Friday, November 3, 3:30 P.M. to 5:00 P.M.

SESSION 455: Presidential Strand Roundtables **Room:** Waianae

Evaluation Capacity in Diverse Organizational Environments

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

Host: Bobby Milstein, Centers for Disease Control and Prevention

Roundtable A (First Rotation): Building Commitment to Evaluation in an Educational Setting: The Implementation of an Organizational Framework for Program Evaluation at Milton Hershey School, Tara L Atkins-Brady, Milton Hershey School

At Milton Hershey School (MHS), a private residential school for needy children, individual school programs support a single mission: to nurture and educate the children we serve. Ultimately, it is evaluation of the development, implementation, and impact of programs that holds us accountable to this mission. Therefore, we must not only build capacity within our organization to evaluate individual programs, but also implement an overall framework for coordinating these evaluations to facilitate, across programs, collaborative, timely collection and analysis of information to inform decision-making. To this end, the MHS Office of Research and Evaluation has developed a process model detailing how we will develop, implement, coordinate, monitor, and refine program evaluations throughout the school. This paper discusses the rationale for this model and its key components, including how we are building commitment - through communication, training, and consultation - to an organizational culture that values evaluation.

Roundtable B (First Rotation): The Challenges and Rewards for the Design and Evaluation of Substance Abuse Treatment: The Provider Perspective, Michael W Kirby Jr, Nancy R VanDeMark, Deborah S John, Ellen Brown, Arapahoe House Inc; and G Nicholas Braucht & Charles W Reichardt, University of Denver

The national call to 'bridge the gap between research and practice' is currently beginning to have a major impact on the structure of and reimbursement for substance abuse treatment. In this context, under the rubric of 'research', is included an array of program evaluation activities. The purpose of this paper is to describe the challenges and rewards, from the perspective of a large non-profit substance abuse treatment provider, in engaging in the evaluation enterprise. This organization, Arapahoe House, is the largest provider of treatment services in Colorado. For over fifteen years Arapahoe House has had an internal department dedicated to program evaluation, and utilizes the services of external evaluation consultants to augment these evaluation efforts. This paper will describe: (1) the steering committee model deployed by Arapahoe House to effectively design and manage evaluation studies; (2) the kinds of challenges that the provider who elects to engage in these evaluation studies must surmount to be successful; and (3) the rewards for doing so.

Roundtable C (First Rotation): The Toughest Evaluation You'll Ever Love: Building Evaluation Capacity in the Peace Corps, Paula Bilinsky & Ana T Coghlan, Peace Corps

The Peace Corps works in over 70 countries through multi-sectoral projects and with a range of diverse stakeholders. Although project planning is initiated and conducted by field staff, Headquarters staff members are currently designing agency-wide evaluation processes and tools. The challenge is twofold: 1) designing monitoring and evaluation guidelines that enable the agency to capture its accomplishments, and 2) increasing project stakeholders' capacity to assess their own efforts. Thus far, staff have designed and conducted training on an integrated set of project planning, monitoring and evaluation guidelines and tools, including a project planning framework and a participatory monitoring and evaluation worksheet. Several important lessons learned include the need to:

* Develop meaningful across-the-board guidelines that are flexible enough for local adaptation
* Focus on concepts and thought processes, and not on precise definitions
* Encourage, facilitate, and track stakeholder adaptation

* Continuously check consensus among agency stakeholders

Roundtable D (First Rotation): The Status of Evaluation in HIV Prevention: Community-based Providers' Perspectives on Challenges to Evaluation and Its Use, Robin L Miller & Barbara J Bedney, University of Illinois at Chicago; and Heather Barton-Villagrana, Illinois Center for Violence Prevention

Over the last decade, evaluators have become focused on building the capacity of not-for-profit organizations to evaluate their own work. Not-for-profit organizations are facing increasing pressure to engage in more sophisticated evaluation activities. Little empirical data is available to describe whether and how such organizations cope with these evaluation pressures and how organizations use evaluation data. In this presentation, we will describe the results of analyses from 41 interviews with a statewide random-sample of not-for-profit HIV prevention service providers. The interviews were largely focused on the development of HIV prevention programs, but also included discussion regarding how the organizations determine that they are meeting their program goals, identify needs, and use evaluation information. Implications for evaluators' capacity development efforts will be discussed.

Roundtable E (First Rotation): Assessing and Increasing the Evaluation Capacity of Health Departments, Gary Uhl & Huey T Chen, Centers for Disease Control and Prevention; Dan Gentry, St Louis University; and David Cotton & Kira Sloop, Macro International

Shrinking resources and the continued transmission of HIV have placed an increased importance on the evaluation of HIV prevention efforts conducted by federally-funded health departments. This paper will describe a current study to assess and consequently build the capacity of health departments to conduct evaluations of their HIV prevention efforts. Presenters will discuss three methods for assessing evaluation capacity: document review, rapid evaluation needs assessment, and case studies. Current evaluation capacity was assessed via a document review of health department's applications for federal funding. The rapid evaluation needs assessment extends the document review with data collected from a survey completed by health department staff. The case studies are currently being developed and will rely heavily on data from the preceding two assessments. The rationale, procedures, and data from each of these methods will be described. Presenters will discuss how these data will be used to increase the evaluation capacity of health departments and to better target the provision of evaluation training and technical assistance.

Roundtable A (Second Rotation): Integrating the Hawaiian Value of "Ohana" To Build Culturally-sensitive Evaluation Capacity at The Maui Farm, Arnold J Love, Independent Consultant; and Paula Ambre & Donna Vida, The Maui Farm Inc

In Hawaii, community-based programs are facing the challenge of building their capacity to evaluate outcomes in an objective and culturally sensitive manner with very limited time and money. This poster describes the experience of The Maui Farm to strengthen its evaluation capacity by developing an affordable longitudinal evaluation model to track immediate and long-term outcomes achieved by program participants and continuously improve program quality. The Maui Farm is a innovative farm-based residential program for youth from troubled families. The Maui Farm evaluation model integrates the Hawaiian value of "ohana" with internal/external evaluation approaches and quantitative/qualitative evaluation methods to permit self-assessment by program participants, guide the day-to-day work of service providers, and meet the demands for accountability by funders and the community. The poster shows how trained community volunteers can help to collect data that promotes objectivity and cultural sensitivity, while involving and informing the community about the outcomes achieved.

Roundtable B (Second Rotation): Assessing Parks as Classrooms: Promoting Program Evaluation Capacity in the National Park Service, Brenda Jochums Slez, University of Massachusetts at Lowell and Elizabeth Hoermann, Lowell National Historical Park

The paper will address the development and implementation of a formative program evaluation model for education programs in national parks. The need for internal program evaluation was recognized by the National Park Service (NPS), especially as staff developed curriculum-based programs to serve the needs of public and private schools throughout the 50 states. The model consists of a book and training course for mid-level managers and/or education specialists. The model provides both a theoretical background and structure for conducting and reporting evaluations. The training course provides exercises and an opportunity to conduct and report an evaluation with feedback from nstructors. The paper reports information on the use of the model over a five year period and discussion of the support necessary to continued use and evelopment of expertise among local park staff.

Roundtable C (Second Rotation): Capacity or Confusion? Brenda Wong & Duane Okamoto, Alu Like Inc

What role can evaluation play in coordinating the delivery of diverse and often overlapping programs? Join internal evaluators for a dialogue focusing on the interaction between assessment, planning, accountability, and the practical implementation of programs within the Native Hawaiian community.

Roundtable D (Second Rotation): Choosing Change: A Self-assessment Manual for Non-profit Organizations, G Kem Lowry, University of Hawaii; Holly Henderson, Integrated Outcomes; and Karen Masaki, Hawaii Community Foundation

In the early 1990s the Hawaii Community Foundation began making 'capacity-building' grants to non-profit arts organizations. Early evaluations of the grantees revealed both an aversion to evaluation and some ignorance of legal requirements of non-profit organizations and management 'best practices'.

Foundation staff and consultants developed a self-assessment manual to encourage staff and boards to review their legal obligations and management practices. This manual, Choosing Change, has since been used by a wide variety of arts and other non-profit organizations. The poster session offers an opportunity to show the manual and to share experiences with others who are encouraging the use self-assessment processes.

Roundtable E (Second Rotation): Community-based Evaluation Teams for Capacity Building, Bennie M Stovall, Neighborhood Information and Sharing Exchange

This discussion will focus on the processes and methods used to evaluate a grassroots community leadership development project. The purposes of the project included the training and preparation of residents (non-traditional leaders) for leadership roles, and their subsequent efforts toward changing community systems. The outcomes measured entailed the impact of the training curriculum to develop leaders, and capacity changes in the community. The primary strategy used was the creation of a community-based evaluation team. The team was composed of community members (including leaders) and the evaluator. A development-in-the-context-evaluation (DICE) method was also used. That is, through a process of mutual learning-doing and decision making, the evaluator and team developed a partnership for implementing the project and its evaluation. The lessons learned provide details about leadership development as a part of community capacity building in general, and evaluation capacity building using teams.

