Thursday, November 2, 1:40 P.M. to 3:10 P.M.

SESSION 255: Presidential Strand Panel **Room:** Lanai

Sponsored by the Topical Interest Group on Non-profit and Foundations Evaluation

Developing Evaluation Capacity in Foundations: The Early Stages

Chair: Patricia Patrizi, Patrizi/McMullan Consulting

The purpose of this panel is to contribute to knowledge about the early stages of developing evaluation capacity within foundations. The panel builds on previous research on the potential of evaluation in foundations. Three case studies are presented which progressively demonstrate early stages of capacity building in foundations. The first case presents a foundation in the earliest stages of developing evaluation capacity as it moves toward defining its own evaluation policies. The second case study describes a foundation further along in capacity building, which it has done through the development and operationalizing of an internal foundation evaluation framework. The last case study presents the early stages of the evaluation of a foundation as a whole in light of strategic evaluation framework. The session concludes with a discussion of the early stages of capacity development and the facilitators that can move foundations to further capacity building.

Panelists: *Using Evaluation Policy to Build Evaluation Capacity*, Judith M Ottoson, Georgia State University

This case presents the efforts of a relatively new foundation to identify the role of evaluation. At its inception, the foundation incorporated evaluations variously in level of frequency, formality, and intent. The results of these early experiences encouraged the foundation to further develop its evaluation capacity. It chose to do this through the development of evaluation policy to answer such questions as: When to evaluation? How much to set aside for evaluation? Where to find evaluation resources? This policy focus in capacity development is less about how to do evaluation than how foundation policy makers think about evaluation. The development of the evaluation policy and its shaping factors are discussed along with the implications.

Integrating Evaluation within the Processes of Foundation Decision Making, Jane Arsenault, Independent Consultant

This presentation will describe a framework that was developed for the Baptist Community Ministries BCM trustee education that integrates evaluation within the processes of foundation decision-making from initial priorities to individual grants to issues of sustainability. The framework defines four quadrants of decision making activity: (1) zones of interest, (2) strategic resource deployment, (3) grants and grant project evaluation, and (4) uses of knowledge. The framework captures visually how these quandrants inter-relate and reinforce one another, first, to shape BCM as a learning organization and, second, to maximize BCM's impact. This framework will be used as the conceptual map for the continuing evolution and implementation of evaluation at BCM.

Operationalizing a Foundation Evaluation Framework, Pamela A Goodman, Baptist Community Ministries

Baptist Community Ministries is a four year old conversion foundation in New Orleans that, while still a young foundation, is at a later stage of development than the Foundation presented in the first case study. The BCM foundation had developed a strong framework of evaluation policy and has made a good start in establishing the function of evaluation within the foundation. Helping trustees at BCM understand evaluation and its uses is a continuing challenge. Following the presentation of the framework for integrating evaluation within the process of foundation decision-making, examples of the ways in which BCM operationalizes the framework through the use of tools and processes will be presented.

Evaluating the Foundation as a Whole, Patricia Patrizi, Patrizi-McMullan Consulting

An evaluation framework which assesses the work of the foundation as a whole will be presented. The assessment encompasses a reexamination of the core purpose and competencies valued by the foundation. The evaluation extends its reach to assess the strategies of the foundation as conceptualized and exemplified by the grants made and how the foundation is perceived by others in the community and fields of interest. Finally the evaluation raises questions regarding how the foundation is organized to achieve its desired aims

SESSION 256: Demonstration **Room:** Akaka

Sponsored by the Topical Interest Group on Human Services Evaluation

Empirically Positioning Programs for Sustainability

Chair: Dale C P Howard, Howard Research and Instructional Systems Inc

Presenter: Dale C P Howard, Howard Research and Instructional Systems Inc
Brian T Yates, American University
Peggy Ann Howard, Howard Research and Instructional Systems Inc

As evaluators, we often are asked to decide whether something is 'worth doing.' To some this simply means finding a justification for continued funding from the same or similar sources. To us, however, evaluation means something different: a commitment to a process of continual re-evaluation of the program and its relationship to the social, economic, and political dimensions of its community. Programs can, we believe, sustain themselves and even flourish in economically hostile environments by continually measuring, analyzing, and optimizing the amounts and types of resources used to implement procedures that change or maintain biopsychosocial processes within program participants that are responsible for targeted outcomes. Positioning for Sustainability is about reducing external resource dependency and increasing internal self-determination and self-reliance in human services programs. Our presentation uses three examples from actual services programs operating in Alberta, Canada.

SESSION 257: Alternative Format **Room:** Koko

Sponsored by the Topical Interest Group on Environmental Program Evaluation

Nonformal Environmental Program Evaluation: Where Are We?

Facilitators: Emmalou Norland, The Ohio State University
L Kate Wiltz, The Ohio State University

In May 2000 a small group of program evaluators, environmental educators, and researchers met at the Teton Science School in Jackson, WY. There, they began a dialogue on the opportunities and challenges in incorporating program evaluation into nonformal environmental education. Join us for a task-oriented exploration of a number of issues important to the emerging field of environmental program evaluation. With an eye on the needs of environmental educators and interpreters, we will look at the relationship between evaluation and research; the definition of evaluation and its role in nonformal environmental education programming; perceived costs and benefits of alternative methods and purposes of evaluation; and how the standards of practice in these fields can inform environmental program evaluation.

SESSION 258: Panel **Room:** Waimea

Sponsored by the Topical Interest Group on Cluster, Multi-site and Multi-level Evaluation

Practical Contrasts of Cluster and Multi-site Evaluations

Chair: Ricardo A Millett, WK Kellogg Foundation

Multi-site evaluation emphasizes confirmation and replication while cluster evaluation emphasizes learning. During this session one presenter will give an example of two national multi-site evaluations (focused on early childhood and high risk youth programs). A second presenter will give an example of a national cluster evaluation (focused on Community-University Partnerships). They will describe the methods and features of each as well as the evaluation capacities needed for conducting such evaluations. A third presenter will describe how a theory of the diffusion of innovations can be useful for both types of evaluation and the capacities needed to apply the theory. The discussant will bring to bear his own experience with multi-site evaluations to further illuminate the evaluation capacity that is needed for conducting each type of evaluation.

Panelists: *Multi-site Evaluations as a Means of Generating Science-based Intervention*, J Fred Springer, Evaluation Management and Training Associates Inc

The growing number of multi-level, multi-site evaluations, and the rapid development of analysis techniques for them, gives us a strong prospective evaluation tool. This presentation demonstrates how different forms of multi-site studies (clinical trials, cross-site evaluations, and meta-analyses) contribute to policy questions at different levels of development in the policy cycle. The presentation uses specific examples of policy lessons developed through two large national cross-site evaluations (Starting Early Starting Smart funded by SAMHSA's Office of Early Childhood and the Casey Family Program and the National Cross-Site Evaluation of High Risk Youth Programs for the prevention of substance abuse funded by the Center for Substance Abuse Prevention). Specific examples are used to demonstrate how multi-site, multi-level analysis contributes directly to the development of science-based lessons of program design and implementation.

Cluster Evaluation of a Community-University Partnership Initiative, Beverly A Parsons, InSites

Cluster evaluation involves sites that are using multiple evaluation models and those models are often evolving. Not all outcomes are specified in advance. The emphasis is on strengthening programs and exploring hypotheses and theories that are in their early stages of development. The work in the sites is fairly autonomous with limited centralized control. This presentation describes the cluster evaluation of a national Community-University Partnership Initiative designed to understand how social services professional development and practice can be shifted to build on a new philosophyùone that is more integrated and family-centered than is typically in operation. The presentation will focus on the evaluation capacities needed to undertake cluster evaluation as illustrated by this example.

