Thursday, November 2, 7:10 A.M. to 8:25 A.M.

SESSION 201: Plenary I **Room:** Lanai

The Hokulea Voyages – What Can Evaluators Learn from the Renaissance of Polynesian Voyaging?

Introduction: Nick Smith, Syracuse University

Presenter: Nainoa Thompson, Polynesian Voyaging Society

Thursday, November 2, 8:40 A.M. to 10:10 A.M.

SESSION 201: Panel **Room:** Lanai

Sponsored by the Topical Interest Group on Minority Issues in Evaluation

Garbage in Garbage Out: Challenges of Evaluation in Organizational Cultures in Which Information Gathering Is Not a Norm

Chair: Anna Marie Madison, University of Massachusetts at Boston

The results of evaluations are no better that the data collected . One of the greatest challenges to the accuracy and reliability of the data collected in the evaluation of community-based organizations is learning how to work with organizations that have no tradition of record keeping and information gathering. The panelists will discuss some of the data collection challenges they have faced in the evaluation of programs implemented in community-based organizations. Each presenter will also discuss the strategies used to overcome the challenges faced. The first presenter will present data gathering challenges and how they were overcome in the evaluation of a statewide multi-site youth arts program. The second presenter provides an overview of data gathering for a national evaluation of five sites of a planning coalition for homeless initiative. The third and fourth presenters will present an analysis of the data collection challenges in the evaluation of HIV/AIDS prevention programs.

Panelists: *We Are Artists! We Don't Do Data: The Challenge of Data Gathering When Two Collide*, Anna Marie Madison, University of Massachusetts at Boston

One of the recent trends in youth programming is the utilization of the arts as a medium for healthy youth development. The arts are used to help youth develop self-confidence and life skills, and to bond with a peer group. The artists working with the youth are dedicated to the work they do with youth, but have very little interest in collecting information about their work. During the last four years efforts have been made to design a participatory evaluation that directly involves artists in the evaluation process. The highlights of the challenges, the strategies to overcome them and the outcomes will be discussed in this presentation.

Data Collection Challenges in Working with a "Different to Reach Population," Joeritta de Almeida & Phoebe Soares, The Better Homes Homes Fund

The Better Homes Fund has been involved in a participatory evaluation of three grassroots organizations funded to do community organizing with homeless, formally homeless, and at-risk families during the last two years. The evaluator has designed a participatory process that takes into consideration the assets of the members involved and has built upon these assets to gather data in each of the communities. This presentation will discuss some of the alternatives to traditional data collection methods used in the evaluation of this initiative and the results of the data gathering effort.

The Challenge of Data Gathering in a Community-based Clinical Setting, Janine Walker, Center for Community Health Education and Research

This presentation provides an analysis of data collection challenges in the evaluation of a community-based HIV/AIDS prevention project. Data collection was challenged by a language barrier and the lack of program staff training in data collection. Some of the problems encounter included data collection errors due to third-party interpreter mistakes in the communication of information, lack of clinical data collection quality controls, and the absence of data on sensitive, personal interview questions about drug use and sexual practices.This presentation will cover the myriad of problems encountered in data gathering and procedures taken to minimize these problems in the future.

Performance Monitoring and Accountability Data Collection Issues: The Case of a State Public Health HIV Prevention Initiative, Robert Woronoff, The Home for Little Wanderers

This presentation describes some of the data collection challenges faced by publicly funded HIV prevention and education programs under state and federal mandates to monitor program outputs. Because the HIV disease is fraught with obstacles and pitfalls which arise from addressing the societal taboos of sexual and drug use behaviors (the primary modes of HIV transmission), an essential element of data collection is that program monitor's partner with the community to understand conditions faced by individuals considered to be "at high risk" of HIV infection. This presentation will discuss the data collection system the Massachusetts Department of Public Health (MDPH) HIV/AIDS Bureau has developed to overcome some of the data challenges to the accuracy of data collection to monitor and evaluate HIV prevention programs.

SESSION 202: MultiPaper **Room:** Akaka

Sponsored by the Topical Interest Group on Human Services Evaluation

Evaluation and Human Service Professionals

Chair: Leon H Ginsburg, University of South Carolina

Presenters: *The Federal GPRA Evaluation Mandate: Is Social Work Ready?* Tamara S Davis, University of Texas at Austin

Evaluation is key to the strategic plans required of all federal agencies under the Government Performance and Results Act (GPRA) of 1993. Since federal agency funding supports a large proportion of human service programs, this mandate has significant funding and resource allocation implications for social work. This paper suggests that social work may not have the necessary structures in place to answer the increasing demands for accountability and the potential fallout from this federal policy. Social work continues to rely on traditional evaluation research methods, while the current evaluation environment calls for inclusion of more flexible options, specifically ongoing and participatory evaluation methods. This paper further suggests that the creation of a National Center for Social Work Research (S. 178) could help bridge evaluation gaps and offer a centralized mechanism for maximizing distribution and monitoring the evaluation activities of federal resources allocated to human/social service delivery systems.

Building Social Work Evaluation Competence, Leon H Ginsburg & Abraham H Wandersman, University of South Carolina

Social workers have major obligations for evaluating their own practice as well as for participating in program evaluations. This roundtable is designed to provide a forum for social workers who attend the American Evaluation Association meeting to discuss their evaluation activities and to identify ways to improve evaluation processes in the field. It is also designed to help identify ways to enhance social work education about evaluation. Part of the impetus for the roundtable is the projected publication, in 2001, of Dr. Ginsberg's text, Social Work Evaluation, which is in press with Allyn and Bacon. Dr. Wandersman is the author of one chapter in the book. A psychologist, he is the author of many scholarly works on evaluation.

A Protocol To Evaluate Cultural Competency Programs in the Human Services Profession, Richard E Njoku, Cuyahoga County Community Mental Health Board

One of the practice implications of serving a growing clientele of ethnic and racial minorities is the move toward training and retraining of human services workers to become culturally competent in their professional practice. In many states across the nation, this move is manifested in a paradigm shift which integrates cultural responsiveness into the overall competency-based, family-centered, training for human services workers.

This paper proposes a protocol to evaluate cultural competency programs in the human services profession, based on Kirkpatrick's (1979, 1994) four-stage training evaluation model. It demonstrates practical approaches to evaluating cultural competency training programs at the reaction, learning, behavior (transfer of learning), and results (organizational change) levels.

Responding to Difficult Evaluation Requests, Vernon R Wiehe, University of Kentucky

As faculty teaching research in schools of social work, we often receive calls from social service agencies needing to have a program evaluated. Some of these requests present problems for the evaluator because they were poorly conceived or developed without measureable objectives. Also, the agency is concerned that if they cannot document the effectivness of the program, they are in danger of losing funding. This conversation will focus on the discussant and participants sharing their approach to such requests. Emphasis will be placed on the process an evaluator might follow in evaluating such a programs. The discussant will share a recent experience with the evaluation of an AIDS Hospice Program.