SESSION 456: MultiPaper **Room:** Akaka

Sponsored by the Topical Interest Group on Crime and Justice

Conceptual Models of Juvenile and Criminal Justice and Uses in Local Capacity Building

Chair: Robert A Kirchner, US Department of Justice

Presenters: *Capacity Building for Juvenile Gun Violence Reduction Partnerships*, David I Sheppard, COSMOS Corporation

In 1997, the Department of Justice's Office of Juvenile Justice and Delinquency Prevention awarded three community demonstration grants to implement the 'Partnerships to Reduce Juvenile Gun Violence' program. The program's goals are to: 1) reduce illegal gun availability to juveniles; 2) reduce juveniles illegally carrying firearms and committing gun-related crimes; and 3) coordinate services for at-risk youth. The national evaluation team provided intensive capacity-building technical assistance to enhance the demonstration sites' partnership structures, and employed a logic model process to enhance their comprehensive strategies, including the identification of appropriate process and impact outcomes. This process allowed the programs to link the efforts of local government, law enforcement, probation, prosecutors, schools, clergy, and other youth-serving organizations. Initial assessments of this process suggest that all of the sites have developed into functioning partnerships, and identified important lessons for creating effective gun violence reduction partnerships across such a large spectrum of stakeholders.

Conceptual Issues in Conducting Evaluation within a Juvenile Justice Context, Ryan M Quist, Claremont Graduate University

In 1967, the Presidential Commission on Law Enforcement and Administration of Justice recommended to President Lyndon Johnson that the juvenile justice system should practice judicious nonintervention and avoid the incarceration of children and adolescents. The commission suggested that such an emphasis should strive for decriminalization, diversion, due process, and deinstitutionalization. As time passed, however, less and less emphasis is placed on rehabilitation. Juvenile justice programs have begun to emphasis safety and security, personal responsibility, and cost effectiveness. The present paper evaluates this change in policy focusing on the resulting implications for juvenile justice evaluation. The accumulated evaluation literature documenting effective juvenile justice interventions is reviewed. The characteristics of these effective programs are contrasted with existing policy. The issues discussed in this paper will prove useful to evaluators, program designers, administrators, and policy makers.

Evaluation and Criminal Justice Program Theory: The Critical Linkages, Robert A Kirchner, US Department of Justice and Roger K Przybylski, Kent Consulting Group

This paper presents criteria and methodology to enhance evaluation designs for criminal justice programs. To produce findings and results that are useful and directly applicable for decisionmaking is often difficult, but not impossible. If program evaluation is to be a fundamental part of effective public policy, the primary responsibility for program evaluation should rest with responsible officials. By integrating program evaluation and program administration, program managers must ensure that program evaluation functions: (1) demonstrate to the satisfaction of oversight officials, the extent to which the program is effectively administered, and (2) support the program manager in producing an effectively administered program. Program managers need better methods to guide program development and to demonstrate success. Policy makers and funding sources need results that answer their questions and identify what works and where to focus future resources. Criteria are recommended that define an effectively managed program, and a set of ideal conditions for program implementation and performance -- conditions to be brought about through linkages between management direction and program logic.

Strategies for Fostering Utilization of Information from an Outcome-based Information System Tracking Delinquent Youths, Lori K Grubstein, Crime and Justice Research Institute

ProDES (The Program Development and Evaluation System) is a unique outcome-based information system that tracks every adjudicated youth in the City of Philadelphia's juvenile justice system and measures delinquency program outcomes, regardless of the program's location. ProDESÆ primary goals which focus on the utilization of data collected by the system are as follows: (1) To facilitate the development of delinquency programs serving Philadelphia youths; (2) To facilitate the rational matching by probation officers and judges of adjudicated Philadelphia youths to programs that can meet the youths' needs and the community's needs; and (3) To facilitate the entire array of delinquency services provided by the City of Philadelphia's juvenile justice system. This paper will explore the extent to which these utilization-focused goals have been met. Barriers and obstacles encountered in attempting to foster use among intended users will be discussed. Finally, strategies for successfully overcoming these challenges will be presented.

SESSION 457: Business Meeting and Panel Presentation **Room:** Koko

Sponsored by the Topical Interest Group on Assessment in Higher Education

Evaluating Transformation in Higher Education Institutions

TIG Chair: Mark E Troy, Texas A&M University
Session Chr: William H Rickards, Alverno College

This session draws on the work of the Kellogg Forum for Higher Education Transformation. This consortium--

composed of five institutions, three higher education study centers, a national foundation, and a national association-has been meeting over the last five years to address issues in the transformation of institutions and systems toward more socially engaged and responsible educational practices. It has drawn on the extensive data bases of UCLA's Higher Education Research Institute as well as the experiences of individual campuses such as the University of Portland and Alverno College, all of whom are members. Through these processes, the questions of evaluation focus on how to capture the experience of institutions in the midst of change and how to make data and observations most effectively usable in these processes. This session will involve representatives from the different participants and explore different models for evaluation practice and how they support learning within a consortium.

Panelists: *Institutional Transformations: Leadership Perspectives and Evaluation Needs*, Andrea Rutledge, Olivet College
Studying Institutional Practices, Donald Tuski, Olivet College

This presentation lays out some of the problems of undertaking major reorganizations in educational practice and studying them at the same time. >From an anthropological perspective, many of these involve actively re-defining working roles of educators and negotiating relationships within practice.

Rethinking University Studies, Judy Patton, Portland State University

At Portland State, the University Studies Program was reorganized to increase its effective role in the undergraduates' experience and the in baccalaureate degree. Implementing the revisions has had implications at a number of levels that become important to the notion of evaluation.

The Consortium as a Strategy for Change: Learning Perspectives and Evaluation Practice, William H Rickards, Alverno College

For Alverno College, this Kellogg forum and other consortium experiences have confirmed the effectiveness of consortia as formats for learning. They provide structures within which to pursue change strategies as well as forums in which to explore the conceptual foundations and empirical implications of change, and they are likely to continue as attractive options for granting agencies. Two important questions, then, emerge regarding how inquiry and evaluation activities contribute to the learning of consortium participants and how that learning of participants is, in itself, evaluated. A set of learning principles—with evaluation applications—is then explored.

SESSION 458: Panel **Room:** Waimea

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

Outcomes Evaluation: Small Wins in an Era of Accountability

Chair: James H Derzon, George Washington University

Whether the effort comes from a desire to add to the knowledge base or the need to meet NGO, agency or federal accountability directives, quantitative outcomes research using comparison group designs remains the primary tool of evaluators for demonstrating intervention effectiveness. Thirty years of evaluation experience, however, has shown how difficult it can be for single studies to demonstrate that effectiveness. When coupled with small numbers of subjects and a p. = .05 criterion, this difficulty was interpreted as "Nothing Works." Because evaluation occurs in a political context, the problem is more than academic. This panel provides a forum for discussing some of the latest thinking on the causes, consequences, presentation, management of likely small effects in an era of accountability.

Panelists: *On the Probability of Small Wins: A Meta-analysis of the Stability of Substance Using and Antisocial Behaviors*, James H Derzon, George Washington University

Prevention practice is predicated on the theory that altering risk factors or their relationship to a distal outcome will reduce the likelihood that a those exposed to intervention will engage in that undesired behavior. For this to be an effective strategy, however, several conditions must be met. First, the risk factors targeted for preventive intervention must, indeed, be predictive of subsequent delinquency. Second, that predictive relationship must involve a causal component rather than a spurious association. Third, the subjects exposed to intervention must be those likely to engage in the undesired outcome. Fourth, the intervention should have sufficient impact on the outcome to be worthwhile. Using evidence from a recent risk-factor meta-analysis of over 300 reports of 155 prospective studies examining the stability and homogeneous continuity of substance using and antisocial behaviors, we will examine the predictive strength of these behaviors and their utility for selecting individuals for preventive intervention.

Statistical Significance and Practical Significance in Outcome Evaluation, Mark W Lipsey, Vanderbilt University

Assuming a research design that provides a credible estimate of program effects, how can we assess the practical magnitude of those effects? Statistical significance, proportion of variance represented, effect size statistics, and other such commonly used statistical frameworks bear no necessary relationship to practical significance. To assess practical significance, statistical information must be translated or interpreted into terms relevant to the practical domain of application. This presentation will examine the relative advantages of various ways of doing that using examples from meta-analysis of outcome evaluations

Small Effects and Statistical Significance: Practical Advice, Dale E Berger, Claremont Graduate University

Null hypothesis tests of statistical significance have come under sharp criticism in recent years. We will examine these criticisms and explore alternatives. A key weakness is that hypothesis testing introduces a false dichotomy between a 'real effect' and 'no effect.' The false dichotomy can lead to a search for moderator variables to reconcile apparent differences which may have no substance. This concern is especially relevant in research where the effects in the population are small and power of the statistical tests are low. A practical alternative is to describe the estimated effect size and the precision of that estimate in comparison to both a null effect and to a meaningful effect specified by the researcher.