Applying Diffusion of Innovation Theory to Cluster and Multi-Site Evaluation, James W Dearing, Michigan State University

Diffusion of innovations is an interdisciplinary research paradigm with particular merit for helping to conceptualize evaluation goals and objectives concerning novel, risky, experimental demonstration projects as well as tested, refined, exemplary demonstration projects. This presentation matches the goals of experimental demonstrations with those of cluster evaluations, and the goals of exemplary demonstrations with those of multi-site evaluations. Specific diffusion concepts are discussed that can be used to help orient each type of evaluation, quite apart from the substantive topics of the programs under study.

Discussant: Bobby Y Milstein, Centers for Disease Control and Prevention

SESSION 259: Panel **Room:** Niihau

Sponsored by the Topical Interest Groups on Evaluation Use & State and Local Government

Lessons From Evaluation Utilization: Facilitating Legislative Use of GPRA Reports

Chair: Elaine L Vaurio, US General Accounting Office

The evaluation utilization literature suggests that, in addition to an evaluation's credibility and timeliness, it's relevance to stakeholders' information needs and the early and continuing involvement of stakeholders in the design and interpretation of the study are critical to their utilization of an evaluation's findings. These variables have been considered important in facilitating evaluation use by agencies but not legislatures. For GPRA to be successful in improving government performance and accountability, however, agencies' performance reports must not only reflect the agencies' own information needs but also those of their legislative committees.

Panelists: *Improving the Usefulness of Performance Reports to Congress*, Elaine L Vaurio & Stephanie Shipman, US General Accounting Office

Our study of the usefulness of 3 agencies' performance information to congressional committees demonstrated that these agencies infrequently employed these practices with regard to their congressional stakeholders. This paper will describe options agencies could use to improve the usefulness of their performance reports to Congress. Some options are (1) to consult with congressional staff on their information needs or on their choice of performance measures, or (2) to clarify the links between expenditures, strategies, and objectives.

Moving GPRA from Promise to Common Practice: Developing a Pragmatic Theory To Support Utilization, Patria de Lancer Julnes, Utah State University

This paper develops a framework to strengthen the utilization of GPRA reports. GPRA presupposes that it would be a good thing for GPRA to be successful, and by implication, that sometimes such efforts do not work. In order for us to understand this problem, we need to have a theory of why previous efforts dealing with performance measurement have not worked. We also need a theory of what functions performance measurement serves when it works (i.e. what does it mean to work?) This paper will include: 1. a description of lessons learned from previous efforts, 2. theory of the function of performance measurement, 3. theory of utilization, and 4. a report on interviews.

The Role of Evaluators in Managing for Results, Maria P Aristigueta, University of Delaware

This paper will examine the role of evaluators in a managing for results system by 1) noting the placement of evaluation in state government; 2) examining the roles given to evaluators in a managing for results system; and 3) discussing the impact that placement may have on the role of evaluation in the state. The study will consist of comparisons among states previously identified with active evaluation functions. In addition, the State of Delaware's efforts to involve evaluation in the system and increase the use and support of the legislatures for the system, will be highlighted.

Lessons for Evaluation Utilization for State Legislative Program Evaluation Reports, Jane Fletcher, Florida State Legislature

Legislative program evaluation (LPE) units exist in over 3/4 of the states and issues pertaining to utility are of prime importance (NDE#81, 1999). These units generally perform a wide range of program evaluation and policy analysis services for their legislatures, and are deeply involved in the major policy issues facing state governments (such as the impact of welfare reform). While the units vary greatly in size and products, they collectively face unique challenges in promoting utilization of their work. These include the need to ensure that reports are fully responsive to legislative information needs and methodologically valid while still being timely and relevant. This presentation will address best practises that units have used to enhance the utilization of their products.

SESSION 260: Tech Talk **Room:** Hilo

Sponsored by the Topical Interest Group on Computer Use in Evaluation

The Development of An Impact Evaluability Relational Database for Conducting Impact Evaluation of HIV Prevention Efforts and Policies at the Municipal, State and National Levels

Chair: Timothy A Akers, Centers for Disease Control and Prevention

Presenters: Timothy A Akers, Centers for Disease Control and Prevention
Choi K Wan, Centers for Disease Control and Prevention
Xiaohong Mao Davis, Centers for Disease Control and Prevention

Since the onset of the HIV/AIDS epidemic CDC has directed a significant amount of resources to health departments and community-based organizations (CBOs) in an effort to combat the spread of the disease. Monitoring and evaluation is critical to this effort. CDC employs many different types of evaluations to address this issue, such as process, outcome monitoring, outcome evaluation, and impact, among others. However, apart from the standard types and methodologies employed for evaluation, CDC is working towards developing a new strategy that can link various types of data from different sources to assess the impact evaluability across HIV/AIDS data systems. In summary, this project is a relational database that integrates different variables from diverse HIV/AIDS databases and displays their variable characteristics and attributes for more broad-bases data analysis.

SESSION 261: MultiPaper **Room:** Puna

Sponsored by the Topical Interest Group on Teaching of Evaluation

Developing an Infrastructure that Supports Program Evaluation

Chair: Sarita Davis-Chukwuka, Clark Atlanta University

Presenters: *Increasing the Capacity To Develop and Evaluate Programs within Community-based Agencies*, Wendy A Etheridge, Anne E Farber & Debra L Stark, University of Pittsburgh

This paper describes a training model for teaching evaluation principles and skills to staff from community-based programs that increases their internal capacity for evaluation. The program model goes beyond providing skills and tools needed to "comply with funder requirements" and helps community-based agencies use evaluation for program development and management. The program consists of seven three-hour sessions that are held bi-weekly over a 14-week period. A significant portion of the training occurs in small groups, and time is allotted to address specific concerns of each agency. Participants apply what they have learned by developing a program logic model and an evaluation plan for a program that they wish to evaluate. Participants have consistently provided highly positive feedback over the three years that the training has been provided. The authors will give a brief overview of the training program and its impact on participants' internal capacity for evaluation.

"The Evaluation Workshops": Building an Evaluation Culture within the Occupational Safety and Health Service in New Zealand, Kathryn M Mickell, Occupational Safety and Health Service

The New Zealand Occupational Safety and Health Service (OSH) is moving towards an outcome focused environment, with less emphasis on the numbers and more on results. One step to achieving this new focus was for OSH staff to monitor the work they do and assess their impact on occupational safety and health. A project was developed to build an evaluation culture within OSH, its purpose being to develop an evaluation capability within the organisation. Evaluation workshops for staff were run in the central office and regional branches. The workshops highlighted that time, resources, ongoing support, realistic expectations and an environment conducive to evaluation are needed to build evaluation capacity effectively. This presentation will focus on the evaluation workshops and discuss the strengths and limitations found in to taking this approach.

Promoting the Transfer of Evaluation Training, Scott A Robbins, University of New Mexico

Evaluation in corporate training often fails to assess the connection between what occurs in the classroom and what occurs where the trainees work. This paper describes the results of a training design innovation intended to remedy this problem. The featured course on training evaluation is a component of a comprehensive five-day training staff development curriculum at a US Department of Energy National Laboratory. Training professionals are required to complete the program in order to meet institutional training staff qualification standards. The training evaluation course deviated from training design tradition by crossing the boundary that typically separates the training classroom from the trainee's workplace. In addition to in-class activities, trainees were expected to complete a field exercise that directly related the evaluation concepts from the course to their work situation. Lessons learned from the featured course will provide other evaluation trainers with ideas and strategies for bridging the workshop--to-workplace divide.