SESSION 203: MultiPaper **Room:** Koko

Sponsored by the Topical Interest Group on Evaluation of Services for Special Needs Populations

Measuring Effects: Reviews of Programs Serving Diverse Populations

Chair: Susan Stoddard, InfoUse

Presenters: *Problem-oriented Evaluation Design: The NIDRR Program Reviews*, Betty Jo Berland, National Institute on Disability and Rehabilitation Research

The National Institute on Disability and Rehabilitation Research is a federal agency that awards grants and contracts to conduct medical, engineering, social, psychological, and service delivery research related to disability. NIDRR is a "research institute in the public interest," balancing the demands of consumer relevance and scientific rigor. NIDRR's challenge was to design an evaluation/accountability mechanism stressing responsiveness to the various constituencies and engaging NIDRR staff in evaluation and program improvement. The resultant model, Program Review, developed with the help of an evaluation consultant, served the multiple purposes of involving diverse constituencies, empowering NIDRR staff, improving linkages to other Federal research agencies, and increasing awareness of evaluation for all participants. The resulting Centers of Excellence Model, embraced by NIDRR and its grantees, has received attention from other agencies and Congress. NIDRR will make Program Review a permanent activity, and a number of other agencies are adapting it.

Forging New Frontiers: Reengineering a Measurement System for Child and Adolescent Services, Leonard Bickman & Ann Doucette, Vanderbilt University

Conducting continuous quality improvement (CQI) assessment often requires financial and human resources as well as specialized measurement approaches. Resources are often pieced together from existing efforts. This paper will describe the efforts of the authors to establish an international measurement system for child and adolescent service sectors that can be used in several settings including physical and behavioral healthcare, foster care, outpatient services, and inpatient care. The authors will provide an overview of three integrated assessment system dimensions (baseline, concurrent, and follow-up). This paper addresses the advantages of using an Item Response Theory (IRT) approach for measurement development, and computerized and web-based applications. The proposed CQI model includes multi-informants as well as process and outcome measures. The outcome framework is meant to be flexible and sensitive enough to adapt to the needs and characteristics of specific child-serving systems. Data sources and instrumentation, financial resources, and sufficient planning will also be addressed.

Implementing Program Review: A Model of Evaluation and Program Management in the National Institute on Disability and Rehabilitation Research, Susan Stoddard, InfoUse

In 1996, the National Institute on Disability and Rehabilitation Research (NIDRR), U.S. Department of Education, initiated a program review process for review and evaluation of NIDRR's national Research Centers. The author, a planner and evaluator, designed and facilitated the reviews. Program Review is based on project planning and management, program evaluation and organizational development. Over one hundred reviews have been completed.

Program Review has gained increasing acceptance in NIDRR program management, and helped changed organizational culture. Program Review enhanced NIDRR's internal evaluation capacity and expanded NIDRR's investment in evaluation activities. NIDRR Grantees have increased their own evaluation activity as a result of preparing for reviews. Program review is now a formal part of the Department of Education's Government Performance and Results Act (GPRA) plan. This paper describes the model, its development, implementation and potential for transfer to other government agencies, research institutes and large

SESSION 204: Panel **Room:** Waimea

Building Evaluation Capacity: Three Technical Assistance Approaches

Chair: Lance Potter, Corporation for National Service

Evaluators at Aguirre International provide technical assistance in evaluation to a wide range of social programs in California and across the country. This panel will examine ways in which three different technical assistance efforts have worked to raise the evaluation capacity of programs at multiple levels. Technical assistance recipients include community-based organizations, public schools, and colleges participating in the Corporation for National Service's AmeriCorps programs, organizations serving youth in out-of-school-time programs through Save the Children, U.S.A. and a network of bilingual public radio producers, Radio Bilingue. Each effort has sought to enhance organizational capacity in different areas of evaluation and through different media, and each has met with different challenges and successes.

Panelists: *Characteristics and Techniques of Assistance that Builds Internal Evaluation* Capacity, Linda Morell & Kenneth Terao, Aguirre International

Aguirre Group has been providing training and technical assistance to local community organizations within the U.S. for the past four years through its contract with Save the Children, U.S. Programs. Throughout this project Aguirre has identified (1) characteristics programs (and individuals conducting evaluations) need to have successful internal evaluations and (2) techniques individuals and groups at multiple levels can use to keep evaluation simple, on track, and meaningful. For example, a technique that has proven effective is using a 'Train-the-Trainers' model that includes not only training senior staff in how to conduct evaluation but also in how to train and provide technical assistance to others in evaluation. This paper outlines the characteristics and techniques that Aguirre has identified which support the creation and maintenance of a manageable internal evaluation system that is both meaningful to stakeholders and dynamic given the changing needs and emphases of a national organization and individual local communities.

Radio Bilingue: Development of Formative Evaluation Capacity, Ed Kissam, Aguirre International

The Aguirre Group's work with the Radio Bilingue network continues its evaluation of the network's multi-cultural health outreach programming to Latino audiences. Aguirre has now trained more than 15 of the station's producers in basic formative evaluation and will be working with them as they begin evaluation implementation. The goal of this effort is that producers' new skills in formative evaluation will foster more reflective and creative production planning, which will maximize attitudinal and behavioral change among targeted audiences. This presentation will report on the outcomes of training and coaching with producers and summarize initial conclusions about the effectiveness of this "bottom-up" strategy for building a community-based organization's ability to incorporate evaluation findings into program operations. The presentation will also give attention to the overall strategic planning implications for the station -- since, like other non-profit community broadcasters, it is competing with commercial entities oriented toward gaining market share.

Raising Capacity of Diverse Organizations: Outcomes, Successes and Challenges, Nicole C Vicinanza, Susan Gabbard & Anna Marie Schmidt, Aguirre International

Since the inception of the AmeriCorps National Service program, the Corporation for National Service has required that all programs conduct objectives based internal outcome evaluation of their services. This evaluation requirement has demanded that programs increase both their evaluation capacity and capability. During the past five years, Aguirre International's Project STAR has provided these programs with technical assistance in developing measurable objectives, planning and designing internal evaluations and collecting and reporting evaluation data. This paper will discuss the findings of a review of AmeriCorps program documentation including whether or not 1) programs increased their ability to write measurable objectives 2) programs conducted evaluation 3) programs reported evaluation results and 4) technical assistance in evaluation is associated with improvements in evaluation capacity and capability. It will also discuss the areas in which programs showed improvement in evaluation related skills and compare these areas with the types of assistance that were provided.

Discussant Lance Potter, Corporation for National Service

SESSION 205: Panel **Room:** Niihau

Sponsored by the Topical Interest Group on Evaluation Use

Learning Organization Construction Ahead! Nuts and Bolts, Tools for Evaluators

Chair: Anna Rockhill, Portland State University

The role of evaluation in creating and maintaining learning organizations has sparked an interest to collect some nuts and bolts as well as tools for implementing a different type of inquiry. Presenters will share their understandings of the relationship between evaluation and learning organizations with tools or studies they have created. Others who are newer to the Preskill & Torres concept of EILO will take a more "nuts and bolts" approach that they have collected along the way. Special attention will be paid to the challenges of transferring evaluation knowledge to organizations and people not trained in the evaluative process.