Discussant: Leonard Bickman, Vanderbilt University

SESSION 459: Demonstration and MultiPaper **Room:** Niihau

Sponsored by the Topical Interest Group on Extension Education Evaluation

Accountability: Demonstration and Presentation on Aspects of Development

Chair: Marilyn Grantham, University of Minnesota

Whether the effort comes from a desire to add to the knowledge base or the need to meet NGO, agency or federal accountability directives, quantitative outcomes research using comparison group designs remains the primary tool of evaluators for demonstrating intervention effectiveness. Thirty years of evaluation experience, however, has shown how difficult it can be for single studies to demonstrate that effectiveness. When coupled with small numbers of subjects and a p. = .05 criterion, this difficulty was interpreted as "Nothing Works." Because evaluation occurs in a political context, the problem is more than academic. This panel provides a forum for discussing some of the latest thinking on the causes, consequences, presentation, management of likely small effects in an era of accountability.

Presenters: *Minnesota Impacts: Using a Web-based Accountability and Reporting Tool To Increase Evaluation Capacity in a State Extension Service*, Jennifer Obst & Marilyn Grantham, University of Minnesota

Minnesota Impacts is a web-based, publicly accessible, searchable database that reports the economic, environmental, and/or quality of life impacts and the linkages between research and Extension outreach at the University of Minnesota. It is a project of both the Minnesota Agricultural Experiment Station and the University of Minnesota Extension Service. MI has increased evaluation capacity among faculty and staff because of the need to provide outcome and impact information. Each MI submission is carefully reviewed by an economist familiar with impact and cost-benefit analysis and a professional writer-editor. This review helps authors to recognize the kinds of information that they are going to need to document impacts and therefore, what kind of evaluation they may need to conduct. This demonstration will discuss what has been learned from the experience of developing Minnesota Impacts and the strengths and weaknesses of a web-based reporting system and show how MI is used.

Developing a Web-based System for Reporting Extension Accomplishments, Rama Radhakrishna & Joan Pinion, Clemson University

(Basic Message)
With major accountability and reporting issues and major changes in the computer support system, program evaluation faculty and staff were faced with a challenging task of developing a new reporting system that essentially captures all of the Extension accomplishments emanating from programming efforts in all counties. The basic message of this presentation will be to: 1) describe the process involved in developing a Web-based system, 2) assess the reactions of users toward the new system and 3) share examples of reports that were generated from the new system to meet state and federal accountability requirements.

(Implications for Evaluation Practice)
This presentation will demonstrate through a four-step process of what it takes to develop a good Web-based reporting system. How important input from county staff and campus faculty is to the development of the system will also be discussed. Involving county staff in the development process and visits to county offices has helped them to understand the importance of accountability and reporting and the use of accountability information for decision making. In addition, involving them in the process has eliminated concerns they had about evaluation, accountability, and reporting. As one agent reported, "communicating with county staff and involving them in the development process has brought credibility and confidence to the reporting system."

(Usefulness of Information to Audience Members)
Our presentation will answer the question what did you learn from what you did and how have you used what you learned to make the system better? Audience members will learn how a team approach is critical to developing a Web-based extension reporting system. The steps followed in developing the Web-based system serves as an excellent model for other organizations to follow in solving accountability and reporting problems. Audience members will also know how user reactions was helpful in refining the system and making the system more user-friendly. Finally, the results will show members how state and federal reports were generated using the new system.

(Other Information)
Good accountability reporting systems are critical to extension program effectiveness. In that it contributes to better programs, better relationships with users, clients, and better working relationships with state and federal agencies. Recently, the South Carolina Budget Control Board which oversees accountability aspects of all public agencies in South Carolina reviewed our new system and recommended that this system can be a model for 57 other state agencies. A summary highlighting the steps involved in developing the new system was published in the November issue of their newsletter, Impact.

Measures of Success: Measuring the Impacts of the Canadian 4-H Program on Members, *Family and Alumni*, John B Collins, University of British Columbia

Canada's 4-H program is a powerful and enduring force for rural youth, families and communities in fostering core objectives such as knowledge and skills, leadership, personal development, and citizenship skills. Additional initiatives in rural leadership, farm safety, environmental awareness, technology transfer, awareness of agricultural production, processing and marketing, and careers in agriculture have been focus topics during the 1990s. Measures of Success was a three-part evaluation to examine how well the overall objectives of the Program are achieved by (1) surveying alumni who were 4-H members from the 1950s through the 1980s, (2) nation-wide telephone interviews with current community and agri-business leaders, 4-H leaders and volunteers, 4-H families and current members, and (3) a package of audit and evaluation tools for reviewing 4-H projects, programs and activities to determine how well they promote the Program's overall goals. More than 1,400 people responded during the cross-Canada information-gathering phase. During the paper presentation, overheads and handouts will document the evaluation outcomes, with a particular focus on how provincial evaluations have now become routinized in 4-H program operations.

SESSION 460: Panel **Room:** Hilo

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

A Comprehensive Strategy for Implementation, Evaluation, and Utilization from Diversity Grantmaking at a Major Foundation.

Chair: Astrid Hendricks-Smith, W K Kellogg Foundation

For the W.K. Kellogg Foundation (WKKF), 'capitalizing on diversity' is more than a philanthropic catch-phrase. Throughout its history, the Foundation has taken special care to promote diversity across all our program areas. The Foundation programs in Latin America and Africa. By identifying Capitalizing on Diversity (COD) as a cross-cutting theme in 1996, this long-standing commitment was formalized. Currently, the WKKF is in a learning and re-envisioning phase as it works toward integrated programming across cross-cutting themes. This phase focuses on measuring progress and culling lessons from the Foundation's diversity-related grantmaking and practices, and the larger organizational context in which this grantmaking occurs. The following panel will describe the context for this work and three very exciting projects: (1) a framework for a WKKF 'learning plan' focused on integration of diversity principles; (2) a new WKKF program called YES Program that will serve as a mechanism for the integration of learning principles, and (3) the ensuing evaluation of COD-YES.

Panelists: *A Framework for a Learning Plan for Diversity Grantmaking at the W.K. Kellogg Foundation*, Hanh Cao Yu & Traci Endo, Social Policy Research Associates

The exploration of how to measure integration of diversity principles within organizations breaks new ground within the evaluation field. Social Policy Research Associates will present a dual-level framework developed for for evaluating diversity grantmaking within the W.K.Kellogg Foundation. On one level, the framework articulates a dynamic model by which Foundation grantmaking themes of access and representation are developed to sustain individuals and communities against poverty. Within this model, ethnic and racial culture and communities are engaged as resources to create paths towards self-determination. The framework then further addresses institutional supports and barriers to sustainable adoption of this diversity grantmaking strategy. While developed to evaluate the integration of COD within WKKF, these supports of shared vision, authorizing environment, and organizational capacity have been adapted to evaluate organizational readiness and progress in addressing diversity-related strategies more broadly. Key elements articulated under this framework served as an evaluation tool for analyzing grantees for another WKKF initiative, COD-YES.

Implementation of Diversity Principles in Programming at the W.K. Kellogg Foundation: The Youth Engagement Strategy Program, Winnie Hernandez-Gallegos, W K Kellogg Foundation

The goals of the Youth Engagement Strategy Program are to empower youth to address racism and cultural issues in their communities and to enhance intergenerational relationships as youth and adults engage in community problem solving around racism. The program supports leadership development and community projects to ensure that all participants strengthen interpersonal skills necessary to initiate and sustain positive race/cultural relations (i.e. learning to deal with difference), acquire practical skills on running a program or activity (i.e. planning, decision making, consensus building, negotiating, listening, holding people accountable), assume responsibility for their own learning and development, and strengthen their own personal identity. Projects are expected to create or strengthen safe forums in communities for youth of all racial, ethnic or cultural backgrounds to take positive action on race and cultural. Finally, the program provides a vehicle for exploring conflicting messages and stereotypes about race and culture.

An Evaluation Strategy for W.K. Kellogg's Youth Engagement Strategy Program, Hanh Cao Yu & Traci Endo, Social Policy Research Associates

Dr. Hanh Cao Yu and Ms. Traci Endo of Social Policy Research Associates will present an evaluation framework for examining COD-YES, or WKKF's initiative to support youth engagement against racism. The framework acts as a lens for evaluating the process of WKKF's involvement with YES, organizational and programmatic practices across different grantees, and youth outcomes. At the foundation level, the framework examines efforts by WKKF to use this initiative to help build broad institutional support for proactive diversity grantmaking. At the organizational and programmatic level, the framework examines how to evaluate and, through technical assistance, support grantees' use of effective practices in youth development and prejudice reduction. At the youth level, the framework examines whether youth participants are developing the social, leadership and multicultural competencies that the literature on racism suggests are essential to an engaged anti-racist identity.

