to undertake II/USP responsibilities and to effect positive change in student achievement.

*Evaluating Success of Schools: Two Alternative Accountability Systems*, Cheryl A Videen & Geoffrey Maruyama, St Paul Public Schools

Schools have long been evaluating their effectiveness. In today's society, accountability systems have attached consequences -- sometimes very public and serious consequences -- to the results of school evaluations. Poorly structured accountability systems can unfairly punish schools and change their reputation. As evaluators, it is important for us to understand the principles and issues related to these systems. This session will describe an accountability system proposed in Minnesota and explore the problems associated with the proposed system. The session will then describe an alternative model, which is in the early stages of being implemented in an urban district in Minnesota. Issues discussed will include 1) what comparisons are being made -- individual growth versus changes across cohorts, 2) how students with differing backgrounds are treated, and 3) defining what changes are necessary to indicate improvement.

Discussant: Sam Stringfield, Johns Hopkins University

**SESSION 587: Panel**

**Room: Waianae**

Sponsored by the Topical Interest Group on Collaborative, Participatory & Empowerment Evaluation

**Evaluation as a Catalyst for Organizational Learning: Challenges in Translating Theory into Practice**

Chair: Julie Chambliss, EMSTAR Research Inc

The processes by which evaluation can be a catalyzing mechanism for organizational learning that results in optimal adaptive organizational change will be explored from a variety of theoretical and applied perspectives. Panel members will focus on lessons learned from their experiences of working with a variety of both public and private sector organizations to promote organizational learning through evaluation. Issues regarding the differing perspectives, interests and motivations of internal versus external participants will be examined. The need to complement participatory self-evaluation with an external perspective through activities such as benchmarking organizational functioning against best practice will be addressed. Tensions that often arise between the needs and values of various stakeholders will be described, and strategies for enhancing the likelihood that organizational learning results in changed outcomes will be presented. The session will close with an interactive discussion among panel members and attendees.

Panelists: *Complements to Participative Inquiry for Organizational Learning: The Value of the External Eye*, E Jane Davidson, Claremont Graduate University

Approaches to the use of evaluation for building organizational learning capacity have to date focused heavily on the use of participative methods that help staff evaluate and learn about their own organizations. Although self-evaluation is a necessary and valuable element of organizational learning, there is also great benefit in complementing such an approach with a form of evaluation that originates from outside the organization. Business organizations in particular need to know how their performance stacks up against competitors, and therefore require evaluation that is driven specifically by those information needs. This presentation argues for the addition of a second type of evaluation for organizational learning - one that focuses on benchmarking organizational functioning against best practice. In this spirit of mixed method evaluation and critical multiplism (Shadish, 1993), the organization can ensure that its ongoing organizational learning activities are balanced with occasional 'reality' checks that triangulate on the internally-driven inquiry.

*Participatory Evaluation as a Catalyst for Organizational Learning*, Nancy Grudens-Schuck, Iowa State University

Important issues related to successful organizational learning in organizations through evaluation emerged in a qualitative study of a sustainable agriculture educational program for farmers in Ontario, Canada. The first issue concerns position and power of participants internal to the organization who became involved in organizational learning catalyzed through a participatory evaluation. Heron's theory of closed and open boundaries is combined with theories of organizational learning processes put forward by Preskill and Torres to arrive at an understanding of how those who become intensively involved in participatory evaluation may differ from those who don't. The second issue of concern relates to the tension between internal actors' persistent desire for "ammunition" type evaluation data for funding and accountability, and the reality of the diversity of types of evaluation results that were valued differently by different players.

*Organizational Learning in Health Care: An Outcome of Process Evaluation*, Melissa Gallison, Independent Consultant

Health care organizations in the U.S. are influenced by diverse incentives in an increasingly competitive climate. Quality standards are process and outcome oriented, and can yield disparate evaluations. Assisting health care organizations to recognize linkages between process and outcome quality indices can occur in a framework of organizational learning, and reduce gaps between program plans and delivery. This presentation will illustrate ways that an evaluation led to process review by a nonprofit health care agency.

Relationship to Building Evaluation Capacity: Consideration of the motivational forces of the request for outcome

evaluation was vital to project planning. Analysis of the organizational environment led to recognition of the value of a participatory approach to evaluation. Value clarification activities uncovered additional motivations and readiness to learn from experience about the influences of incongruent accountability standards, and were associated with an impetus to update the agency mission and goals by the board and members.