SESSION 262: MultiPaper **Room:** Kohala

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

A Funny Thing Happened on the Way to the Forum: Examining the Intersection of the Roles of Evaluators and Stakeholders

Chair: J Bradley Cousins, University of Ottawa

Presenters: *Revisting 'Responsiveness' in Evaluation*, J Bradley Cousins, University of Ottawa

Stakeholder participation in evaluation is increasingly prevalent as different forms of collaborative, participatory and empowerment evaluation gain popularity. Such approaches involve those trained in evaluation methods and applications working in partnership with program practitioners or other stakeholders who typically do not have such training. Collaborative forms of evaluation are touted to be responsive to stakeholders' needs and concerns by virtue of their inclusive character. Yet critics would argue that responsive evaluations can and should be successfully carried out without the direct involvement of non-evaluation personnel in the knowledge production function. Stake (1998), for example, suggests that he "would not put stakeholder evaluation into the same pocket" as collaborative and participatory evaluation. He advocates a more paternalistic role for the evaluator in "responsive evaluation". This paper traces the use of the construct "responsiveness' in evaluation from its early uses (e.g., Stake, 1973) through to the contemporary evaluation literature. The purpose of this expedition is to examine the evolution of the construct and to explore implications for evaluation theory and practice.

Navigating the Ghosts: What To Be Aware of When Building Capacity within Underserved Communities, Ruth E Forman, Pauline E Brooks, Sally R Durgan & Tenzing Donyo, The California Endowment

Increasingly the words 'community based,' 'participatory,' 'collaborative,' and 'capacity building' come up in the discussion of community health programs and initiatives. Relatedly, evaluation paradigms are shifting to meet these changes. But challenges arise when putting these terms into practice, especially within traditionally underserved communities. Some of these challenges include: 1) a community's history with research and evaluation; 2) expectations on the part of stakeholders regarding evaluation, and 3) the limited ability for an emerging collaborative paradigm to deal with stakeholders' conflicting agendas. Furthermore, what happens when ôcapacity buildingö for one stakeholder means something different than for another? These challenges house unspoken and unforeseen 'ghosts' which will hinder capacity building -- program capacity building or evaluation capacity building -- in underserved communities. This paper presents lessons learned from real life challenges and suggests ways for navigating the new territory.

Increasing Evaluation Capacity in Community Based Organizations: A Report from the Trenches, Louise A Godbold, Independent Consultant

"Participatory" and "empowerment" have become buzzwords in the world of evaluation. However modish it is to bandy these words in theory, this paper argues that when it comes to increasing capacity for evaluation in Community Based Organizations (CBOs) there is no other real choice. A "top-down" approach will not serve to convince CBO staff or community members that evaluation holds "anything in it for them" and will encourage them to do only the bare minimum to meet the funding source requirements -- or less if they can get away with it. The paper outlines three areas where there are opportunities for the determined evaluator to expose CBO staff to an experiential appreciation of the benefits of evaluation: the program conceptualization/proposal writing stage; the program stage; and the dissemination of findings stage. There are also some cautionary tales about the dangers an unwary evaluator is likely to meet along the way.

Evaluators as Stewards of Citizen Deliberation, Cheryl MacNeil, The Sage Colleges

In this paper I take the position that we have become a morally ill society - a collection of individuals who relegate citizen responsibility to a professional imperative. In response, I propose evaluators ought to be stewards of citizen deliberation in evaluation practice. Stewardship is the exercise of comprehensive and dedicated responsibility for a valued practice institution, resources relation or group...the responsibilities of stewardship go beyond mutual concern and call for deeper commitment to selected objects of moral concern (Selznick, 1998). Using examples from different evaluation studies, the paper will address how stewardship of citizen deliberation in evaluation practice can contribute to a greater public good. Concepts and practices for building citizen capacity during evaluation will be explored. Emphasis will be placed on creating deliberative forums that allow citizens of different privilege and perspective to share evaluative wisdom and personal insight. Selznick, P. (1998). Social Justice: A Communitarian Perspective. In A. Etzioni (ed). The Essential Communitarian. Oxford, Engalnd: Rowan and Littlefield Publishers, Inc. pp. 61-72.

Discussant: David Fetterman, Stanford University

SESSION 263: MultiPaper **Room:** Kona

Sponsored by the Topical Interest Groups on Theories of Evaluation & International and Cross-cultural Evaluation

Theoretical and Methodological Considerations for International Evaluations

Chair: Prachee Mukherjee, Colorado Springs School District 11

Presenters: *A Model and Method of Evaluative Accounts: Impact of National Literacy Mission of India*, Harbans S Bhola, Indiana University at Bloomington

Post-positivist discourses on evaluation should include statements of epistemology and ideology, the mean-and-ends calculus of the program being evaluated, justifications for choice of models for elaboration of discourse, and methods for data collection and interpretation. It is further suggested that post-positivist discourses be related to both program's context and program's history. This would require shifting from an evaluation "study" (a discrete, stand-alone, time-bound exercise) to an evaluative "account" (covering a time period in the life-history of a program, building on all earlier evaluation studies of a program). The model and method are tested in developing an evaluative account of the National Literacy Mission (NLM) of India during the last dacade to capture impact by design, impact by interaction, and impact by emergence.

Politicians Require Evaluation. But Do They Afford It? Savior Chircop, University of Malta

Different political systems look towards the US for inspiration attempting to emulate the success/systems of American Economy. However the underlying assumptions of such an economic-political framework are not easily applicable to international situations. (Chircop, 1990). This paper continues the discussion of such a dissonance in the context of a publicly funded National TV/Radio broadcasting network within a small Island Economy. The paper attempts to illustrate the hidden agendas enmeshed in such a decision making process. Furthermore, the divergent cultural and economic assumptions of the US and continental Europe are highlighted. Could the use of a wider stakeholder analysis make the outcome of such evaluations more relevant?

Human Rights Conventions: Monitoring and Evaluation Implications, Mahesh S Patel, United Nations Children's Fund

The development paradigm has become increasingly inclusive. An initial emphasis on economic issues was broadened by inclusion of social development issues such as nutrition, education and health. Another expansion is now taking place, focussing on freedoms and human rights such as political participation, access to information and due process of law. Several taxonomies of human rights are available. But these are often not well suited to monitoring and evaluation purposes. A taxonomy adapted from Amartya Sen's book "Freedom and Development" is proposed as a useful step in this direction. Rights based and rights promoting development programmes and a means of prioritizing their activities are discussed. Some current rights based monitoring and evaluation tools are presented and related to the Human Rights Conventions for which they are used. Several 'blind spots' for which new tools and concepts will be needed are noted.

SESSION 264: Panel **Room:** Ewa

Sponsored by the Topical Interest Group on Health Evaluation

Evaluation When the Community Is In Charge: Lessons from Evaluating Community-based Health Programs

Chair: Ross F Conner, University of California at Irvine

Increasingly, communities as a whole, not separate individuals, are the focus for health promotion programs and, therefore, also the focus for the evaluation of these programs. Because most past evaluation work has concentrated on individual-level processes and changes, this new focus on the community presents challenges and opportunities. In this session, community-based health program evaluators will draw upon on their experiences to highlight opportunities and challenges. Conner and Tanjasiri will discuss the development, implementation and analysis of measures of community-level changes in community decision-making processes. Smith will discuss the process of developing evaluation measures that are meaningful to community members and credible to outside decision-makers. Easterling will present approaches for involving community-based organizations in the evaluation of locally relevant outcomes in the midst of larger multi-site initiatives. Kuo and Campbell will discuss the challenges of clarifying intended program goals while balancing the interests of a philanthropic foundation and grant recipients

Panelists: *The Development, Implementation and Analysis of Community-level Measures of Change*, Ross F Conner & Sora Park Tanjasiri, University of California at Irvine

Conner, joined by Tanjasiri and other colleagues, evaluated the 7-year long, 28-community Colorado Healthy

Communities Initiative, which concluded at the end of last year. This initiative is typical of, and a model for, many "healthy communities" health promotion programs occurring around the US, as well as elsewhere in the world. In these programs, communities as a whole decide what the focuses of their programs should be, usually defining "health" very broadly. This presents a number of challenges for evaluation because the "program" is not specified at the outset and because communities usually have very ambitious, large-scale goals. In this AEA presentation, Conner and Tanjasiri will focus on one aspect of the evaluation: the development, implementation and analysis of community-level measures of changes in community decision-making (i.e., in the ways communities go about doing "community business") and the challenges of linking such changes to the local healthy communities program activities.