SESSION 461: Panel **Room:** Puna

Sponsored by the Topical Interest Group on Minority Issues in Evaluation

Education, Health and Evaluation Capacity: Re-thinking Evaluation Models in Diverse Latino Communities

Chair: Emma V Sanchez, San Francisco State University

As the global economy creates waves of immigration from Latin America into the U.S., dynamic Latino communities are established in urban and rural America. Working with these evolving communities demands multicultural perspectives and dynamic models of evaluation that provide a lens of understanding the diverse countries and experiences that these communities offer to the evaluation. Evaluation capacity requires active promotion and development of community members working as colleagues within these communities. This panel will present findings and lessons learned from evaluations in the areas of public health, education and community organizing. The panelists will discuss their levels of insertion in diverse Latino communities in California and El Salvador, as well as the opportunities and complexities involved in the emic and edic perspectives in the evaluation process. Implications for evaluation capacity will be proposed.

Panelists: *Evaluation Capacity Issues in Diverse Latino Communities: Employing an Equity Approach*, Emma V Sanchez, San Francisco State University

As the U.S. Latino population grows at accelerated rates, public health practitioners and educators are increasingly challenged to effectively respond to the cultural and linguistic needs of this population. Evaluation capacity with Latino communities requires efforts that must transcend translation services. While many evaluations have made contributions to the field, little has been discussed about the challenges and opportunities encountered in evaluation capacity efforts involving Latino communities. This paper explores the challenges encountered in evaluating service programs and community building initiatives with California multi-ethnic, Latino communities; analyzes the need for transcending translation, and the realities involved in building evaluation capacity with these communities; suggests that evaluation vocabulary is grounded in a cultural and political context; and proposes equity approaches to begin to overcome some of these challenges.

The Evaluation Triad: Involving Diverse Latino Perspectives in a School-based Family Literacy Program, Betty Pazmiño, Cesar Chavez Elementary School

A fairly recent area of interest, the notion of family literacy is attributed to educational research by Denny Taylor and others who have done investigations of families and their relationship to school literacies. School-based family literacy is complex due to the multitude of needs that are represented by the diverse families within public schools in inner-city urban environments. Based on qualitative research evaluation involving Latino populations in a multicultural school, this paper will discuss:

* findings and analysis of the triad relationship that exists within a school-based family literacy program involving teachers, parents and school children coming from different immigration backgrounds and countries.
* Challenges involved in language translation and cultural interpretations of the parent perspectives within an educational setting.
* Contradictions involved in the paradoxical relationships between monolingual Spanish parents and their school environment.

Implications for evaluation capacity and school-based Latino parent leadership formation will be presented.

Lessons Learned from Evaluations with Latino Communities: An Edic Perspective, Alice Linsmeier, E R Taylor Elementary School

The debates on bilingual education and cultural competency have helped bring to the surface many important issues related to Latino communities in the United States. It is essential that those working with Latino communities go beyond the limitations of these debates, moving toward building models that are founded on equitable structures starting from within the Latino communities. Thus, evaluation capacity requires a thorough understanding of the challenges and opportunities of building relationships with Latino communities. This paper will discuss findings from evaluations conducted in El Salvador and the U.S., including the lessons learned in:
* establishing "the community" as the protagonist of the evaluation,
* promoting community leaders in the evaluation design, interpretation of findings, creation of programs and a structure for continual evaluation.
The opportunities and challenges involved in these evaluations will be offered from an edic perspective.

SESSION 462: Panel **Room:** Kohala

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

Evaluating Statewide Initiatives that Use County/Community Partnerships To Achieve Important Health and Educational Goals

Chair: Abraham H Wandersman, University of South Carolina

Several states have implemented statewide early childhood initiatives through partnerships (public/private; county/community) that work locally to improve services and conditions for children birth through 5 and their families. The long-term goal of many of these efforts is that children come to school healthy and prepared to succeed. However, the ways in which partnerships are allowed to address this goal are quite diverse. States have acknowledged that, with flexibility, local partnerships can be more responsive to their local needs and more likely to design and deliver interventions that are effective. Evaluation of these initiatives is a challenge. Some lack well-defined goals or target so many goals as to be unrealistic. The services and interventions delivered by these initiatives range from intense to relatively weak, yet strong outcomes are expected by all. Disentangling the effects of the "new" initiative from existing systems of care is difficult. The evaluators of 3 state initiatives -- South Carolina's First Steps, North Carolina's Smart Start, and California's Proposition 13 -- will describe their efforts to meet the evaluation challenges posed by their states. Presenters will include the conceptual model behind the initiative, the design of their evaluations, and strategies that have proven successful in meeting the challenges of these multi-dimensional partnership initiatives.

Panelists: *Evaluating North Carolina's Smart Start Initiative*, Donna M Bryant & Kathleen Y Bernier, University of North Carolina at Chapel Hill

Smart Start began in 1993-94 in 18 of NC's 100 counties. Through a competitive process, new counties have been funded each year so that now all 100 counties receive fund to provide new, improved, or expanded services to children under 6 and their families. The budget for FY99-00 is $254 million. Short and longer-term results derived from the FPG's multi-component longitudinal statewide evaluation have played a major role in providing state legislators with "evidence" to support expansion of the Smart Start program. The NC presentation will include an overview of the variety of qualitative and quantitative studies conducted as apart of the statewide evaluation, with a focus on those that have seemed the most influential to policymakers. The role of the statewide evaluators in helping local partnerships build capacity for local evaluation so that local results can influence decision-making will be briefly described.

Measuring Results: California's Children and Families Act, Jane Henderson, California Children and Families Commission

In November, 1998, California voters passed Proposition 10, the California Children and Families Act. Funded by a surtax on cigarettes and tobacco products, the Act generates approximately $700 million annually to support services for children and their families prenatally to age five. The Act is based on a results based accountability model. It is designed to help communities achieve results in three broad areas: Improved Family Functioning/Strong Families, Improved Health/Healthy Children, and Improved Child Development/Children Learning and Ready for School. As required by the Act, the State Children and Families Commission has identified the results to be achieved by the Act, including short term and long range outcomes. The Commission has also identified multiple sets of indicators which can be used to gauge results locally and statewide. The California presentation will include an overview of issues that need to be Considered in evaluating multifaceted early childhood initiatives -- where the traditional tools of evaluation have limited application; an account of what has been learned from the California effort; and a brief consideration of what is needed from the national, state, and local perspective to support this type of evaluation effort.

Evaluating the Grant Process in First Steps: The South Carolina Initiative, Matthew Chinman & Joy Kaufman, Yale University; Abraham H Wandersman, Pamela S Imm, Arlene B Andrews & Baron Holmes, University of South Carolina

Qualitative and quantitative techniques are being used to evaluate the grantmaking and grant implementation process. Components of the process include grantwriting, grant reviewing, grant implementation and training and technical assistance. Linkages between internal and external evaluation will also be described.

Discussant: David Fetterman, Stanford University

SESSION 463: MultiPaper **Room:** Kona

Sponsored by the Topical Interest Group on Environmental Program Evaluation

Evaluating Environmental Policies and Programs

Chair: Claude F Bennett, US Department of Agriculture

Presenters: *Management and Accountability Indicators for Intergovernmental Programs: The National Extension Targeted Water Quality Program*, Claude F Bennett, US Department of Agriculture; Mary G Marshall, Texas A&M University; James C Davis, Anglican Church of Canada; Julie Jones, Inver Hills Community College

Intergovernmental programs initiated by the federal government, i.e., those to be implemented through projects conducted by state government units, must provide for order across their component projects as well as freedom for individual projects to appropriately respond to their respective, varied situations. The paper presentation describes indicators for the U.S. Department of Agriculture's 53 state/territory National Extension Targeted Water Quality Program. The indicators identified commonalties across state projects as well as variations, or diversity, among these projects. Indicators characterized university extension projects (e.g., educational efforts to encourage improved management of dairy cattles wastes) relative to project outputs, outcome targets and associated outcomes. Illustrations of the types of information produced by the indicators will be presented. The paper evaluates the indicators for the Targeted Program as well as USDA'sapproach to developing them. Finally, suggestions for future development of indicators for intergovernmental programs are presented.

Evaluating Environmental Policy Instruments, Per A Mickwitz,, Finnish Environment Institute

The paper departures from some key characteristics of environmental problems and the policy instruments that can be used to tackle these problems. Environmental problems often: involve public goods and externalities; have long time frames; are complex; involve huge uncertainties; and concern geographically remote regions. The main part of the paper analyses the implications of these features for different aspects of the evaluation of environmental policy instruments. These include: evaluation methods; criteria; and the impact problem. Throughout the paper an evaluation of the Finnish carbon dioxide tax introduced in 1990 is used as an example. If evaluations of environmental policy are undertaken without due considerations of the specifics involved the risk is large that little impacts and low effectiveness will be found. Properly undertaken evaluations can, however, help us make better decisions on old, as well as, new environmental problems.