Panelists: *'Untraditional Tools' Needed for the Evaluator's Kit*, Anna Rockhill, Portland State University

EILO promises to improve evaluation by enhancing both the quality of information and the likelihood of utilization. However, simply getting started presents researchers with a number of challenges that may require 'untraditional tools,' stuff they didn't teach in graduate school. We will share stories 'from the trenches' of some of the difficulties we encountered in implementing EILO within an innovative foster care project spanning 4 different organizations and involving a diverse array of participants including management, caseworkers and foster parents.

* Our goals in instituting EILO
* Brief overview of our EILO workplan
* Description of 'getting started'
* Some potential challenges Challenges to getting buy-in from participants
* Cynicism about the prospects for organizational change
* Lack of confidence in their ability to carry-out research
* Concern about workload Challenges to team-building
* Interpersonal conflicts
* Power differences
* Cultural issues Challenges facing research staff
* 'Letting go'
* Teaching research

Putting Evaluative Inquiry into Practice: Helping New and Non-evaluators Use EI in their Work with Leadership Development Initiatives, Jennifer W Martineau, Center for Creative Leadership

Faced with increased requests from clients to incorporate evaluation into leadership development initiatives, yet at the same time no enlargement of its evaluation staff, the Center for Creative Leadership made an investment in the development of a new tool. Named 'The Evaluation Toolkit for Leadership Development', the goal for this tool is to provide the Center's training and business development staff with the background knowledge and evaluation methods they need to do this work with clients. Using the Preskill and Torres framework for Evaluative Inquiry, the Evaluation Toolkit provides 1) education about evaluation and 2) an evaluation design process within three familiar areas:

1) Focusing the Evaluation;

2) Carrying Out the Evaluation; and
3) Applying Learning.
This presentation demonstrates one organization's use of the Evaluative Inquiry framework in its solution to the question: "How do we ensure high-quality evaluations of our services when we are not always the ones to perform the work?"

Evaluation of Embedded Events as an Organizational Learning Mechanism, Scott A Robbins, University of New Mexico

One use for evaluation findings is to reinforce the learning that occurs in organizations when they respond to crises and opportunities. When applied in this manner evaluation serves as an organizational learning mechanism. This presentation describes recent research conducted to test and refine a comprehensive strategy for evaluating organizational learning that results from an organization's response to different types of events. The evaluation strategy examined two dimensions of organizational learning. Using embedded organizational events as units of analysis the study examined organizational learning as both a dynamic network process and as a change in organizational knowledge structure. Interview, observation, and archival data were used to analyze both dimensions of organizational learning. Following the analysis of each response episode the organization members who were involved were informed of the changes that occurred at the various sites of organizational memory. By using evaluation results in this manner organizational learning can be reinforced.

Increasing Awareness of Effective Evaluation through an Internal Participatory Learning System, Rebecca P Barnette, Atlanta Alliance on Developmental Disabilities

The Objective/Goal of this Evaluation is to increase staff awareness of effective evaluation in their everyday lives and how this awareness will lead to a more effective and efficient workplace.   The Materials and Methods include:
* Personal experiences: observation, staff perception survey results, "interviews"
* Tim T. Robinson and J. Bradley Cousins, Internal Participatory Evaluation as an Organizational Learning System: A longitudinal Case Study(as a guide book for my own understanding)
* Full disclosure, explaining current trends patiently and thoroughly, answering questions (open-door policy)
* Training staff (not necessarily a formal training, but teaching them what to look for early on--grant proposal stage, program restructuring).
Results indicated that colleagues who ask questions for clarification and advice instead of feel like they cannot; who are willing to listen and to contribute rather than avoid evaluation and be afraid of it; more ideas (from everyone) that benefit the organization instead of just one person in charge of "thinking" (with evaluation in mind). An Evaluator's biggest job is creating an amiable environment. It can be done by treating staff with mutual respect and being available to them. Also, the most common obstacle I have found among my co-workers has been that of non-separation. They knew the evaluation concepts well, but they did not know how to separate themselves from the equation and be more objective. I have learned to be patient when I explain things in that I not only explain the issue, but I also explain my thought process in getting to that issue.

Discussant: Hallie Preskill, University of New Mexico

SESSION 206: Demonstration **Room:** Hilo

Sponsored by the Topical Interest Group on Non-profit and Foundations Evaluation
Turning Information into Action: A Model and Case Study for Integrating Community Indicators into Grant Making
Chair: John B Bare, John S and James L Knight Foundation
Presenters: John B Bare, John S and James L Knight Foundation
Lizabeth L Sklaroff, John S and James L Knight Foundation

To add value to society, foundations need new models for identifying and integrating relevant information into planning. The demonstration will present a model for 'turning information into action' to improve quality of life in communities. At Knight Foundation, the evaluation team is directing a multimethod approach that incorporates information on community needs and nonprofit capacities into program and evaluation planning. It's the market research behind the grant making. The effort builds on the Foundation's 26-city Community Indicators project and ongoing capacity-building research. The demonstration will provide attendees with a clear explanation of the model, along with actual cases describing how the Foundation is applying it in two specific communities. The demonstration will be interactive. In the case, the presenter will facilitate a discussion of decision strategies. To illustrate the model's strengths and weaknesses, the discussion will address the implications of decisions funders must make in applying the model.

SESSION 207: Think Tank **Room:** Puna

Sponsored by the Topical Interest Group on Business and Industry

Aligning Corporate Evaluations to Business Goals To Build and Maximize Evaluation Capacity

Chair: Vanessa D Moss-Summers, Xerox Corporation

Facilitators: Trude J Fawson, The Training Impact Group
Toni Hodges, Bell Atlantic
Sharon Marie May, Compaq Computer Corporation
Bill McCook, Lotus Development
Ferdinand Tesoro, Dell

Evaluators in the business and corporate communities face the continual challenges of marketing, implementing, and aligning their work with the overall goals of their respective organizations. To address these challenges, several evaluators from Fortune 500 companies formed the "Business Alignment Group" (BAG) (a sub-group of a business/industry Evaluation Consortium) to define a framework to focus evaluations, document best practices, and share tools. The framework examines attributes, processes, outcomes/success factors, caveats, and competencies required when positioning, executing, and "leading the organization" with effective evaluation studies. The proposed panel discussion will address the elements of evaluation capacity (motivational, organizational environment, workforce and professional development, resources/supports, and learning from experience) in terms of aligning, positioning, executing, and leading with goals and priorities of represented organizations. Discussion will address mapping of evaluation attributes to framework and capacity elements. Additionally, a forum will be provided to gather feedback on the strategy/framework.