Community Indicators in the Evaluation of a Canadian Healthy Communities Initiative, Neale B Smith, Regional Public Health - David Thompson Health Region

The Healthy Communities Initiative (HCI) in the David Thompson Health Region (DTHR) began in 1998 with the participation of four rural and one urban community in central Alberta. The HCI involved community visioning, selection of key priority areas, and the implementation and evaluation of community action plans to address these health priorities. Facilitation, evaluation support, and funding for action plans were provided to community steering groups by the DTHR. The health region as well conducted parallel evaluation of the HCI as a health promotion/community development strategy. Both the region and communities have learned much from the evaluation, including the importance of community capacity building, and the challenges of ensuring broad participation and achieving effective intersectoral cooperation. In the context of these findings, Smith's presentation will highlight the development of community indicators -- evaluation measures that are meaningful for community members and yet credible to outside funders, service providers and decision-makers.

Moving from the Evaluation of Community-based Initiatives to Community-based Evaluation, Doug Easterling, University of North Carolina at Greensboro

Most community-based initiatives operate at two levels: the macro and the micro. At the macro level (corresponding to the perspective of funders and government agencies), the critical outcomes for evaluation relate to changes in programming, decision making and capacity across multiple communities. At the micro level (corresponding to the perspective of an individual community), the desired outcomes tend to be much concrete and community-specific (e.g., increasing the number of clients served by effective programs). In keeping with the logic and values of community-based initiatives, it is important that local evaluation reflect the questions, theories and assumptions of local program developers. Easterling will present examples of approaches that have proven effective in engaging the staff of community-based programs in the development of meaningful evaluation studies.

Increasing Evaluability: Balancing Community and Funder Interests, Victor Kuo & Martha Campbell, The James Irvine Foundation

A major challenge of evaluating any type of program with multiple interests, including those that are community-based, is that of specifying intended outcomes. Challenges are multiplied when a funder-grantee relationship is involved. The increased attention and focus on evaluation at The James Irvine Foundation, a large California private foundation, has required foundation staff to articulate clearly their goals; it also requires staff to work collaboratively with community members and program providers to specify intended outcomes. Evaluation unit staff members, Campbell and Kuo, will share challenges of clarifying definitions of "outcomes" and increasing usage of evaluation concepts among foundation staff as well as grantees. Examples from health and other foundation program areas will be shared.

SESSION 265: Panel **Room:** Honolulu

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

Using Case Mix in Outcome Monitoring: Findings from Three Longitudinal Studies of Substance Abuse

Chair: Michael L Dennis, Chestnut Health Systems Inc

While research often focuses on narrow homogenous clinical subgroups, actual programs typically serve more heterogenous client populations that may move across multiple levels of care or in and out of treatment. This panel will use data from three longitudinal studies of substance abuse treatment to explore three different uses of case mix adjustments to account for differences in severity and/or expect outcomes: 1) a cluster analysis of measures of clinical severity to predict treatment outcomes and allow comparisons of clients served across different levels of care; 2) a cluster analysis of days of using different substances to identify patterns of use, examine how the profile of use pattern have changed over a 10 year period, and discuss the implications for program planning; and 3) a segment analysis to predict outcome status and develop expected outcomes across multiple outpatient programs from several different sites.

Panelists: *Predicting Outcomes in Adults and Adolescent Treatment with Case Mix Versus Level of Care: Findings from the Drug Outcome Monitoring Study (DOMS)*, Michael L Dennis, Christy K Scott, Mark D Godley & Rod Funk, Chestnut Health Systems Inc

This paper evaluates a case mix adjustment based on a cluster analysis of presenting clinical issues (patient level data) with the initial level of care (program level data) in terms of predicting individual outcomes. The data are from the Drug Outcome Monitoring Study (DOMS) that conducted from 1995-1997. They include intake and follow-up assessments (93.4% of those attempted) with the Global Appraisal of Individual Needs (GAIN) on 424 adolescents and adults from 21 outpatient, methadone, and inpatient treatment units. Our case mix adjustment was worse then level of care in predicting retention (R2 =.68 vs. .88), but was otherwise comparable in terms of days in subsequent treatment (.25 vs. .24), alcohol use (.14 vs. .14), drug use (.15 vs. .18), illegal activity (.42 vs. .41) and employment (.61 vs. .60). This is important because in a continuum of care system patients typically go through multiple levels of care so the old approach is increasingly less useful.

Developing Alcohol/Drug Use Profiles from Assessment and Outcome Monitoring Data: Ten-year Trend in a Large Multi-functional Treatment Center, Brian Rush, Patrica Cleland & R Shaw Moxam, Centre for Addiction and Mental Health; and Michael L Dennis, Chestnut Health Systems Inc

Developing typologies of substance abuse has a long history in the addiction field As part of an outcome monitoring system developed for a large multi- functional treatment centre in Toronto, Canada, cluster analysis was used to analyse frequency of use across 13 drug categories (days of use in past 90 days) to develop groupings of clients useful for case-mix adjustments to compare outcomes appropriately and for other quality improvement purposes. While the full outcome monitoring system has been in place for three years, standardized assessment data on frequency of drug use are available on all clients for the past decade. This paper briefly describes the classification approach, its utility for case-mix adjustment and the longitudinal trends in alcohol/drug use patterns. The longitudinal patterns of use in this treatment population are compared to trends in population drug use in the centre's catchment area over the same period.

M*odeling Case Mix at Intake Based on Outcome Functioning*, Mark A Foss, Christy K Scott & Michael L Dennis, Chestnut Health Systems Inc

Problems and service needs vary widely among those presenting for substance abuse treatment. Differences in the composition of clientele can make comparison of treatment programs, even at the same level-of-care and seemingly from the same community, problematic. Case mix adjustment and risk assessment are two approaches, using different techniques, used to compensate for individual differences. This study presents a method for developing an "at-risk case mix model" base on client functioning post-intake (e.g., drug use, employment, health). A taxonomy based on client characteristics at intake is developed that distinguish between clients who are functioning more successfully six months post-intake from those who are not. A subset of data from a national study of substance abuse treatment systems in eight metropolitan areas was used for this study. As a demonstration, results comparing treatment outcomes among those receiving outpatient services across cites are presented- with and without an at-risk case mix adjustment.

Discussant: Lee B Sechrest, University of Arizona

SESSION 266: Roundtables **Room:** Kahuku

Minority, International and Cross-cultural Issues in Evaluation

(This session includes two 45-minute rotations of roundtables. The Host will ask the tables to rotate at 2:25.)

Host: To be announced

Roundtable A (First Rotation): Evaluating a National Land Mine Awareness Campaign in Mozambique: Key Results and Reflections, Tristi C Nichols, Manitou Inc

The paper presentation describes results and reflections of an evaluation conducted in Mozambique of a National Landmine Awareness Campaign. My discussion covers three points: 1) the context in which accidents associated with landmines/explosives shaped the awareness campaign; 2) the multiple media/materials used to communicate awareness messages; and 3) the multiple methods used to uncover program processes and the extent of community awareness over a four-year period (surveys, observations, and interviews). I am an African-American female, and my role as an independent evaluator was strengthened in part because of my language proficiency and prior experiences working in Africa. It appeared that the staff/partners were unguarded during interviews or informal interactions, as we established a high level of trust in each other and could confide in each other about confidential matters. After understanding the multiple staff/partner constructions, I made conscious attempts to improve staff's capacity to conduct assessments of program impact.