Discussant: Julie Chambliss, EMSTAR Research Inc

**SESSION 588: MultiPaper**

**Room: Molokai**

Sponsored by the Topical Interest Group on International and Cross-cultural Evaluation

**Evaluation of Education Programs in Different Cultural Contexts**

Chair: Thomaz K Chianca, Ayrton Senna Foundation

Presenters: *Evaluating a National After School Program in Brazil: A Participatory Approach*, Thomaz K Chianca & Eduardo Marino, Ayrton Senna Foundation; and Margareth Goldberg, Instituto Ayrton Senna

In 1997, Ayrton Senna Institute (IAS), launched a strategy of using evaluation as a tool to improve one of its major programs. The Program was designed to help developing children and adolescents' personal, social, productive and cognitive capacities, using sports as the pedagogic method of an integral educational approach. It involves six projects based on major universities spread throughout Brazil. The strategy included building evaluation capacity of local projects' leaders, through evaluation training seminars and individualized support for developing local evaluation plans. In 1999, the Institute sponsored a summative evaluation of the Program. The evaluation was coordinated by external consultants working closely to IAS staff as well as local projects directors. The process of determining evaluation questions, indicators, and data collection, analysis and interpretation strategies was developed considering major stakeholders. The paper presents the main lessons learned from IAS' evaluation strategy, including perspectives from different stakeholders.

*Strategies for Evaluating a Parenting Program with Immigrant Populations*, Angela G Montagno & Jim Gaudin, University of Georgia

The number of immigrants being served by social service agencies in the United States is rapidly growing. With this growth, agencies are encouraged to develop programs that are culturally sensitive and appropriate in serving the needs of these populations. This paper discusses the challenges of developing and evaluating culturally sensitive programs for immigrant populations. It specifically focuses on a parenting program serving Mexican immigrants in Athens, Georgia. The use of a nonrandomized pretest/posttest design from a standardized and an adapted instrument is discussed to measure parent behaviors and attitudes before and after the parenting program. Special emphasis is given to the selection of culturally sensitive curriculum topics for the program measuring the frequency of parent/child interactions and parent opinions of their relationships with their children. Demographics, such as age and length of time in the country are considered in relationship outcomes.

*Development of a Research Agenda for American Indian/Alaska Native Education*, E William Strang, Westat

For evaluation capacity to be increased requires, at a minimum, asking relevant evaluation questions. This, in turn, requires making the affected community a partner from the beginning. Under the authority of Executive Order 13096, federal agencies have been directed to work with Native Americans to improve the academic achievement of American Indian and Alaska Native students. One of the strategies used to achieve this goal is to develop a comprehensive research agenda, and one of the central assumptions of that process is to involve Native American communities now in establishing the agenda and, in the future, to develop the capacities of those communities to implement it. This presentation reports on the process of establishing the research agenda, from the perspective of the U.S. Department of Education's contractor, with a focus on identifying evaluation strengths and gaps in the American Indian and Alaska Native communities.

Discussant: Thereza Penna-Firme, CESGRANRIO Foundation

Sponsored by the Topical Interest Group on Health Evaluation

**Logic Models: Where No Evaluator Has Gone Before (With Apologies to Gene Roddenberry)**

Chair: Molly Engle, Oregon State University

Logic models have come of age. It is rare that a conversation on evaluative inquiry can take place without including something about logic models. Now that logic models have entered into the main stream, the question occurs, "What next?" This session will focus on the "what next?". Panelists will discuss the application of logic models to missions of organizations, using logic models for managing for results, using logic models as a learning orientation rather than results orientation, and as an approach for constructing and applying program theory.

Panelists: *Logic Modeling: Looking for Love in all the Wrong Places*, John A McLaughlin, University of Southern California

Program managers across the private and public sectors are being asked to describe their programs in new ways. Funders and other stakeholders want managers to present a logical argument for why and how the program is addressing specific problems. At the same time managers are challenged to present a framework for monitoring and evaluating their programs to enable them to detect error and do something about it - managing for success.

While planners and evaluators have employed it for quite some time, the logic model recently has become highly regarded as a tool to help managers develop and tell their performance story. The purpose of this presentation is to explore the utility of logic modeling for communicating how a program will work to solve the problem it is targeting and for identifying key performance information to guide monitoring and evaluation of the program. An alternative approach will be described that the author believes better satisfies the managers' need to have information available to continuously improve their programs using a learning orientation rather than results orientation.

*Mission-level Logic Systems: Dr. Spock Confronts the Irrational Universe*, Michael Quinn Patton, The Union Institute

Logic models have traditionally focused, like evaluation, at the project and program level. Increasingly, evaluation will need to address mission level evaluation for mission-driven organizations. How well does the logic of logic models and theories of change work at the mission level?