SESSION 208: MultiPaper **Room:** Kohala

Sponsored by the Topical Interest Group on Evaluation Use

Assessment Issues in Evaluation Utilization

Chair: Carolyn Huie Hofstetter, University of California at Berkley

Presenters: *Performance Measurement for Accountability of Undergraduate Psychology Programs: Moving from Promise to Meaningful Implementation*, Shawn R Edmondson, Utah State University

Concerns and criticisms of the effectiveness of higher education have resulted in legislative mandates for assessment and accountability of undergraduate programs. As Sheehan (1995) has noted, however, many programs have struggled in attempting to develop effective assessment programs. Two concerns are paramount in this struggle: (1) developing valid indicators of performance and (2) developing procedures by which the information provided by these indicators is actually used. After summarizing the efforts to validate an exit competency assessment of undergraduate psychology majors, this paper addresses the question of strengthening the use of this information by students, by faculty, and by university administrators.

Uses of Assessments in Standards-based Evaluation, Carolyn Huie Hofstetter, University of California at Berkley

Evaluations of educational programs, particularly school-based programs, typically include student assessment information. Some evaluations, in fact, rely heavily on students' standardized test scores to make judgments about the overall performance of the schools over time. In California, for example, schools are given an Academic Performance Index (API) score that is comprised largely of students' test scores on standardized tests linked with statewide content-based standards. Critics have noted that reliance on these test scores produces unfair indicators of student and school performance, particularly for schools with large numbers of English Language Learners (ELLs) who traditionally exhibit lower performance due to the language demands. This paper will examine issues related to the use and misuse of assessments in accountability-based evaluation efforts, notably large-scale testing programs in states with large populations of English language learners.

Utilization of State Criterion Referenced Test Scores: Does Testing Lead to Instructional Adjustments? Louise R Moulding, Utah State University

Utah, like many states, uses criterion-referenced tests for secondary mathematics as part of the educational accountability system. The stated purpose of these CRTs is to improve student achievement through improved instruction based on teacher interpretation of test scores. However, no study had been conducted in Utah to determine if teachers are indeed using the results. As part of a study to collect validity evidence for the Utah pre-algebra criterion referenced tests, pre-algebra teachers from a large district were interviewed to determine the degree to which the scores are used to make adjustments to instruction. Issues such as timeliness of receipt of results, confidence in the test scores, ability to interpret the results forthe purpose of instructional planning, and actual instructional adjustments emerged. The utilization of tests scores by teachers is an important issue if accountability systems are to succeed in the stated purpose -- improvement of instruction.

SESSION 209: Panel **Room:** Kona

Sponsored by the Topical Interest Group on Extension Education Evaluation

The Implications of New Funding Sources for Extension (In-house) Evaluators

Chair: John A Michael, US Department of Agriculture

Panelists: John A Michael, US Department of Agriculture
Claude F Bennett, US Department of Agriculture
Marilyn Grantham, University of Minnesota
H Michael Harrington, University of Hawaii At Manoa

Federal support for Extension is gradually shifting away from formula funding to competitive grants. JOHN MICHAEL will summarize key changes that have taken place and suggest possible implications for extension evaluators (e.g., shifting responsibilities to include program planning and proposal writing; "lobbying" for the inclusion of evaluation work in grant competitions; evaluating entities other than extension educational programs). An Extension administrator will describe initial responses being made to funding changes and suggest how evaluators might help the organization make a successful transition. An extension specialist will then assess the meaning and impact of federal funding changes for evaluators affiliated with state land-grant institutions, identify action alternatives and make recommendations. CLAUDE BENNETT will explore implications for relations between research and extension.

SESSION 210: Panel **Room:** Ewa

Sponsored by the Topical Interest Group on Health Evaluation

Program Design and Evaluation Services: A Case Study in Building Evaluation Capacity in a County and State Health Department

Chair: Michael J Stark, Oregon Health Division

In this panel, we present the challenges, successes, and examples of projects undertaken by Program Design and Evaluation Services, an interdisciplinary, inter-agency evaluation and research work group, jointly established by the Multnomah County Health Department and the Oregon Health Division less than a decade ago. Beginning with 3 staff, PDES has grown to over 25 persons, including doctoral-level anthropologists, psychologists, health educators, sociologists, and epidemiologists. The panel presentation begins with an overview of PDES structure, a brief description of our current projects, an explication of factors which contributed to the success of PDES, and challenges along the way. Panel members will then present examples of specific evaluations, including issues in the implementation of the evaluations, results of the studies, and how the projects benefit the county and/or state health departments.

Panelists: *Program Design and Evaluation Services Overview*, Michael J Stark, Oregon Health Division

PDES' mission is to design and evaluate public health interventions, disseminate results, and develop funds for new projects. PDES currently is evaluating over 20 public health and health-related programs. Factors associated with PDES' success include:1) as an inter-agency team, we bring evaluation, epidemiology and surveillance capacity to the local health department, and population access and intervention capacity to the state health department, 2) by being behavioral scientists as well as evaluators, we bring social science theory and practice to respond to public health problems, 3) by being flexible we conduct projects across different areas and organizational structures, 4) as published scientists we secure grant funds, and 4) by collaborating with external agencies we bring capacity and funding to our home institutions. Challenges include reporting to two agencies, having to be self-supporting, and being an academic-like behavioral science unit within a public health environment.

Special Evaluations for the County Health Department, David W Dowler, Multnomah County Health Department

Special projects conducted by PDES for the Multnomah County Health Department have presented unique challenges as well as opportunities for building evaluation capacity. One recent project examined the effects of fee increases at primary health clinics on utilization and satisfaction of vulnerable patient groups. Surveys were administered across 7 clinic sites at 2 time points. A second project evaluates attitudinal and demographic characteristics of patients referred for partner notification interviews with County Disease Intervention specialists. In this study, qualitative data regarding factors associated with gonorrhea infection is collected via supplemental structured interviews by these in-place staff. The primary challenges for both projects arise from budgetary and methodological imperatives to partner with County staff for data collection protocols. This requires balance between research and clinical realities, and the need to operate under community-based evaluation principles. This process has provided helpful results to the Health Department, and built stronger ties for future evaluation collaboration with the County.

Evaluation Projects for External Customers: Juvenile Violence Prevention Programs Evaluation, Haiou He, Multnomah County Health Department

PDES provides evaluation services to non-public health agencies that deal with matters related to health. One example is that beginning in 1996, PDES has been working with the Oregon's Criminal Justice Services Division to develop a statewide strategy to evaluate juvenile violence prevention programs. A major challenge encountered in collaborating with CJSD, the Governor's Drug and Violent Crime Advisory Board, and local government /community-based was having different understanding and expectation of evaluation goals and plans. Lessons learned include that evaluators need to be flexible and be willing to tailor evaluation approaches to community needs and capacities. Through PDES' technical assistance and program monitoring, program sites developed process measures that show if programs are being implemented according to the plan. Program outcomes show declines in students' discipline or violence-related referrals, positive changes in students' school and classroom behaviors, improvements in school attendance, and reduction in delinquency recidivism.