Roundtable B (First Rotation): Challenges of Evaluating Issues of Marginalized Populations

Presenters: *Documenting Outcomes of Domestic Violence in the Hmong Community: A Participatory Action Research Project*, Kalyani Rai, University of Wisconsin at Milwaukee

The goal of this paper is to share the results of a participatory action research project in which team of individuals representing domestic violence victims, advocates, community leaders, service providers, court system and law enforcement worked together to develop processes to document outcomes to measure the effectiveness of domestic violence initiatives in Milwaukee's Hmong community. It describes the methods we use and the processes we develop that address the goals of community participation, grassroots education and empowerment. Critical issues and new insights will be offered to others who are pursuing a similar approach. Implications of these results on developing culturally sensitive program planning and evaluation will be discussed.

The Impact of Mass Media Intervention Programs on Marginalized Members of Society, Ryan M Quist & William D Crano, Claremont Graduate University

In the United States, the government often has called upon the mass media to ameliorate important social problems. For example, millions of dollars are being spent on a mass-media approach to prevent drug abuse and increase knowledge of AIDS in young adolescents, especially adolescents of low SES and low acculturation, groups in which drug abuse AIDS risk is inordinately high. Our evaluation suggests that the most marginalized members of society are least likely to profit from mass mediated social intervention programs. Latino Americans at lower levels of acculturation exhibit patterns of media usage quite at variance with those of higher acculturated Latinos and Anglo-Americans. In addition, low acculturated Latino Americans demonstrate less knowledge of AIDS, less knowledge about the transmission of AIDS, and more prejudice against people with AIDS (PWA). This is problematic, because this pattern of low knowledge and anti-PWA attitudes has previously been associated with greater HIV risk.

Structural adjustment programs (SAP) represent the application of a standard package of measures or "conditionalities" imposed by the International Monetary Fund (IMF) on countries that seek World Bank financial support to address economic problems. The rationale underlying SAP is that markets are efficient whereas government interventions in resource allocation are distorting and inefficient. SAPs are envisioned to break the culture of government intervention in economic and other sector activities and spur long-term economic growth. Does the occurrence of SAP result in subsequent economic growth? The objective of the paper is to evaluate the impact of structural adjustment programs on economic growth in Africa. The study uses pooled cross-sectional time-series analysis to statistically estimate the impact of various independent variables including the implementation of SAPs on economic growth. This study is important because a better understanding of what generates economic growth in Africa, with its large population size, has the potential of making significant contribution to human welfare.

Roundtable C (First Rotation): Considering the Role of Culture in Program Evaluations

Presenters: *Multi-cultural Evaluation of Maori Tauira (Patients) in Queen Mary Hospital's Parallel Treatment Program for Addictive Disorders*, Gerald A Bunn, Evaluations and Auditing Services Ltd

The issue of evaluating Mental Health (A&D) programmes involving culture, or ethnicity, often demand attention being given to political as well as clinical issues. This is particularly true when treating Maori patients (Tauira) in New Zealand. Maori, the tangata whenua (people of the land), have a unique political relationship with their colonial partners. The Treaty of Waitangi, signed by the Queen of England and a number of Maori Chiefs, specify a partnership role that is singularly different than that which is experienced by the world's indigenous peoples, specifically its second clause which relates to health issues. In developing clinical methodologies, and subsequently evaluation efforts, for Maori, recognition was given to the fact that instruments developed along Eurocentric approaches "cannot be blindly applied to people of color." (Suzuki, et al, 1996) This report will review the steps taken in developing evaluation instruments and processes for Maori patients treated at Queen Mary Hospital's Parallel Programmes. Maori are treated in the hospital's Taha Maori Programme, while Pakeha (non-Maori) are treated in its Hanmer Programme. These separate, but interrelated, efforts are described by Maori as being a "basket weave", whereby each strand of flax (culture) retains its own color and design, but when woven together form a unique pattern (New Zealand society). A summary of findings will be presented along with observations relative to clinical and non-clinical issues that emerged over the five years of development (1994-1999).

The Assessment of Culture as a Protective Factor among Native Americans: The Survey of Nez Perce Culture, Elizabeth J Harris, Evaluation Management and Training Associates Inc and Joyce McFarland, Nez Perce Tribe

Evaluation of Native American programs in the past have often not considered the distinct history and unique culture in designing evaluation instruments. Therefore in our evaluation of Students for Success, a High Risk Youth Project funded by the Center for Substance Abuse Prevention, culture was seen as being extremely relevant and a definite area of interest with respect to the use of alcohol, tobacco and other drugs. Although the effects of acculturation in relation to substance abuse have been investigated among Spanish speaking peoples (Padilla & Snyder, 1992), no such studies prior to ours were conducted among Native Americans. We thus developed a Survey of Nez Perce Culture. This paper will present the survey, describe the methods by which it was developed, the process of pilot testing and the psychometric properties among both adult and adolescent populations.

Roundtable D (First Rotation): Crossing the Racial Divide: Minority Admissions in the 21st Century, Richard J Braxton, University of Wisconsin at Madison

Colleges and universities that successfully recruit and select minorities have institutional philosophies and structures that place a high value on diversity in their student bodies and embrace a system wide view of recruitment and student selection. More specifically, according to Fenske, R.H., Geranios, C.A., and Keller, Jonathan, E. (1997), successful minority recruitment and retention programs have the following components: 1) financial assistance, 2) collaboration among agencies and programs, 3) support services, and 4) information about college for parents and students. The purpose of this study is to assess the minority recruitment and selection practices of selective research universities by using the Fenske, et. al. (1997) activity categories to further specify factors affecting their success in attracting and retaining minority students. This study is designed to answer the following question: How do the recruitment and selection strategies and structures of research universities that are most successful in their efforts to enroll minority students differ from those found in less successful institutions? Case studies will be used to capture the full-context of the admissions and selection process. Admissions officers and other student affairs personnel will be interviewed for this study to discover how they recruit and select minority students. An interview protocol will be designed to provide each interviewee with information about the research and how it will be conducted (Creswell 1994).

Roundtable E (First Rotation): Development of a Rubric for Evaluating Cultural Aspects of Curriculum Modules for Navajo

Students, Rebecca H Zittle, Center for Educational Evaluation and Research; Elvira Largie, Navajo Education Technology Consortium; and Frank Zittle, Cameron University

The TECH Share project of the Navajo Education Technology Consortium (NETC) is developing standards based culturally relevant and appropriate multimedia curriculum modules for grades K-12 for delivery via the Internet. The target population for the project is teachers and students at underserved schools located primarily within the Navajo Nation. Employing a collaborative approach to the evaluation, the evaluators used a nominal group to involve stakeholders in the development of a rubric for evaluating the cultural aspects of the modules. In order to overcome barriers of time and place, the evaluators used electronic mail to facilitate the technique, enabling stakeholders from across three states to participate. Stakeholder representatives included project management, project participants, and the general Navajo community. Nominal group participants will discuss their experience of developing the rubric and how it contributed to evaluation capacity building.

Roundtable A (Second Rotation): The Role of Cultural Competence in Program Evaluation Settings

Presenters: *Cultural Competence in Human Service Programs: Its Definition, Measurement, and Role in Program Evaluation*, Cheryl A Hosley, Wilder Research Center

This presentation will summarize the results of an examination of cultural competence undertaken by the Wilder Research Center. A definition of cultural competence based on an extensive literature review and the results of a staff survey will be presented. Several strategies for measurement of cultural competence will be described, including the use of information collected from program clients and extracted from program records. Data will be presented summarizing the results of these measures across a variety of human service programs, including mental health services and residential treatment programs. These data will be used to illustrate the nature of cultural competence and the relationships between cultural competence and other program evaluation elements, including client satisfaction and service effectiveness. The presentation will conclude with recommendations for incorporating cultural competence into program evaluation and for using the results to enhance the provision of services to a culturally diverse clientele.