Evaluating a Market Based Energy Program, Jane S Peters & Marjorie R McRae, Research Into Action; and Ben Bronfman, Northwest Energy Efficiency Alliance

The Architecture + Energy (A+E) program, offered by the American Institute of Architects Portland, encourages the market to produce more energy efficient buildings. The program began awarding good design in 1993 when utilities offered incentives to improve building design. In 1995, most utilities reduced their funding and shifted to trying to influence the market to support energy efficient design. Since 1997, the Northwest Energy Efficiency Alliance has funded the A+E program as a market transformation program. Throughout A+E's history, participants expressed high levels of satisfaction with the program. However, a market transformation program must influence more than participants, there must be market effects. Evaluating market effects is challenging. One approach is to use a market assessment as the baseline and to provide an ongoing means to track whether market effects have occurred. This paper will discuss how to market based programs can be evaluated using the A+E program.
.

Evaluation, Effectiveness and Global Environmental Regimes, Mikael Ro̊màn, Uppsala University and Evert Vedung, Institute for Housing Studies

The paper concerns the operation of global environmental regulatory regimes. More specifically, it focuses on how their effectiveness might be evaluated. As such, the paper constitutes thereby an attempt to cross-fertilize environmental regime theory, as conceptualized in the International Relation literature, and evaluation research. The authors claim that effectiveness can be assessed in various stages in the policy process. Consequently, they discuss diverse conceptions of effectiveness, such as impact, outcome, output and process effectiveness. In this effort, the 1985 Vienna Convention and the 1987 Montreal Protocol for the Protection of the Stratospheric Ozone Layer is taken as a case illustrating the general argument.

SESSION 464: Alternative Format **Room:** Ewa

Sponsored by the Topical Interest Groups on Health Evaluation & Minority Issues in Evaluation

The Native Hawaiian Health Program: A Dialogue on Evaluation Issues

Chair: Barbara Wells, Health Resources and Services Administration
Presenters: Meiling Chang, Hui No Ke Ola Pono
Sita Nissanka, Kamehameha Schools Bernice Pauahi Bishop Estate
Hardy Spoehr, Papa Ola Lokahi
JoAnn Tsark, Papa Ola Lokahi

The mission of the Native Hawaiian Health Care Program (with partial support from HRSA/BPHC) is to improve the physical, mental, and spiritual health of Native Hawaiians by advocating, initiating and maintaining culturally appropriate strategic actions. This session will begin through providing background presentations on two of the program's major components: the Native Hawaiian Health Care Systems (NHHCS) and the Native Hawaiian Health Scholarship Program (NHHSP). Presentations will then address the approaches used and issues faced in evaluating each of these components. For NHHCS, we will address the collection of uniform data on clinical process and outcome measures and the need for capacity building. For NHHSP, we will present the methodologies being planned for assessing the experiences and outcomes of Native Hawaiian health scholars. Finally, we will engage the attendees in an extensive dialogue about the numerous issues identified in the presentations, including participatory evaluation and culturally appropriate methodologies.

How to:
The reason that we are characterizing this session as an "alternative format" is that we expect to have spend at least half of the session engaging the participants in an open dialogue on the unique aspects of evaluation within the Native Hawaiian community. We contrast this with the traditional panel with didactic presentations and Q&A's. The 90 minute session will flow as follows. The presentations will be timed to last no longer than 45 minutes; this will be possible because the presentations will only frame the issues for later discussion. For the remaining 45 minutes of the session, we intend to have an open dialogue between the Native Hawaiian presenters and the audience around the evaluation issues identified in the presentations. These issues include the unique needs of the Native Hawaiian population, empowerment, and evaluation capacity building to name a few. Ultimately, we will address: What does one need to understand in order to participate in evaluation within the Native Hawaiian community?

SESSION 465: Panel **Room:** Honolulu

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

Maximizing the Capacity of Multi-Site Evaluations -- II: Validity Strategies used to Enhance the Quality of the CMHS Housing Initiative for Persons with Serious Mental Illness

(Part I to this session is to be held Friday at 9:30 in the Honolulu Ballroom.)

Chair: Debra J Rog, Vanderbilt University

Although multi-site studies offer unique methodological challenges and validity concerns, they also offer unique strengths. In the CMHS Housing Initiative, in particular, its design offered several opportunities for assessing the validity of the measures of changes in the target population and the fidelity of the interventions. The features of the Initiative that lend themselves to different validity approaches include a diverse sites in different regions of the country with varying target populations, cross-site adoption of a uniform set of resident and housing measures, site specific studies that used complimentary but different measurement strategies on some key measurement domains, and the involvement of an ongoing consumer perspective at both the individual and cross-site levels. Each of these features will be highlighted by one of the four presenters in the panel. The chair will introduce the session, offering a transition from the earlier panel, and also serve as the discussant.

Panelists: *A Cross-site Approach to Assessing Construct Validity: The CMHS Housing Initiative Validity Synthesis*, Michael S Shafer, University of Arizona

Sparked by interviewer concerns and preliminary baseline results, participants in the CMHS Housing Initiative implemented a cross-site validity process to provide a critical assessment of the core resident interview used across the six sites. In each site, the data coordinator conducted focus groups with their interviewers and consumer representatives to review the protocol item by item with respect to vocabulary problems (i.e., words that are unfamiliar to residents, obscure in meaning, etc.), situation dependent questions, ambiguous meaning, circumstantial interference (external factors that affect its completion), and questions that may not apply at the moment or may reflect changing attitudes over time. The validity issues were summarized cross-site and used to inform an analysis plan. This presentation will describe the validity process, the key concerns that emerged, and how they affected the cross-site analyses.

Assessing the Validity of the Cross-site Findings: Insights from Site-specific Ethnographic Work, Kim Hopper, Nathan S Kline Institute

In addition to having some set of uniform measures across diverse sites, one of the benefits of a multi-site study is having site-specific studies that can offer additional insights into the cross-site findings. In the Housing Initiative, several sites conducted studies that can be used to assess the validity of some of the cross-site results. In one study, ethnographies conducted in selected housing programs has illuminated some of the limitations of the standardized protocols used to assess resident outcomes as well as to assess the operation of the housing programs. This presentation will share some of the major insights offered through this work, including how choice in services is negotiated, how drug and alcohol issues are handled, and the impact of the local environment and neighborhood on outcomes.

Assessing the Validity of Cross-site Findings: Insights from Site-Specific Studies of Housing and Fidelity, Linda Frisman, Connecticut Department of Mental Health and Addiction Services

The work of the Connecticut Housing Initiative Study provides a second example of how a site-specific study can offer insights into cross-site findings. The presenter will describe two aspects of their site specific evaluation that helps assess the validity of the cross-site housing fidelity and implementation assessment (described in Panel I). First, the project conducted a separate cross-sectional fidelity assessment of each program with all residents currently living in the program. Because the cross-site fidelity approach relies on key informants in each housing program, these data provide a basis for assessing the validity of that approach. Second, Connecticut is conducting a study of how the social environment of supported housing relates to individual outcomes. These data will also shed light on the validity of the cross-site fidelity approach as well as offer additional insights into the dimensions measured.

The Role of Consumers in Improving the Validity of Individual and Multi-site Efforts, Elaine Myers, Portland State University

A key feature of the Housing Initiative has been the ongoing, real involvement of consumers (in the form of a Consumer Panel) in every aspect of the multi-site study, from its initial design to the analysis and interpretation of results. In this presentation, a consumer representative will describe the role the Consumer s had in enhancing the validity of the measurement, including their role in selecting appropriate instrumentation, highlighting and addressing vocabulary concerns, and increasing the sensitivity of interviewers to consumer issues by participating in the cross-site interviewer training. The presentation will also address how the Panel is involved in assessing the validity of the collected data through their participation in the validity synthesis as well as their work in assisting in the interpretation of baseline and interim outcome results.