Clinical Services and Prevention Evaluations: Introducing Behavioral Methods of Disease Control into a County Health Department, John A Dougherty, Multnomah County Health Department

In 1995, PDES expanded to include quantitative and qualitative evaluation of behavioral clinical and prevention services within the Multnomah County Health Dept. Outcomes included increased mental health assessment in the HIV Clinic, initiation of program evaluation by TB Clinic staff, and development of motivational and harm-reduction approaches to substance use in primary care clinics. Based on PDES evaluations, the Department has initiated a behavioral health program in primary care that includes functional and pathologic assessments of health behavior, and an emphasis on client self-management of chronic illness. Studies of lead dust prevalence and evaluation of the County's lead program led to increased resident participation in home lead hazard management. PDES also facilitated the diffusion of effective behavioral technologies across programs, used qualitative approaches to reduce conflict within and between organizations, and identified the negative side effects of rapid resource introduction on health programs

SESSION 211: Panel **Room:** Honolulu

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

Access to Chemical Dependency Treatment for Adolescents in the Public Sector: The Interplay of Multiple Systems

Chair: Katherine Laws, RMC Research Corporation

Researchers from RMC Research Corporation working on a CSAT funded study will explore the interplay of social service, juvenile justice, and health care systems in providing access to chemical dependency treatment for adolescents in the public sector. A mixed methods approach employing quantitative and qualitative data is used to examine the issues of access and availability of adolescent chemical dependency treatment services under two different administrative systems: managed care in Oregon and fee-for-service in Washington. Qualitative research efforts in both states have provided valuable insight into the impact of integration and coordination of services across the systems of care that affect adolescents' access to these treatment services. Treatment provider experiences working within these two systems have implications for improving the capacity of the field to meet the increasing needs of this population. Examples of effectively integrating and coordinating treatment services for adolescents in rural and urban counties are also featured.

Panelists: *Access to Chemical Dependency Treatment Services for the Medicaid-eligible Adolescent Population under Managed Care*, Dennis D Deck, RMC Research Corporation

States have increasingly turned to managed care to curtail the costs of publicly funded chemical dependency treatment, though critics have raised concerns about reduced access to services. The Oregon legislature added a capitated chemical dependency benefit to the Oregon Health Plan in 1995, providing an opportunity to examine the impact of managed care on access to treatment. This presentation will describe the statewide trends in access to treatment for Medicaid-eligible adolescents in Oregon over a seven-year period, before and after implementation of the capitated benefit. These trends are contrasted with results from Washington, where Medicaid-covered treatment services are reimbursed under fee-for-service and the approach to utilization management is different. The surprising results are discussed within the context of a model of service utilization developed by Aday and Anderson.

Chemical Dependency Treatment for Adolescents in the Public Sector: Access and Availability of Services under Managed Care and Fee-for-service Systems, Ryan D'Ambrosio, RMC Research Corporation

Two of the central issues of publicly funded chemical dependency treatment are access to and availability of services. When considering special populations these two issues come under even more scrutiny. This presentation will explore and examine the differing impacts two distinct administration systems have on access to and the availability of chemical dependency treatment services for adolescents aged 13-17. The analysis of these issues is drawn from state data of chemical dependency treatment admission rates and use patterns as well as from extensive qualitative interviews with key informants working in both systems. The use of qualitative data has proven to be invaluable in interpreting quantitative data and informing the analysis of state-level data on access and use of treatment services. Discovering the differing impacts of managed care and fee-for-service systems on publicly funded chemical dependency treatment enables researchers to inform public policy on these issues.

Coordination and Integration of Services to Adolescents in Need of Chemical Dependency Treatment in Oregon and Washington, Jane Grover, RMC Research Corporation

This presentation explores the interplay of service delivery systems in providing treatment and support services to adolescents from low-income families. In Oregon, public policy initiatives to encourage coordination and collaboration between chemical dependency treatment providers, mental health, child protective services, juvenile justice, schools and families are underway in 22 counties. Extensive qualitative interview data compares the integrated services in Oregon's rural Jackson County with successful collaboration in Washington's King County, Seattle's diverse and complex metropolitan area. Qualitative findings concerning what drives collaboration efforts, facilitating factors and barriers form the basis for this paper. Themes include the role of personal relationships, formal agreements, legislation, common protocols, interagency staffing, and other factors in enhancing or impeding collaboration. Implications for the field in terms of addressing gaps in adolescent treatment services and removing barriers to successful service integration will be explored.

Service Providers' Experiences in Providing Chemical Dependency Treatment to Publicly Funded Adolescents in Oregon and Washington: Implications for Improving the Provision of Treatment, Katherine Laws, RMC Research Corporation

In May of 1995 Oregon's publicly funded chemical dependency treatment system transitioned from Medicaid fee-for-service to managed care under the Oregon Health Plan. Prior to the implementation of managed care chemical dependency treatment providers in the state expressed concern about this transition, fearing that the quality of treatment would suffer, that the reimbursement rates for treatment would not be adequate, and that managed care organizations would not understand the ongoing needs of this population. This presentation will describe the service providers' experiences in providing or accessing chemical dependency treatment services for adolescents during this transition period to managed care. Service providers' experiences in Washington during this same time period will also be explored, comparing the similarities and differences between the fee-for-service system and the managed care system.

SESSION 212: Alternative Format **Room:** Kahuku

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

Creating Communities of Inquiry in Communities of Practice

Facilitators: Victor J Friedman, Ruppin Institute

Jay Rothman, Action Evaluation Research Institute

The object of the community of inquiry is to jointly develop a 'theory of action' for a particular problem/dilemma/issue in evaluation. This community of inquiry will look at the issue of encouraging self-evaluation and ways of integrating evaluation into the world of practice. Prior to the session, each participant must submit a brief personal case illustrating a successful or unsuccessful attempt to encourage such integration. The session chairperson will analyze the cases in advance, generate a composite 'map' of their implicit theories of action, and present this map at the beginning of the session. As a community of inquiry, the participants will test the map's validity by sharing their own experiential data and critically inquiring into the experiences of others. The ensuing dialogue will be aimed at refining the implicit theory and designing action 'experiments' for testing new strategies and generating knowledge.

SESSION 213: Panel **Room:** Oahu

Synthesizing Data: Quantitative and Qualitative Measures

Chair: Alicia Morales, University of Arizona

Panelists: *Investigating Rules and Principles for Combining Qualitative and Quantitative Data*, Alicia Morales, University of Arizona

Insistence on the necessity of combining data, e.g., qualitative and quanitative data, and on "taking all the variables into account," is ubiquitous. What is not ubiquitous are any rules or principles for doing that combining. Examination of many texts resulted in few rules and fewer actual guidelines. Humans, even when intelligent and well-trained, are poor at mental construction of what must amount to regression equations. The bases for establishing rules for combining information are examined, and several principles are adduced, along with recommendations for the kinds of development efforts that will be required if the insistence is to be taken seriously.

Synthesizing Multiple Dependent Measures, Erinn Herberman, University of Arizona

Analysis of multiple dependent variables is often a problem, e.g., requiring adjustment of alpha and producing confusing and conflicting results. Often the problem stems from failure to specify correct models for those variables. The model that is specified can have profound implications both for proper analysis of data and for subsequent interpretation of findings. Simplex, common factor, emergent variable, and independent effects models are possible specifications. Proper analyses for these models will be presented in the context of specific applications in research, and a novel, sequential path analysis will be described. Implications for interpretation will be discussed.