Evaluation Issues in Cultural Competence, Saumitra Sengupta, City and County of San Francisco and Rodney K Hopson, Duquesne University

This roundtable discussion will build upon the skills training session presented last year by the facilitators of this roundtable. The participants will identify and develop strategies in the areas of cultural competence in social service delivery.

Roundtable B (Second Rotation): Enhancing Cultural Diversity and Cultural Competence

Presenters: *A Model for Training and Evaluating Mexican Cultural Competence in Paraprofessional Human Service Agency Staff*, Angela G Montagno & David Boyle, University of Georgia

The number of Mexicans entering the United States in need of social services is increasing rapidly. This creates the need to educate human service agency staff, both professional and para-professional, in the Mexican culture. Previous research has focused on developing and evaluating cultural competence trainings for professional agency staff working with various populations in order to increase their practice expertise with clients. This paper discusses a cultural competence model for training paraprofessional staff in human service agencies. It uses a modified version of a cultural competence scale originally developed for social workers to measure cultural competence in paraprofessional human service agency staff before and after a training held in Northeast Atlanta.

Evaluation Participation in Culturally Diverse, Intergenerational Programs, Richard H Salo, The Ohio State University

As the population of older Americans increases in age and size, more advanced and specific procedures must be developed to evaluate culturally diverse, intergenerational programs. These programs, a vital ingredient for dynamic family ties, provide services and a heightened sense of community for both young and old. Beginning with a positive view of oneself, the individual becomes empowered in the community by countering perceived hopelessness with a newly discovered competence. If we are to continue to empower individuals and communities through intergenerational programs, we must be able to evaluate program effectiveness and recommend improvements. Scriven noted that program evaluation "must determine the worth or merit of the program being evaluated." Evaluators must meet this challenge. Our hypothesis is that program evaluators in collaboration with administrators will provide needed evaluation services in culturally diverse intergenerational programs to increase capacity and efficiency beginning in the year 2000 and beyond.

Roundtable C (Second Rotation): Considerations in Multi-cultural Evaluation

Presenters: *Ethical Issues in the Development of Standards for Multicultural Evaluation*, Elizabeth W Beale & James V Wright, Auburn University at Montgomery

This paper addresses issues in evaluation as they relate to multicultural populations. Approaches in evaluation for multicultural populations are neither systematic nor well defined, and frequently fail to adequately assess the gifts and talents of cultures other than Euro-centric. Ethical responsibility in evaluation recognizes the inherent dignity and worth of all cultures and attempts to identify these attributes, including awareness and knowledge of personal and structural biases that impact outcomes related to multicultural populations. This paper provides an overview of multicultural competencies as they relate to an organization's commitment to multicultural awareness and knowledge of personal and structural inequities. A three-tiered model for addressing competencies of awareness and knowledge is incorporated into strategies that promote skill development and multicultural education. The model includes recognition of biases that inhibit achievement, factual information about various cultural groups, and the integration of awareness and knowledge competencies into the development of standards that accurately assess multicultural populations.

An Analysis of Multicultural Competency (MCC) Training in Counselor Education Programs in a Midwest, Tri-state Region of the United States, Jennifer E Williams & Michelle Neace, University of Cincinnati

While counseling supervision and service delivery have been researched extensively by the mental health profession, the body of knowledge related to providing training for future counselors, especially cultural competence instruction, remains underdeveloped. This study investigates various methods of multicultural competence training in regional counselor education programs. The resulting data provides an in-depth understanding of how counselor educators attempt to integrate multicultural awareness, skills, and knowledge into the curriculum. Through this evaluation, the investigator compares Council for Accreditation of Counseling and Related Educational Programs (CACREP) accredited institutions' counselor education programs with non-CACREP accredited programs to broaden awareness of the types and potential effectiveness of some of the multicultural competence instructional methods currently in use in Indiana, Kentucky, and Ohio.

Roundtable D (Second Rotation): Evaluation Context and Brazilian Universities

Presenters: *Partnering for Change. Lessons Learned from the Training Partnership between the World Bank and the University of San Paolo, Brazil*, Mita Marra, The World Bank

Training partnerships between international technical assistance bodies and education institutions are pursued as an alternative, participatory way for human and institutional development and as a challenge of internationalization by institutions in the field. Building upon the experience of the training partnership between the World Bank and the University of San Paolo for the course on Macroeconomic Management to Brazilian central bankers, the paper analyzes training institutional capacity building and sustainability vis-α-vis the current social science literature on development, organizational theory, and political economy of partnership. The focus of the analysis will be on financial issues which are of crucial importance for the success of the common undertaking. Finally, the paper draws some policy recommendations which are relevant to the American and international audience in the light of the intensification of partnerships for knowledge generation and sharing throughout the world.

Evaluation in Brazilian Public Universities: Questions Related to Support and Conflicts, Iduina Mont' Alverne Chaves, Federal Fluminense University and Sueli Barbosa Thomaz, University of the State of Rio De Janeiro

This paper presents some considerations related to problems caused by the Evaluation Program being implemented in Brazilian Public Universities (Programa de Avaliaтпo das Universidades Brasileiras) since 1995. This Program defines evaluation as (1) a continuous process of developing the academic performance of the teachers, (2) this process takes into consideration the institutional identity (3) and involves accountability In this sense evaluation has a pedagogical character and it is essential to the institution development. Besides, the theoretical perspectives of the project include activities such as teaching, research, continuing education and management. The program focus on increasing the quality of the institution, taking into account its internal side, I mean, the institution structure, the pedagogical project and human resources. However the implementation of this evaluation project in Brazilian public universities has been quite different from its original proposals.. Neither the faculties nor the students are involved in the evaluation proceedings, there is no training courses to prepare professional evaluators and the evaluating groups operating the Program lack in knowledge of the recent evaluation theories. In sum, the government evaluation process in practice, is being diverted from its original aims and is viewed by the faculties and administrative personel as a process to punish or reward the university teacher by the volume of his production. There are few specialization courses on evaluation in Brazil and this leads to a process of evaluation without a proper academic knowledge and suitable preparation of evaluators to implementing such ambitious programs throughout Brazil. The people involved in these programs have been inadequately trained and operate under too many and varied constraints. I think this study can clarify and point out the weaknesses of the program aiming to contribute to improve such systems.

Roundtable E (Second Rotation): Evaluating Development Programs

Presenters: *Evaluation as a Participatory and Learning Tool in UNCDF*, Sophie de Caen, United Nations Capital Development Fund

The United Nations Capital Development Fund (UNCDF) is an organization working under the overall direction of the Administrator of the United Nations Development Programme. UNCDF's goal is to help reduce poverty in the poorest countries and has two pillars: local governance and microfinance. Empowering local communities through access to basic services, microfinance and participation in decision-making is key to attaining this goal. UNCDF's guiding principles in evaluation include: transparency and accountability; learning from experience; participation; capacity building; and mainting independant external evaluations. The paper briefly describes UNCDF's evaluation process for local development programmes while focusing on innovations such as annual project self-assessments; pre-evaluation qualitative studies; learning from thematic evaluations; and stakeholder ownership of evaluation results. The paper also presents UNCDF's strategy for generating lessons learned and feeding them back into projects, towards central government to impact decision-making on issues related to decentralization and back into UNCDF programming and policy refinement.