*Modeling for Results*, Joseph S Wholey, US General Accounting Office

This paper will discuss roles for logic modeling in overcoming challenges that may arise at each of the steps in managing for results: (1) developing a reasonable level of agreement on results-oriented goals and strategies for achieving goals; (2) measuring and evaluating performance; and (3) using performance measurement and evaluation information to improve program effectiveness, strengthen accountability, and support policy decision making. Examples will be given to illustrate use of logic models in developing agreement on strategies for coordinated efforts in crosscutting programs, in which different agencies share goals; using intermediate measures (outputs and outcomes) to show progress or contributions to intended results; and developing partnerships to leverage or mitigate the influence of factors outside program control.

*Constructing and Applying Program Theory for Program Planning and Development*, Huey T Chen, Centers for Disease Control and Prevention

The construction of program theory for planning and development purposes requires a close working relationship between stakeholders and evaluators. This paper discusses strategies and procedures for evaluators to get stakeholders' input into the theorizing process and the application of program theory for program planning and development purposes.

Discussant: Molly Engle, Oregon State University

**A Dialogue on Guiding Principle E: Responsibilities for General and Public Welfare**

Chair: Jody L Fitzpatrick, University of Colorado at Denver

One of the most controversial principles in the Guiding Principles for Evaluators, and one which opens new territory for evaluators, is Principle E. It states: "Evaluators articulate and take into account the diversity of interests and values that may be related to the general and public welfare." This principle has been hailed by some (Covert, 1995; House, 1995) and viewed with skepticism by others (Rossi, 1995; Knott, 1995). It requires more than simply involving and considering various stakeholders. The principle asks that evaluators "maintain a balance between client needs and other needs", and "consider not only the immediate operations and outcomes of whatever is being evaluated, but also the broad assumptions, implications, and potential side effects of it." These panel members, working in diverse settings and representing different points of view, will discuss their interpretation of this principle and the manner in which they apply it in their own work.

Panelists: *Guiding Principle E: What We Did and Why We Did It*, William R Shadish, University of Memphis

The 1994 Task Force on Guiding Principles for Evaluators consisted of four members representing academia (William Shadish, Dianna Newman), private sector evaluation (Mary Ann Scheirer) and public sector evaluation (Chris Wye). This talk will reconstruct some of the Task Force thinking that went into Principle E, including the decision to add it and its subparagraphs, and especially some carefully crafted phrases that may contain more hidden meaning than is obvious in retrospect.

*Should Principle E Be Critical for Evaluators in Business and Industry?* Darlene F Russ-Eft, AchieveGlobal Inc

The Guiding Principles for Evaluators suggests attention to "the general and public welfare." Furthermore, the Standards on Ethics and Integrity of the Academy of Human Resource Development (1999) indicate "responsibilities to the community, the society -- and the planet." As evaluators in a for-profit business that develops training materials and interventions, our work focuses on impacts of training interventions on trainees and the organization. This training helps people in their leadership, teamwork, customer service, and sales performance. Such performance improvements benefit the person and the organization. Presumably, improved functioning in the organization leads to better products and services to customers and the general public. In some cases, our evaluations would examine impacts on customer and the general public. However, I would argue that including the general public in each and every evaluation would violate standards on Cost Effectiveness (F3) and Fiscal Responsibility (P8) as outlined in The Program Evaluation Standards.

*Why We Need Principle E*, Ernest House, University of Colorado

I will discuss the reason why we need the principle in the evaluation field, how the principle can be implemented, some examples of implementation, the problems evaluators encounter in implementing it, what objections to the principle exist, and why these objections should not prevail.

*Which Side Are You On? Truth, Justice and the Way of Evaluation*, David M Chavis, Association for the Study and Development of Community

The evaluation of social programs, especially those that take a systems or environmental approach, often take place in ethnic minority and other disenfranchised communities. The client is often government and foundations seeking to solve complex problems with the least amount of risk. Often these initiatives confront and are directly affected by social inequities, injustices, and other conflicts that exist in the community. These inequities have a direct bearing on the process and outcomes of the initiatives. Reporting these conflicts and issues may increase the risk to the client. The evaluator can both bear witness and give voice to these issues on behalf of a community when reporting the "truth." When compromise or balance cannot be reached, the principles do not offer advice. Examples of the presenter's experience in handling these matters and the need for personal decisions and professional support in these situations will be discussed.

Discussant: Jody L Fitzpatrick, University of Colorado at Denver