Combining Data: Examples from a Longitudinal Study of Child and Adolescent Development, Jennifer Bailey, University of Arizona

The task of combining qualitative and quantitative data can be a challenging one, especially when dealing with social scientific constructs that are difficult to measure. This presentation will provide applied examples of methods of data combination drawn from previous and pending research on a longitudinal sample of children and adolescents. There will be a focus on constructs that are difficult or controversial with regards to measurement and that often require qualitative data gathering techniques. Specific constructs will include childhood sexual abuse trauma, exposure to interparental violence, and peer violence and aggression. Potential applications of the strategies used to other situations will be discussed.

SESSION 214: Panel **Room:** Waialua

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

One REL's Response to the OERI Interim Review: A Quality Assurance Model for SERVE Projects and Products

Chair: Jerry Natkin, University of North Carolina at Greensboro

At the 1999 AEA annual conference, there was an OERI sponsored panel discussion of the Interim Review process undergone by the Regional Educational Laboratories in the spring of 1999. The presentations in the proposed panel session carries the discussion forward by describing how SERVE, the Regional Laboratory for the southeastern states, incorporated interim review panel recommendations into planning, development, needs assessment, and evaluation. The Interim Evaluation of the SouthEastern Regional Vision for Education Synthesis Report (June 1999) had numerous references to the need for SERVE to develop a quality assurance (QA) process for services. Since late summer 1999, the SERVE Evaluation Unit in collaboration with the Assessment, Accountability, and Standards program area has been working to develop a QA process for services. The four phases of the process are:

Phase I - Quality assurance planning
Phase II - Defining the model/product
Phase III - Piloting the model/product
Phase IV - Field Test Modal/Product (Research)

Critical components of this Quality Assurance Model are:

* SERVE Needs Assessment Process
* Research Base
* Clear and Explicit Outcomes
* Project Objectives
* Collaboration internally and externally
* Evaluation
* Evidence of sustainability - "Reflective Development Model"

Panelists: *Background and Rationale for the Development of a Quality Assurance Model*, Jerry Natkin, University of North Carolina at Greensboro

All the DIR evaluation criteria recommendations specifically had references to the need for a QA process. One overarching theme of the evaluation was the importance of developing a lab-wide QA process for R&D REL services. Dr. Jerry Natkin will provide the background and rationale that led SERVE to the development of "A Quality Assurance Model for SERVE Projects and Products." His presentation will focus on:

* the need for a quality assurance process,
* developing a flow diagram of the process,
* using logic models/"reflective development models" to describe the project, and
* how much of the work done within the laboratory supported a quality assurance process.

SERVE Needs Assessment, Susan Crawford, University of North Carolina at Greensboro

Dr. Susan Crawford will discuss SERVE's Evaluation Unit comprehensive needs assessment designed to:

* Provide a solid foundation of regional needs data
* Provide input to program planning
* Incorporate SEA and Legislative Activities
* Utilize findings from current literature
* Develop a mechanism to collect and routinely update information
* Provide needs data for strategic planning
* Address all renewals of existing contracts and provide the foundation for new contracts and grant applications

SERVE uses a variety of methods to assess needs. These methods include:

* Delphi needs assessment
* Assessing customers' needs through survey feedback
* Oversampling of designated stakeholder groups to address specialized needs
* Reviewing of annual reports, adding state/national political and legislative perspectives by SERVE state policy analysts
* Reviewing research literature, legislation, regional/national demographics, evaluation studies

Assessing the Impact of REL Projects on Student Achievement, Patrick H Harman, University of North Carolina at Greensboro

An examination of the impact of REL projects on student achievement is no trivial matter. The very nature of these initiatives makes it difficult to make unqualified causal statements about their impact, as different variables are involved that result in differing intensities and implementation at particular sites. Thus, the ability of a randomized experiment, generally considered the "gold standard" for inferring a causal effect of a treatment on a population, to untangle these effects seems unlikely. SERVE evaluators have attempted to determine the impact of its programs on achievement by using a regression - discrepancy model where the difference between predicted and actual achievement levels is used as a measure of program impact. Thus, positive residuals are considered evidence of site impact. Evidence for program impact is attained if a majority of site residuals are positive. Dr. Harman will discuss the pros and cons of approaching impact studies with this particular methodology and how SERVE evaluators have improved the model over time.

Piloting the Quality Assurance Process with One SERVE R&D Project, Pamela B Finney, University of North Carolina at Greensboro

In early fall 1999, the SERVE Evaluation Unit began working with an R&D project selected as a pilot for the Quality Assurance Model for SERVE Projects and Products. The project, "Building School Capacity through Effective On-Site Teacher Development," was in the formative stage. Evaluation staff and the project director began a collaboration that has resulted in many "lessons learned." A description of the project, results of the formative evaluation, and implications for use of the quality assurance process for future R&D projects and products will be discussed.

SESSION 215: Panel **Room:** Waianae

Sponsored by the Topical Interest Group on State and Local Government

Maximizing Evaluation Expertise through Organizational Skill-blending

Chair: John M Owen, University of Melbourne

In recent years, New Zealand has stood out in public policy and management circles for the breadth and depth of its public sector reforms. Massive institutional restructuring accompanied a comprehensive shift from input-based to results-based policy making. From virtual non-existence twenty years ago, professional evaluators have established a small, but respected, presence. Nevertheless, accountability requirements and policy-driven interest have spurred the search for strategies to make the most of the relatively limited organizational and professional evaluation capacity. The twin challenges of a major change in policy direction and a relative dearth of existing capacity have been met in the social policy area with complementary strategies. The first focuses on interdepartmental collaboration and the second on quickly growing evaluation skills and culture in a single large organization. The New Zealand experience contains useful insights for practitioners in progressive state and national governments.

Panelists: *Making Ends Meet in New Zealand*, Penelope J Hawkins, Ministry of Social Policy

State sector reforms have led to two linked demands for evaluation studies. Parliament requires regular reports from departments to account for outputs and demonstrate the achievement of outcomes. Parliament requires departments to demonstrate that policy advice is grounded in a good understanding of complex, changing social realities. This presentation describes the overall demand for evaluation expertise, with a particular focus on the social policy sector. In this area, the continuing economic and social gaps between indigenous Maori people and New Zealanders of European descent present both conceptual and technical challenges for evaluators. The presentation then provides an overview of the evaluation profession, characterized by a small pool of evaluation expertise and limited available training. Government agencies rely heavily on overseas hiring of experienced professionals and extensive 'on-the-job' training. Government agencies in the social sector are currently working together to address the challenge-capacity gap.