The Impact of Structural Adjustment Programs on Economic Growth in Sub-Saharan Africa, Al Bavon, University of North Texas

Structural adjustment programs (SAP) represent the application of a standard package of measures or "conditionalities" imposed by the International Monetary Fund (IMF) on countries that seek World Bank financial support to address economic problems. The rationale underlying SAP is that markets are efficient whereas government interventions in resource allocation are distorting and inefficient. SAPs are envisioned to break the culture of government intervention in economic and other sector activities and spur long-term economic growth. Does the occurrence of SAP result in subsequent economic growth? The objective of the paper is to evaluate the impact of structural adjustment programs on economic growth in Africa. The study uses pooled cross-sectional time-series analysis to statistically estimate the impact of various independent variables including the implementation of SAPs on economic growth. This study is important because a better understanding of what generates economic growth in Africa, with its large population size, has the potential of making significant contribution to human welfare.

SESSION 267: Tech Talk **Room:** Oahu

Sponsored by the Topical Interest Group on Quantitative Methods: Theory and Design

Innovative Applications of Structural Equation Modeling Techniques to Explore National Survey Data

Chair: Astrid Hendricks-Smith, WK Kellogg Foundation

Presenter: Cynthia C Phillips, Third Sector Strategies
Jeffrey Greene, Third Sector Strategies

Presenters will share the findings of their dissertation work at Western Michigan University. This session will explore how SEM can be used to explore and understand respondents mental models from a variety of perspectives: systemic educational reform and medical certification. These dissertations represent cutting-edge evaluation methodologies that apply SEM in new and innovative ways of interest to a variety of evaluation audiences and as such the presenters will include a basic primer on SEM and illustrate its use in relatively non-technical terms for participants.

SESSION 268: Panel **Room:** Waialua

Sponsored by the Topical Interest Groups on Pre-K-12 Educational Evaluation & Collaborative, Participatory and Empowerment Evaluation

Building Evaluation Capacity for Staff in PreK-12 Schools: Experiences and Challenges

Chair: Jean A King, University of Minnesota

Building capacity for PreK-12 educators to engage meaningfully in program evaluation and data use for school improvement is challenging work. Issues we have identified in attempting this work will be raised through presentation of the experiences of evaluators or evaluation consultants from six different contexts, both within and outside school settings. Issues that will be explored include training and technical assistance, structural conditions and supports needed for ongoing evaluation, and roadblocks such as distrust of the products of measurement and evaluation and the perceived incompatibility of high stakes accountability measures and evaluation for school improvement. Following the panel's presentation, the audience will be engaged in small group consideration of the issues raised, with feedback and discussion carried into the Pre-K-12 TIG meeting.

Panelists: *Where There is a Will, There is a Way: PreK-12 School Evaluation*, James R Sanders, Western Michigan University

Program evaluation does not need to be complex or inordinately time consuming. It does not require extensive technical training. What it does require is a DESIRE to improve one's school, a WILLINGNESS to work collegially, careful ATTENTION TO DETAIL, and BASIC KNOWLEDGE of how school program evaluations should be carried out.

Professional Development for Internal Participatory Evaluation: A Competence-building Approach, Astri M Indrebo, Lillehammer University College

The paper examines the effects of a professional development training program, building evaluation capacity in schools. Researchers from the Department of Education and Social Science at the University College of Lillehammer in Norway collaborated with the regional school authority to arrange training courses, which focused on enhancing the schools own capacity for doing systematic, collaborative inquiry. Teams of teachers and school leaders from the same school were attending a training program spaced over one year, and during this time the teams collectively planned and conducted collaborative evaluation projects in their respective schools together with their colleagues. The paper, building on a three-year longitudinal case study of the involved schools, will discuss some of the challenges and barriers met implementing this evaluation strategy.

Complexity in Teacher Professional Development: A Qualitative Study, Michael Sikes, Cultural Education Collaborative

Presenters will share the findings of their dissertation work at Western Michigan University. This session will explore how SEM can be used to explore and understand respondents mental models from a variety of perspectives: systemic educational reform and medical certification. These dissertations represent cutting-edge evaluation methodologies that apply SEM in new and innovative ways of interest to a variety of evaluation audiences and as such the presenters will include a basic primer on SEM and illustrate its use in relatively non-technical terms for participants.

Using a Continuous Improvement Model for Building School Capacity for Evaluation, Margarida F Southard, Leon County Schools

Building the capacity for evaluation in K-12 schools has been a priority for Leon County schools since we adopted standards-based reform and developed a data-driven school culture focusing on improving student performance. The purpose of this endeavor is to enable teachers to design instructional strategies to meet the needs of individual students. Using the framework for developing school improvement plans (problem solving model), the district constructed a model for developing school capacity for evaluation, ranging from use of data to track individual student progress to school-based research. The paper will present the components of this model, the activities and strategies being implemented, the constraints district and schools face for their implementation and next steps. Topics for discussion will also include: role of evaluator, the increase demand for testing and accountability, and the technology to access and analyze student performance.

Building Head Start Capacity for Outcome Evaluation, Mary L Marx & Debbie Zorn, University of Cincinnati

An evaluation framework which assesses the work of the foundation as a whole will be presented. The assessment encompasses a reexamination of the core purpose and competencies valued by the foundation. The evaluation extends its reach to assess the strategies of the foundation as conceptualized and exemplified by the grants made and how the foundation is perceived by others in the community and fields of interest. Finally the evaluation raises questions regarding how the foundation is organized to achieve its desired aims

Building Evaluation Capacity in Lake Woebegone ISD: How Many Minnesotans Does It Take To Use Data? Jean A. King, University of Minnesota

The challenges of building evaluation capacity in Minnesota's third largest school district is the topic of Jean King's paper. Building on Louis, Kruse, and Bryk's notion of professional community (1995), the district's Assessment Department has worked to build the structural conditions and social and human resources to support the process of evaluation at the district office and in its 40 buildings and to facilitate the use of data and evaluation information. The paper will describe the evolution of the Department's developmental approach to program evaluation and discuss specific challenges at four levels (district, program, building, and classroom).

SESSION 269: Panel **Room:** Waianae

Sponsored by the Topical Interest Groups on State and Local Government & Evaluation Use

There Is More to Performance Measurement than Just Counting Inputs and Outputs

Chair: Rakesh Mohan, Washington State Joint Legislative Audit and Review Committee

Performance measurement poses many questions and challenges to evaluators and to users of such information. One cannot just count and compare inputs and outputs of a program! The panelists will discuss answers to such questions as-Who uses performance measurement information and for what purpose? How do you measure performance of hard-to-measure programs? What are the barriers to implementing effective performance measurement systems? What are the effects of the use and misuse of performance information?

Panelists: *Who Wants To Know What Information, Why, When, and in What Format: The Performance Measurement Dilemma*, John A McLaughlin, University of Southern California

Measuring the performance of social programs has become a high priority throughout the public, private and not-for-profit sectors nationally and internationally. Generally speaking, the purpose of performance measurement is to reduce uncertainties about the social program of interest. But who's uncertain about the program? Why do they care? A major dilemma challenging those of us who are charged with the development, implementation and use of effective performance measurement systems is to clarify at the start of the planning process who wants to know what about the program, why, when, and in what format. The solution of the dilemma will lead the developer down a pathway that is both procedurally and substantively different.

Measuring Hard-to-measure Outcomes, Joseph S Wholey, US General Accounting Office

This paper will explore the use of qualitative methods, impact evaluation, and cost-benefit analysis in measuring agency and program performance. Examples will be drawn from federal agencies and programs in which there are difficulties in using typical performance measurement systems to measure performance.