Discussant: Debra J Rog, Vanderbilt University

SESSION 466: Roundtables **Room:** Kahuku

Evaluation Issues in International and Domestic Federal, State and Local Government

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

Host: Martin Ørvim, Office of the Auditor General of Norway

Roundtable A (First Rotation): Evaluating the Government Performance and Results Act

Presenters: *Benchmarking Federal Agency Progress on the Government Performance and Results Act (GPRA)*, Doris M Werwie, US Department of Education

March 31, 2000 marked the first time that all Federal agencies reported to Congress on the Government Performance and Results Act (GPRA). Each Federal agency had to demonstrate the actual progress made toward accomplishing established goals. After several years of planning; developing goals, identifying indicators and data sources, agencies were required to report on the quantifiable changes on each performance measure. This paper presents the results obtained by a benchmarking team comprised of key human resource staff from ten federal agencies. The team began by identifying benchmarking criteria and selecting five agencies to review. Benchmarking information and data were collected on the process each agency used to move GPRA gaols and objectives to specific strategies, which were then implemented by individual programs. The team studied the relationship of GPRA planning to budget and the alignment of agencywide goals and objectives with program goals. Data quality issues were also examined. The results provide a realistic look at the current status of strategic planning within the Federal government. It also identifies some best practices which will be helpful for across agency improvement and implementation of the Results Act objectives.New Foundations for Evaluation: The GPRA Evaluation Environment in the US Department of Education, Steven W Zwillinger, United States Department of Education and Natalia Pane, American Institutes for Research

New Foundations for Evaluation: The GPRA Evaluation Environment in the US Department of Education, Steven W Zwillinger, United States Department of Education and Natalia Pane, American Institutes for Research

The Government and Performance Results Act recognizes the complementary nature of program evaluation and performance measurement. Both are important components of an effective performance measurement system. Before the enactment of GPRA, the US Department of Education was conducting substantial numbers of program evaluation studies and made these activities part of a comprehensive strategy to capture performance data. We will present the role of evaluation within the department as it is evolving to serve five major purposes in the post-GPRA

environment:
1. Provide information beyond performance measures.
2. Validate performance data and refine performance indicators.
3. Address strategic, not programmatic, goals.
4. Guide program improvement.
5. Capture effects over the long-term.

We will also discuss strategies that the department has developed to advance the use of evaluation data. These strategies fall under the categories of supporting and improving performance measurement, using technology to improve response time, and improving capacity building.

Roundtable B (First Rotation): Issues in Assessing International Civil and Public Service Reforms

Presenters: *Qualitative and Quantitative Criteria for Assessment of International Civil Service Reform Programs*, Sutherland I Miller, The Services Group

Throughout the world governments are attempting to reform public agencies to become more cost effective and better serve the needs of constituents. Usually, reform efforts are designed to achieve such goals as boosting efficiency, cutting costs, increasing transparency in decision-making, and improving the use of information. Yet when outside actors, including donor agencies and business and community groups, try to assess the relative failure or success of reform efforts clear and consistent evaluation criteria are often elusive. Some criteria -- such as the processing time an agency takes to approve of an application and the existence of guidelines for completing a process -- can be tracked over time. But the more abstract and qualitative indicators that would validate progress in changing perspective of civil servants, improving customer service, and increasing responsiveness are often more difficult to demonstrate. This roundtable proposal suggests a discussion of ways in which civil service reform can be assessed internationally. Through the presentation of three case examples from Malawi, Jordan, and Tanzania various quantitative measurements will be presented and discussed. Participants will also be asked to share experiences and perspectives related to how to gauge and present qualitative measurements to demonstrate the success or failure in fomenting organizational change.

Creating a Competitive Market: What Can and Should be Evaluated? Amanda M Wolf, Victoria University of Wellington and David R Turner, New Zealand Department of Labour

Most public policies require ongoing, yet discrete, government involvement. Evaluating day-to-day operations and results can aid decisions. Policies that create competitive markets, however, pose different evaluation challenges. Regulated markets are complex open systems. It is difficult, perhaps paradoxical, to link government policy and activities to results. How can evaluators avoid over-promising what they can deliver while effectively targeting what can and should be assessed? In 1999, the New Zealand government shifted from monopoly provider of workplace accident insurance to light-handed regulator. The policy's implementation logic centers on communicating rights and responsibilities and facilitating bargaining among market players. Together with the policy's objectives to reduce injury rates and spread the premium burden more fairly, the logic suggests some premises and objectives for an evaluation strategy. These are compared with the actual strategy. The conclusions bear generally on evaluators' capacity to assess market-oriented policies.

Roundtable C (First Rotation): Mixed-method Approaches to Governmental Evaluations

Presenters: *Building Capacity for Intersectoral Program Evaluation: A Saskatchewan Example*, Lori S Ebbesen, Meredith Moore & Joan Feather, University of Saskatchewan; Felecia R Watson & April Barry, Saskatchewan Health; Bill Werry, ADMs Forum on Human Services; and Susan Hetu, Saskatchewan Social Services

Although working intersectorally in human services is supported conceptually and structurally in the province of Saskatchewan, Canada, challenges are evident in evaluating intersectoral programs. To understand the realities of and to build capacity for intersectoral program evaluation, a partnership planning group undertook a 2-staged process. First, a consultative process was facilitated with five stakeholder groups. Five key issues emerged from these consultations: finding common ground, de-mystifying evaluation, blending evaluation across multiple levels, establishing causality, and costs and benefits. Second, a satellite telecast workshop was held to explore and address the five key evaluation issues using a complex, multi-level Saskatchewan program as a case example. The consultative process and the satellite workshop reached over 700 individuals from across the province, and contributed to increased capacity in evaluating intersectoral programs. Participants were able to collectively address rewards and challenges of intersectoral programs, focus on common problems and solutions, and explore evaluation designs at multiple levels.

Evaluation of the Alliance for the Field (Alianza para el Campo) in Mexico, Murari Suvedi, Michigan State University; and Pablo-Ramirez Moreno, Horacio Santoyo Cortes & Jose Maria Salas Gonzalez, Universidad de Chapingo

The Alliance for the Field is the Federal Government of Mexico's policy to promote agricultural and rural development. It is designed to address the problem of low levels of agricultural productivity and technology adoption by farmers. During 1998, twenty agricultural development programs were implemented under "the Alliance for the Field" in the 32 states of Mexico. This evaluation was conducted to assess the impacts of the 1998 programs. The study utilized both qualitative and quantitative methods of data collection. An uniform methodology was used to gather information from 334 programs in all 32 states of the Republic. Data were collected to solicit the opinions about program impacts from farmers, research and extension officials, input suppliers and members of producer organizations. In this paper, we will summarize the methodological approach for this national evaluation. We will also share the major findings related to program performance, including client profiles and opinions, as well as perceived socioeconomic and environment impacts.

Roundtable D (First Rotation): Responding to Federally Mandated Audits

Presenters: *Evaluating the Performance of Legal Service Grantees*, C Eric Kirkland & Pamela M Schaal, Legal Services Corporation

In October 1999, Congress mandated an assessment of the accuracy of case service reporting by Legal Service Corporation's grantees, to be conducted by the Office of Inspector General. This paper presents the findings of this assessment and discusses the methodological hurdles necessitated by attorney-client privilege considerations. The cornerstone of the assessment is two-stage cluster sample of grantees, selected from the population of 237 grantees who provide the national legal aid service for the poor. Data were collected via an online submission process, to support case-level sampling, and by on-site visits to review case files. Findings were reported to Congress with respect the most common error types, including unsupported cases, no client name, untimely closing, no legal services provided, duplicate cases, no assessment of eligibility, and clerical error. Lessons learned will focus on the challenges of managing a multidisciplinary evaluation, integrating the work of auditors, lawyers, information technologists, and evaluators.

Managing for Results: A Methodological Design and Software for Assessing the Quality of Performance Reporting in the Canadian Federal Government, Werner J Müller-Clemm & Carol Motuz, Office of the Auditor General of Canada

As part of a government-wide Value-for-Money audit, the Office of the Auditor General of Canada (OAG) was interested in developing a robust methodology for assessing performance reporting in the Canadian federal government. Recently, a number of jurisdictions have been faced with the challenge of assessing public sector performance reporting. The OAG's identified need was to be able to assess both the quality and the progress made in performance reporting across government. We had to build upon an essentially judgmental process, by developing a rigorous methodological approach to evaluating performance reporting. The solution was to design an index or similar standardized measurement approach allowing assessments or judgements to be made in a clearly defined and systematic manner; resulting in more reliable and valid data collection and analysis. We describe the methodological design, the software prototype, preliminary findings, and ôlessons learnedö from the implementation of our rating methodology. We address the utility to stakeholders of the information produced. Finally, advantages and disadvantages of this approach will be discussed, as well as its applicability to other jurisdictions.

Roundtable E (First Rotation): Increasing the Voice of Evaluation in the Policy Process

Presenters: *Evaluating Social Science Research within the Context of the Government Performance and Results Act*, Irwin Feller, Pennsylvania State University

Federal agencies' required linkage of results to budget requests stemming from the Government Performance and Results Act (GPRA), combined with the National Science Foundation's recent articulation of 'connections between discoveries and their use in service to society' as a principal GPRA goal and proposal selection criterion, place new importance on the development of methods for evaluating the impacts of federal government sponsorship of academic research. This challenge is especially formidable in the case of social science research, given historical and contemporary political opposition to its funding and continuing differences among leading social scientists about the impact of social science research on public policy decision-making. This paper presents an analytical framework and plan for assessing the societal impacts of academic social science research. The framework links Weiss's conceptualization of "instrumental" and "enlightenment" forms of knowledge utilization, the economics of science and technological innovation, and selected case histories to demonstrate the cumulative impacts of multiple studies and interconnections between private and public sector uses.