Theory and Practice of Inter-department Collaboration in Evaluation, David R Turner, Department of Labour

Given limited staff and dollar resources available for evaluation in the New Zealand central government, inter-department collaboration offers a way to increase evaluation capacity beyond what each agency could achieve on its own. This study builds on the work of Eugene Bardach, a prominent policy analysis theorist, who proposes a theory of interagency collaborative capacity in his 1998 Getting Agencies to Work Together. However, none of his cases provide instances of evaluators working together. A case study of the ongoing collaboration between three New Zealand departments examines key elements such as the development of effective working relationships at different levels, mutual intelligibility and trust, managing resources, and joint problem solving. The case study highlights the applicability of Bardach's theory to evaluation. While there are similarities between evaluation and other policy work, there are also differences. The conclusions draw potential lessons for improving collaborative evaluation.

Developing a Regional Evaluation Strategy, Kate McKegg, Department of Work and Income

The Department of Work and Income is one of the largest public sector organizations in New Zealand. In a constantly changing policy and operational context, a regional evaluation strategy has been developed to help balance external demands for accountability and internal demands for information that will assist in operational improvement. The strategy creates opportunities for field staff to participate in evaluations. Staff help decide on evaluation objectives and contribute their operational knowledge to the conduct of evaluations. Staff at all levels of the department have been involved in a range of evaluation activities, including evaluation concepts training, participating in evaluation working and steering groups, providing research assistance, sponsoring program evaluation, and providing input to evaluations. An evaluation needs assessment process has also begun, to tailor evaluation approaches and methods to the context and needs of the Department.

SESSION 216: Panel **Room:** Molokai

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

Building bridges: Practitioner knowledge construction and the academy

Chair: Susan R Boser, Cornell University

Research based in the academy brings with it systemic elements that inhibit opportunities for knowledge-construction that is generated by or with practitioners. Yet if, as a field, we are going to find ways to build evaluation capacity within and for organizations, we need also need to find ways to build bridges between practitioner and academic knowledge construction. This panel will feature presenters from Cornell University who each have found ways to integrate their commitment to practitioner theory building into their doctoral work at a this major research institution. The panelists will present these strategies and some of their "lessons learned", and discuss the implications for higher education. To enhance the opportunity for an interactive experience with attendees, these papers will be posted on the website of the Cornell Participatory Action Research Network (www.parnet.org) in advance of the conference.

Panelists: *Creating a Theory of Change Based on Practitioner Research*, Betsy Crane, Cornell University

As Senior Trainer and Collaboration Manager for the New York State Family Development Training and Credentialling Program, Betsy Crane focused her dissertation on the development of a theory of change for the program. She created a Program Logic Model based on feedback she had received about the program from community-based trainers, and her own knowledge as co-creator of the program and as someone who had worked in human services for 25 years. Interviews and focus groups conducted with a wide variety of other stakeholders, including workers/trainees and families/help-seekers, enabled her to build on this model to create a more grounded representation of the program and its outcomes. This practitioner-based research and theory development will be discussed in terms of its utility for construct development in program evaluation

How Extension Educators View Knowledge Generation, Heidi Haugen, Cornell University

Along with her position as an information and technology specialist with Cornell Cooperative Extension (CCE), Heidi Haugen focused her dissertation work on extension educators' epistemology - how educators within the land grant Cooperative Extension System view their relationship to knowledge generation within their educator roles. This work provided a fascinating exploration into educators' perceptions of their learning and knowing, of extension processes, and of the role of the university in community-based learning and community development practice. This inquiry also helped all its participants to better understand just what it means to "enable people to improve their lives and communities through partnerships that put experience and research knowledge to work," as the current CCE mission statement reads. Haugen will summarize the implications of her research for the continued evolution of Extension in it role as the structural link between the university and the community.

Action, Reflection, Interpretation and Representation: Quandaries in Praxis, Richard Kiely, Cornell University

As a doctoral candidate in Adult Education, Richard Kiely seeks to integrate the principles of participatory and collaborative approaches in his own work as a student. Kiely will share his experience utilizing this approach for a project in a qualitative methods class. He and his partners found this experience -- the collaborative doing of inquiry design, choice of methods, defining questions, gathering data, interpreting, representing -- called for more reflective attention than the original course-based framework intended. They found the co-generation of knowledge to be a dance of multiple subjectivities, replete with confusions and unanticipated dynamics of power and relationship. This paper explores how what we study and how we do it are inextricably part of what it means to be professional evaluators, practitioners, graduate students and people working to build bridges in the midst of diversity, power and multiple conceptions of "what is," and "what ought to be."

Building Connections through Research: An Action Research Dissertation, Susan R Boser, Cornell University

After 20 years as human service practitioner, Susan Boser used a doctoral dissertation as an opportunity to both support a regional, public agency action research project, and also to promote an agenda for enhancing the presence and legitimization of practitioner research within the academy. This action research project, funded by the Robert Wood Johnson Foundation, had two objectives. First, it sought to determine the feasibility of creating a regional, publicly managed, non-profit entity that could use managed care strategies to redesign the administration of public moneys for human services. Second, it endeavored to affect the policy-making processes at the state level by increasing the presence, credibility and capacity of the local stakeholders. Boser will present a summary of some of the challenges she encountered in endeavoring to link such practitioner research with doctoral dissertation requirements, and the strategies she found to support this connection.

SESSION 217: MultiPaper **Room:** Maui

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

Empowerment Evaluation: High Tech and Low Tech Capacity Building

Chair: David Fetterman, Stanford University

Empowerment evaluation is explicitly designed to build capacity. Funders, program staff members, and participants learn how to evaluate their own programs. Empowerment evaluations may have a high-tech focus using web-based surveys or a low-tech focus using traditional self-help evaluation manuals. They typically involve three basic steps: 1) developing a mission (values); 2) taking stock (assessing program components and establishing an evaluative baseline); and 3) planning for the future (establishing goals, strategies, and evidence to monitor change over time). This approach to evaluation helps people gain mastery over their own lives. Individuals who participate in an empowerment evaluation internalize the logic of evaluation. This helps them determine the proper use of specific evaluation tools. The data collected from this process is used to inform decision making. In the process of conducting an empowerment evaluation, participants are building capacity that is generalizable to other activities and programs in the community.

Presenters: *Empowerment Evaluation and Capacity Building*, Margret Dugan, Empowerment Evaluation Institute

Empowerment evaluation is a capacity building process that changes who is involved in the evaluation, what information is gathered and valued, how the information is handled and interpreted, and when the assessment process interacts with the process of improving the program. Developmental capacity building is fundamental to all empowerment evaluations. Because empowerment is a process by which individuals gain mastery or control over their own lives as well an increased ability to play an assertive role in controlling resources and decisions in their own context, it often requires skills and processes that participants need to develop. This brings new perspectives and challenges to the work of the evaluation. Outcomes most frequently associated with evaluation capacity include greater: participant and programmatic flexibility; evidence-based decisions by participants; shared responsibility for improved outcomes; higher standards of performance for both participants and the program itself; clarity about and commitment to the program; and the use of logic models to define program assumptions, processes and outcomes. Overall, empowerment evaluation is designed to increase workforce capacity.