Lessons Learned from Local Government Performance Measurement Use, David J Bernstein, Montgomery County Department of Finance

This presentation will summarize the results of five case studies, including cross-case analysis, of local governments with extensive experience using performance measures. The research was sponsored by the Governmental Accounting Standards Board under a grant from the Alfred P. Sloan Foundation. The study involved site visits, secondary analysis of existing case studies, and analysis of documents that incorporated performance measures. The case study design involved in-depth analysis of "descriptive" multi-case studies (Yin, 1993, p. 22) and an "explanation building" data analysis strategy (U.S. GAO, 1990, p. 63). The case studies were used to construct an "organizational biography" (Schein, 1984, p. 456) from documents and interviews to identify periods of culture formation within each organization in order to assess how performance measures were used, and the effect of use. Information will be presented on use of performance measures, the effect of use, performance measurement quality, barriers to implementation, and lessons learned.

Discussant: Rakesh Mohan, Washington State Joint Legislative Audit and Review Committee

SESSION 270: Panel **Room:** Molokai

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

Building Evaluation Capacities in International Aid Organizations and in Developing Countries - Emerging Lessons and Some Future Challenges

Chair: Patrick G Grasso, The World Bank

Evaluation has become increasingly important to developing and transitioning countries. While there are a number of factors that drive this demand, two are primary. First, many donors, both national and international, are asking questions about the effectiveness of their aid, and calling for accountability from both recipients and providers. Second, evaluation is being viewed as an important part of sound governance. However, developing countries typically lack the capacity to implement sound evaluation programs. This panel looks at experiences with and lessons learned from efforts on the part of donor organizations and developing country governments to build evaluation capacity.

Panelists: *Building Evaluation Capacity at the World Bank: Lessons Learned in the Context of Institutional Change*, Francois Binder, The World Bank

Since the early 1990s the World Bank has recognized the need to strengthen monitoring and evaluation of operations in order to improve development effectiveness. Several studies during the mid 1990s underscored the dismal track record of M&E at the Bank. Although much progress has been made in recent years to strengthen the use of performance indicators, recent work demonstrates that M&E systems remain weak. This presentation will highlight some of the underlying causes of weak M&E, emerging success stories, and current challenges for improving M&E.

Building Evaluation Capacity: Challenges Facing Developing Countries, Osvaldo N Feinstein & Keith Mackay, The World Bank

The new paradigms for development include partnerships and a focus on results. A key constraint is the enormous capacity-building challenges facing developing countries, including for monitoring and evaluation, a severe shortage of skilled evaluators and social scientists. And the level of demand from governments - important to achieve a high level of evaluation utilization - is low. There also is challenge in international aid organizations to achieve evaluation utilization internally, and to support countries in their own capacity-building. This presentation looks at the experience with evaluation capacity-building of both sets of players in the development business, governments and aid organizations, and considers synergies between both sets of activities.

Evaluation Capacity Building Using Distance Learning Techniques, Eckert William, The World Bank

The World Bank and other international organizations recognize the growing demand for evaluation capacity, and have offered assistance in the form of training and technical assistance. These donor groups have come to recognize that effective evaluation assistance can be expensive to deliver and faces problems with access to those most in need of assistance. To overcome these barriers to providing effective evaluation capacity building, the World Bank Institute in cooperation with the Bank's Operations Evaluation Department provides training and other assistance using Distance Learning technologies. This presentation will describe the use of these technologies in evaluation capacity building and assess their cost and other effectiveness vis-a-vis more traditional (face-to-face) methods of training and technical assistance.

Building Evaluation Capacity: The USAID in Bulgaria and Georgia, Laurie E Ekstrand, US General Accounting Office

This presentation looks at the process of building evaluation capacity in two countries with different economies in transition. It compares experiences from U.S. Agency for International Development-supported efforts to build capacity in Bulgaria and Georgia.

Discussant: Patrick G Grasso, The World Bank

SESSION 271: Alternative Format **Room:** Maui

Sponsored by the Topical Interest Group on Independent Consulting

So You Want to Be an Independent Consultant?

Chair: Belle Brett, Brett Consulting Group

Presenters: Jo Ann Doino-Ingersoll, Strategic Research
Melanie A Hwalek, SPEC Associates
Kris Putnam, Independent Consultant
Michael Scriven, Claremont Graduate University

Building capacity for Pre-K-12 educators to engage meaningfully in program evaluation and data use for school improvement is challenging work. Issues we have identified in attempting this work will be raised through presentation of the experiences of evaluators or evaluation consultants from six different contexts, both within and outside school settings. Issues that will be explored include training and technical assistance, structural conditions and supports needed for ongoing evaluation, and roadblocks such as distrust of the products of measurement and evaluation and the perceived incompatibility of high stakes accountability measures and evaluation for school improvement. Following the panel's presentation, the audience will be engaged in small group consideration of the issues raised, with feedback and discussion carried into the Pre-K-12 TIG meeting.

SESSION 272: Panel **Room:** Kauai

Sponsored by the Topical Interest Group on Qualitative Methods

Mixed Methods Analysis Using Concept Mapping

Chair: William Trochim, Cornell University

Concept mapping is a mixed methodology in which multiple stakeholders can either generate or code text information into a set of discrete text entities, individually organize these into themes or categories, and rate each text entity on any number of dimensions. The data are analyzed with multidimensional scaling and hierarchical cluster analysis. The main result is a thematic 'map' portraying all of the original text entities, how they are grouped into themes based on stakeholder sorting, and how they are rated. The three papers in this panel describe real-world applications of concept mapping in qualitative research: in a participatory, collaborative, multiple stakeholder qualitative analysis (Lowe); as a basis for developing in-depth interviews (Penney); and to both analyze open-ended survey questions and simultaneously do a form of network analysis based on the open-ended responses (Jackson). The discussion will critically assess the presentations and consider additional qualitative and mixed-methods applications.

Panelists: *Participatory, Multiple Stakeholder Qualitative Data Analysis Using Concept Mapping*, Staci Lowe, Cornell University

The overall objective of this project is to test the usefulness of using concept mapping as a method of qualitative data analysis. The substantive content is drawn from a study which looks at how urban adolescents are experiencing and perceiving welfare reform policy changes. With concept mapping, one can broaden the diversity of traditional analytical processes by including participation from a range of constituents, in effect creating a participatory, collaborative, mixed-methods form of thematic analysis. Using verbatim interview transcripts, a range of stakeholders (e.g., evaluation experts, youth, and administrators) will participate in the three-stage mapping process comprised of data reduction, sorting, and rating. Their input will be entered into the concept mapping software and a visual map will be generated, showing the interrelationships among all the data, how they are organized into clusters, and their relative importance as judged by participants. Findings related to process and outcomes will be presented.

Using Mixed-methodology Concept Mapping for Community-based Planning and Evaluation, Naomi G Penney, Cornell University

This paper will describe a multiple case study evaluation of four community-based coalitions to assess the ways and to what extent target-population community members have on coalition planning and outcomes. Concept mapping was used to assess coalition-member perception of community member role and its effect on coalition goal attainment. The results of the mapping project then served as a basis for in-depth interviews with coalition members to assist with map interpretation and a deeper understanding of the circumstances in which target-population community members can best serve coalition planning and outcomes.

The Application of Concept Mapping to the Analysis of Open-ended Survey Questions: Developing a Coding Scheme, Assessing Similarity, and Linking People to Ideas, Kristin Jackson, Cornell University

The purpose of this paper is twofold. First, it will demonstrate how the concept mapping methodology (Trochim 1989) can be used to develop a less researcher-biased structured conceptualization in coding open-ended survey responses than is possible with traditional thematic analysis. Decisions about data reduction, coder selection, and data interpretation will be discussed through an example of team-based data. Reliability and validity issues associated with the two methods will be compared. Second, the analysis will demonstrate how the concept mapping method can be extended to link people with categories of ideas. The same data gathered from teams will be used to show how each individual's responses can be transformed into a map of the similarities between all individual respondents which in turn can be used to assess the similarity of responses at the team-level. Finally, it will be demonstrated how the method can be applied to assess changes in individual and team responses between measurement intervals.

Discussant: William Trochim, Cornell University