Program Evaluation: Extracting Lessons from Central Debates in Order To Improve Policy Implication of Reports, Olav Andreas Kvitastein, Norwegian School of Economics and Business Administration

The article outlines central debates in program evaluation and extracts their implications for governmental evaluation management. That is, in a wide and fuzzy field of research, the article aims at helping manager navigate in an ocean of different schools and contradicting views of research and science. Such a point of departure does not necessarily produce an objective state-of-art overview. Rather, it provides judgements of schools and positions that protagonists of the various schools may object to. The descriptions of methods and tools are aimed at serving as product information for those who buy evaluation services. The core problem this approach is trying to address is usefulness i.e. to what extent is evaluation management carried out in a manner that facilitates learning and the correction of errors of the past? If the intentions behind evaluations are that reports should have policy implications, it is of vital importance to governmental evaluation management that managers are offered means that enables them to communicate with the research community. Due to the difference between academic discourse and practical applications, what is brute simplification in the research community may be beneficial to evaluation management

Roundtable A (Second Rotation): Improving Governmental Responsiveness to Indigenous Populations through Evaluation: The Maori Experience

Presenters: *Monitoring and Evaluating Government Performance for Indigenous Peoples: The New Zealand Experience*, David W Earle, Ministry for Maori Social and Economic Development

Improving the social and economic status of Mäori -- the indigenous people -- is a key priority for Government in New Zealand. One of the key strategies of the current Government to achieve this is to increase accountability for, and effectiveness of, Government spending on Mäori. State sector accountability for improving Mäori outcomes has recently been strengthened in New Zealand. The Ministry for Mäori Economic and Social Development, the agency which provides advice to Government on Mäori issues, will have greater involvement in determining departmental accountability for outcomes through purchasing and annual reporting processes, agency reviews, and effectiveness audits of programmes and services. This session will look at the challenges of developing and implementing mechanisms to assess and influence Government performance in improving outcomes for Mäori. It will consider how performance accountability, agency reviews and effectiveness audits can best inform Government policy and enhance the provision of programmes and services to Mäori.

Ticking the Box: Accountability Issues for the Government and Community Providers, David W Earle, Ministry for Maori Social and Economic Development

Many government agencies are now contracting out a lot of their service delivery functions to community providers. This includes an increase in the indigenous people of New Zealand, Mäori, delivering services to their communities.

However, government agencies continue to design the programmes for delivery and set the accountability functions within their guidelines. This generally involves Mäori providers accounting for funding through reporting on programme outputs. This is known to Mäori providers as 'ticking the box'. Consequently Mäori are given little opportunity to provide information in terms of outcomes for their clients. Mäori providers have a number of objectives for their clients and communities. However, little of what they are concerned about is reported on except through occasional evaluations. There is a strong feeling amongst Mäori providers that monitoring systems need to be developed that acknowledge Mäori cultural paradigms. This is particularly important in the present environment where the Government is aiming for increased community and Mäori capacity.

Roundtable B (Second Rotation): Using Evaluation To Inform Governmental Program Development

Presenters: *An Approach to Utilizing Quality Assurance Reviews for Program Development*, Mahassen Ahmad, Texas Department of Human Services

Quality Assurance (QA) reviews are usually internal and formative evaluations which are used as feedback to managers. A formalized approach for conducting QA reviews of Special Nutrition Programs field offices was developed and implemented. The purpose of the approach was to extend the scope of the reviews from assessment of the quality of program implementation to finding means of program development and improvement. This presentation will outline the QA review approach. It will show how the approach not only ensured program compliance but also was a means for achieving consistency and standardization between the field offices, detecting policy and/or procedure development requirements, identifying training and technical assistance needs, and relating necessary individual office corrective action(s). Lessons learned, as a result of applying this approach, will also be presented. The presentation is related to the conference theme of 'Increasing Evaluation Capacity'. It highlights the importance of the QA review efforts as self-evaluation skills used to manage and develop social programs.

Perspectives on Land Use: A Survey of Decision-makers in Northeast Michigan, Murari Suvedi & Sarah Genschaw, Michigan State University

Rural communities are being challenged with an increase in population growth and development. Michigan State University Extension conducted an opinion survey of public officials involved in land and water resources planning in eight Northeast Michigan counties to learn about the issues of land use and community growth. The survey was mailed to 796 public officials that included members of the planning and zoning commissions, elected boards and council members, employed county staff, and Board of Appeal positions. It had a response rate of 69%. Findings indicated that the majority of officials involved in land and water resources planning are over 50 years of age and have been living in the area over 30 years. The majority felt that there has been growth pressure in the area and it will increase over the next five years. Poor public understanding of land use challenges was identified as a major barrier to land use. Respondents identified several educational needs to improve their decision-making skills and abilities and suggested ways to address local land use issues.

Roundtable C (Second Rotation): Roundtable C (Second Rotation): Putting Evaluation on Developing Countries' Reform Agenda, Arild O Hauge, Habberstad Management Consultants

The establishment of evaluation capacity is entering the agenda of public sector reform in an increasing number of developing countries. However, the availability of technical skills and methodology are only a small part of the challenge. More critical is the presence of policy-makers, budget planners and operational managers who want to MAKE USE OF evaluation as part of their decision-making process. Their demand for evaluation is in turn dependent on the prevailing governance culture and public service performance incentive system. In addressing capacity development status, needs and opportunities, our starting point is therefore the broader national context and includes analysis at the levels of: a) policy environment, b) infrastructure of institutional arrangements, and c) technical evaluation skills. Against this framework we explore dimensions of program cycle management; accountability and transparency; and decentralization and participation.

Roundtable D (Second Rotation): Cultural Influences, Evaluation and Government Change

Presenters: *Not Telling It Like It Is: The Impact of Tacit Knowledge on Basic Education Policy and Behavior Change in the HIV/AIDS Epidemic*, Kate A Spring, United Nations Children's Fund and Karin Hyde, Latilewa Consulting Ltd

"When an old man dies, it is as if a library were burnt down," is an African saying that captures the reality of information storage and retrieval in Africa. Oral tradition has survived generations and lessons have been learned and relearned while the process of documentation stays in a non-written form. How are major lessons learned and documented for posterity? How is the knowledge that is held by the "wise" and "experienced" captured and reproduced? What effects do these nuggets of knowledge have on development of social policies? How and why has a particular knowledge survived? Who benefits? These are just a few of the many questions surrounding the transfer of information in Africa. The authors explore two areas of great importance -- policy for basic education and behaviour change in light of the HIV /AIDS epidemic. They seek to document and expose areas in which that tacit knowledge has some influence.

The Use of Proxy Variables To Measure Project Impact: A Valid Alternative Method for Outcome Evaluation, When Final Outcomes Are Not Yet Apparent, Rosa Maria Nuñez-Urquiza, Instituto Nacional de Salud Publica

The introduction will address the delicate issue of the social impact of "project impact evaluations in the Latin American Countries" by reviewing cases studies. Then, a methodology to construct indicators to measure intermediate effects, will be proposed.

Roundtable E (Second Rotation): Meta-evaluation of Large-scale R&D Programs, Heug Deug Hong, University of Manchester

The aims of this paper are to evaluate the evaluation systems, namely meta-evaluation, of national large-scale R&D programs and to find out factors for good evaluation system of large-scale R&D programs. In particular, the focus is to divert from program-based goal-oriented evaluation to a strategic evaluation system to measure the healthy condition of an R&D program.

An evaluation system is a social process and the way in which it is organized is critical to its outcome. The organization of evaluation includes the issues of what is being evaluated, by whom, by which criteria, for whom and to what purpose. In this paper, the core research questions are as follows;
- Which areas of activities of the programme promotion were evaluated?
- How and under which conditions were the evaluation carried out?
- How were the results of the evaluation utilized?

The programs will be consider in terms of the interface between evaluation and the wider policy-making process and particularly the utilization of the evaluation findings and the impacts. The programs for Case Study are Advanced Technology Program(US), Alvey Program(UK), and Brite-Euram program(EU). The academic contribution of this paper is to find out a good evaluation framework, particularly for large-scale research programme and suggest a policy direction for a good evaluation system.