Empowerment and Web Based Evaluation, Melissa Eiler & David Fetterman, Stanford University

Empowerment evaluation has an explicit focus on capacity building. A recent empowerment evaluation of a non-profit organization highlights the capacity of this approach to enable program staff members to evaluate themselves and improve program performance. The nonprofit uses the web to connect university researchers with community based organizations in need of research. The goals include engaging students and other researchers in community challenges and real-world needs and improving the effectiveness of nonprofits and public agencies. In addition to facilitating an empowerment evaluation, the nonprofit agency has hired the same evaluation team to conduct an external assessment of their services. This evaluation provides a test case concerning the compatibility of internal and external evaluation. It also highlights the value of web technology to facilitate both the internal and external evaluations. For example, one of the components of our external evaluation is an online survey. We developed this survey using a free on line service, Flashbase, which not only allows survey design, but also serves as the database for survey responses. Within one day of emailing a notice about the web-based survey to all users of the nonprofit's service, we had received 84 responses. The preliminary findings can already be presented to the client in both numerical and graphical formats. This tool was shared with the clients, in the process of presenting preliminary evaluation findings, further enhancing their capacity to function as a reflective, learning organization.

Empowerment Evaluation: Getting to Outcomes Guidebook, Pamela Imm & Abraham Wandersman, University of South Carolina

Guidebooks are typically written to build capacity. The Getting To Outcomes: Methods and Tools for Planning, Evaluation, and Accountability (GTO) guidebook is no exception. The GTO manual provides a user-friendly model of results-based accountability that can be useful to those working in many fields (e.g., treatment, prevention, and education). It was developed as a guidebook to provide practitioners with methods and tools to develop a comprehensive and systematic approach to accountability. The guidebook is designed to help practitioners plan, implement, and evaluate their programs to achieve results. GTO promotes user friendly methods for planning, implementation, evaluation, and sustainability. The material in GTO is based in part on a review of 35 books and/or manuals on evaluation GTO is based on answering 10 accountability questions that are logically linked and must be addressed to promote program accountability, focusing on needs and resources, goals, science and best practices, fit, capacity, plan, implementation, outcome evaluation, continuous quality improvement, and sustainability.

Discussant: David Fetterman, Stanford University

SESSION 218: MultiPaper **Room:** Kauai

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

Evaluating Mathematics Programs in a Number of Settings: What Adds Up?

Chair: Kathryn Race, Teachers Academy for Mathematics and Science

Presenters: *Differences in Fourth Grade Mathematics Achievement among Chilean Elementary Schools: An Application of Hierarchical Linear Models (HLMs)*, Janet Cadiz, University of California at Los Angeles

Because there are few external studies about the System of Assessing the Quality of Education in Chile (SIMCE), the Chilean Ministry of Education has little possibility of refuting or supporting its own SIMCE analyses. Existing analyses have primarily used only ANOVA and regression. Today, multilevel modeling techniques offer more sophisticated possibilities for analyzing the SIMCE assessment data. For example, questions related to the extent of differences in 4th grade math achievement among Chilean elementary schools with respect to the type of school administration (i.e. municipal, private subsidized, and private schools), and other factors can be better answered. Indeed, the examination of multilevel effects will add to new insights about student achievement. Therefore, the purpose of this study is to report on the application of Hierarchical Linear Models (HLMs) for reanalyzing 4th grade math student achievement by using the 1996 SIMCE data.

The Effects of the Connected Mathematics Project in Middle School Mathematics Achievement, Lesa Maria Covington Clarkson, University of Minnesota

The purpose of this study is to examine the three-year effect of the Connected Mathematics Project (CMP) on the mathematics achievement of middle school students in the Minneapolis Public School District. This will be accomplished by 1) comparing the mathematics achievement of eighth graders who have completed three years of CMP with the mathematics achievement of eighth graders who have completed three years of a traditional mathematics curriculum; 2) comparing the mathematics achievement of historically underrepresented students in both curricula; and 3) comparing the interaction and communication patterns in the two types of classrooms.

Sycamore Elementary Math Tutoring Program Evaluation, Rebecca M Eddy & Tiffany Hinz, Claremont Graduate University

Several math tutoring programs have been developed to address the deficiency of mathematics scores observed in elementary school students, however, few have subjected themselves to a critical formative evaluation of the program's merit. A new mathematics tutoring program was implemented in an elementary school in Southern California. The program consisted of undergraduate students from a local college tutoring small groups of students while enrolled in a class taught by the program director. The tutors used no set curriculum or strategies; however, topics addressed were based on state-mandated guidelines. The evaluation includes a synthesis of internal and external perspectives and uses the Key Evaluation Checklist as a guide for components essential to the evaluation. Evaluation methods included surveys to parents and tutors, interviews with the school principal, program director, parents, teachers and students, qualitative analysis of tutor field notes, pre-test and post-test scores with the intervention school and a comparable control group.

Assessing Student Perceptions of Classroom Methods and Activities in the Context of an Outcomes-based Evaluation, Kathryn Race, Teachers Academy for Mathematics and Science

Within the context of a larger evaluation, student perceptions were assessed regarding activities and methods used by teachers during mathematics and science instruction and other related activities within their school. 2,150 students from 10 elementary schools within the Chicago Public School system completed a 47-item questionnaire. The construct structure (based on factor analysis) identified three dimensions, i.e., Nontraditional Pedagogical Approaches in the Classroom, Hands-on/Cooperative Learning, and School/Learning Environment (internal reliability coefficients of .84, .82, and .75, respectively). Analysis showed a negative linear relationship between grade level and each of these sub-scales. That is, student perceptions declined from third through eighth grade ($r = -.30, -.32$, and $-.42$, respectively); mean attitude scores across grades were statistically significant ($p < .001$) with effect sizes of .10, .14, and .20. These attitude data mirrored performance based on state standardized math scores for these individual schools and statewide data. Implications of these findings are discussed.

Community Math Nights: Evaluating the Effectiveness of School-community Involvement, Patricia Kusimo & M Joy Riffle, AEL Inc

This evaluation supported school change collaborators as they attempted to promote high academic achievement of ethnic minority children who have been traditionally underserved by schools. Overall 73% of the 3rd to 8th grade students in a rural, public school district performed below the mean on the most recent Tennessee Comprehensive Assessment Program test. Fewer than 50% of 9th graders passed the math competency portion of the state test required for graduation. As part of this initiative, school district and educational laboratory personnel collaborated to design a community involvement program to improve mathematics instruction and parental support for innovative instruction. Three Community Math Nights were held across the county. At each, parents engaged with their children and teachers in mathematics activities that teachers planned to use. Parents' and students' attitudes toward math were assessed via a survey and anecdotal reports. To evaluate the success and impact of the initiative, survey data were collected from teachers and school personnel, and a focus group session was conducted with the district leadership team. Results showed that parents and children had much better attitudes toward math than anticipated. Parents and teachers' opinions of the math nights were equally positive but only one school chose to conduct similar math nights in the future. Implications and recommendations for the evaluation of collaborative efforts when one partner is not fully vested in the project are discussed.

Discussant: Naida Tushnet, WestEd