SESSION 467: Business Meeting and Presentation **Room:** Oahu

Sponsored by the Topical Interest Group for Graduate Students

Mentoring 101: How To Create Mutually Beneficial Relationships

TIG Chair: Liesel A Ritchie, Mississippi State University
Program Chairs: Deborah L Wasserman, The Ohio State University
Kerry L Kriener-Althen, University of Minnesota
Presenters: Jody L Fitzpatrick, University of Colorado at Denver
Eunice Rodriquez, Cornell University
Martha Tompkins, Virginia Beach City Public Schools
Jack A McKillip, Southern Illinois University at Carbondale

Cliff Carr, Independent Consultant
Doug Hoy, National Museum of Science & Technology

How does program evaluation differ from social science research? Many experienced evaluators would respond that research skills are only part of the necessary skill set for successful evaluators and that many skills are learned from actually doing evaluation work. And while students can learn the mechanics of doing research and evaluation in graduate school, there is no formalized structure to learn the craft of program evaluation. Mentorships offer graduate students and novice evaluators the opportunity to learn the art of doing program evaluation with the guidance and supervision of more experienced evaluators. This panel session is designed to assist graduate students, novice evaluators, and seasoned evaluators wanting to engage in mentorships. The panel is composed of university, non-university-public, and non-university-private experienced mentors. Each panelist will talk briefly about his or her experiences, followed by interactive dialogue with the audience

SESSION 468: Panel **Room:** Waialua

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

Increasing Evaluation Capacity and Data Based Decision Making in a Decentralized Policy System: the Case of Southwestern Pennsylvania's Regional Education Index

Chair: Christopher D Nelson, Western Michigan University

A recent report by the National Research Council calls for schools to become self-improving organizations. A key component of organizational self-improvement is the ability to collect, analyze, and draw evaluative conclusions from, data. Yet, relatively few schools have the capacity to perform these tasks themselves. The Regional Education Index (REI) project, sponsored by a regional non-profit organization, is a three-pronged effort to enhance the capacity for self-evaluation and data based decision making in Southwestern Pennsylvania schools. The three facets of the project are: (1) an annual public report of regional benchmark indicators, (2) a web-based education information system, and (3) a technical assistance project that helps a targeted group of schools identify and meet data management and evaluation needs. Panelists will represent the perspectives of both designers and stakeholders, and will discuss the challenges associated with facilitating self-evaluation capacity in a decentralized policy system such as education.

Panelists: *Capacity Building and the Need for Political and Organizational Leadership*, Karen S McIntyre, Education, Policy and Issues Center

This presentation, by the president of the non-profit that instigated the REI project, will outline the motivations behind the project and how it serves the organization's mission of facilitating educational excellence in a decentralized regional educational governance system. Design of the REI was predicated on the assumption that indicator systems will be effective only if stakeholders view them as legitimate. Hence, the presentation discusses some of political and organizational challenges involved in developing indicator systems in decentralized political systems and some of the processes and strategies used to overcome them. The presentation will also outline various stakeholder responses, including attempts to use project-generated data in school and district decisions.

Developing Integrated Approaches to Capacity Building in Schools: Implementing a Theory of Change, Jonathan Walkush, Education, Policy and Issues Center

The author, chief policy analyst with the non-profit organization responsible for the REI, articulates a theory of change in which each of the project's three components serves a unique but complementary and mutually reinforcing function in the overall project design. First, a high-profile annual report both provides decision-makers with rudimentary benchmarking data and helps create demand for data based decision making among schools' "constituents." Recognizing that fixed-format reports can be of limited use to administrators, however, the second part of the project includes a web-based education information system that allows administrators to customize the data for their purposes. Finally, individualized school-level consulting with local evaluation talent helps schools identify data needs and develop school-level databases. The consulting teams also help schools develop basic data analysis and interpretation skills, foster trust among teachers and administrators, and, more generally, help develop a "culture" of self-evaluation.

Capacity Building in the Trenches: A School Superintendent's Perspective, Stanley Herman, Woodland Hills School District

The author, a local superintendent and member of the board that designed the REI, provides a school administrator's perspective on data based decision making and the REI's approach to fostering it. The presentation begins by noting briefly that few superintendents receive practical instruction in the application of data based decision-making in their graduate school training. In particular, statistics courses leave superintendents unprepared for the multiple and imperfect data sources that cross their desks. Next, the presentation outlines the types of data most useful to school administrators. The bulk of the presentation outlines the ways in which the author's school district utilized the REI data. These responses include changes in professional development practices as well as modification of budget priorities.

Capacity Building from a State-level Perspective: A Policy Maker's View, Ron Cowell, Education Policy and Leadership Center

The author -- a former member of the Pennsylvania state legislature and the state board of education, and current head of a state-level education policy non-profit -- provides the perspective of a stakeholder concerned with state-level education policy. The presentation begins by discussing what state-level policy makers responsible for the "big picture" can learn from data and from projects like the REI. The author then discusses how the interests of state policy makers and their utilization of data is typically unlike that of district and school administrators. Next, the presentation considers the use of data to initiate and support statewide education policy. Finally, the presentation identifies various barriers that thwart data based decision making - both in schools and in government institutions - and outlines a number of steps the REI team is taking to influence data collection practices by the state department of education.

Discussant: Christopher D Nelson, Western Michigan University

SESSION 469: Panel **Room:** Lanai

Sponsored by the Topical Interest Group on Evaluation Use

Big 'P', Little 'p': Evaluating the Effectiveness of Grant ‘P’rograms through Examining Individual Grantee ‘p’rograms

Chair: Robert G Orwin, Battelle Centers for Public Health Research and Evaluation

Federal, state, and local governments as well as foundations often make grants to multiple grantees around a single program concept. The concept is outlined in a Request for Applications and applicants respond by proposing how they plan to implement the concept. Subsequent evaluations of these Programs often involve exploring individual implementations to determine whether they (1) conform to the Program concept and requirements and (2) are contributing to achievement of the grant Program's overall goals and objectives. However, each grantee adapts to a different program environment, develops at a different pace, and encounters different barriers and facilitators to implementation. Panelists will present challenges and lessons learned from evaluations that adopted creative approaches to accommodating the uniqueness of varied "p"rogram implementations, while also documenting the effectiveness of the "P"rogram as a whole. Techniques developed to involve stakeholders in defining program "success" and to design systems for ongoing monitoring will be shared.

Panelists: *Challenges and Lessons Learned in Refining an Evaluation Strategy for the National Program of Cancer Registries*, Carlyn E Orians, Joanne Abed, Mary Kay Dugan & John Rose, Battelle Centers for Public Health Research and Evaluation; and Leah Simpson, Irene Hall, Mary Hutton, Hannah Weir & Kathleen McDavid, Centers for Disease Control and Prevention

This presentation describes the challenges and lessons learned in refining an evaluation strategy to support an ongoing, legislatively mandated program - the National Program of Cancer Registries - in which program success is dependent on the progress of individual state-based registries. The aim was to work closely with federal program staff as they reviewed objectives and evaluation criteria for the program's first five-year project period and to facilitate the program's plans to continue and enhance the program's evaluation during its second five-year project period. The challenge was to effectively balance the following evaluation goals: (1) monitor compliance with authorizing legislation; (2) objectively measure program outcomes; (3) monitor development of the infrastructure necessary to achieve desired outcomes; and (4) emphasize use of registry data to support cancer control. Moderated staff meetings and work groups were used to engage staff in a dialogue on these issues.

Finding the Big Picture: Program Evaluation in Diverse Environments, Mary Odell Butler, Battelle Centers for Public Health Research and Evaluation

Program evaluators often are asked to infer the effectiveness of federal programs by examining program implementations in diverse settings and at different stages of implementation. This paper discusses an evaluation of the Centers for Disease Control and Prevention (CDC) Field Epidemiology Training Program (FETP) which seeks to protect the US against imported infectious disease by providing technical assistance for epidemiology training in 17 countries around the world. CDC requested the evaluation in order to capture the lessons learned across all of these national programs, hoping to use this information to improve established programs and implement new ones. The paper considers the methodological challenges of defining the evaluation in such a way that the objectives of the core program could be assessed by examining implementations that were evolving on their own trajectories, partially determined by a core program design but necessarily adapted to local conditions.

The International Cooperative Biodiversity Groups Program of the John E. Fogarty International Center at the National Institutes of Health: A Broad Mandate to Grantees and an Emphasis on Innovation, Joanne Abed, Battelle Centers for Public Health Research and Evaluation & J Kirby Weldon, Office of International Science Policy and Analysis

The International Cooperative Biodiversity Groups (ICBG) Program awarded cooperative agreements to five grantees engaged in a complex undertaking to develop alternative uses for natural products harvested from endangered environments. The ICBGs were to promote the identification of natural products with promising commercial uses, but in an innovative manner that would utilize expertise across multiple disciplines, protect indigenous knowledge and rights, enhance source-country capacity, and build a sustainable infrastructure. Among the aspects of the program that posed a challenge to evaluators were the broad mandate, the multidisciplinary nature of the Program, and varied source-country contexts and implementations. Additional challenges included measuring indirect effects, assessing long-term outcomes, and operationalizing 'elusive' indicators. Among the techniques used to address these challenges was electronic data entry by respondents into matrices that (singly) highlight individual program achievements and (together) document effectiveness of the Program overall. The matrices were subsequently incorporated into routine progress reporting by grantees.

SESSION 470: Business Meeting and Demonstration **Room:** Molokai

Sponsored by the Topical Interest Group on Theories of Evaluation

The Synthesis Problem: Issues and Methods in the Combination of Evaluation Results into Overall Evaluative Conclusions

TIG Chairs: Barbara Lee, University of South Florida
Prachee Mukherjee, Colorado Springs School District 11

Session Chr: Michael Scriven, Claremont Graduate University

Presenters: Michael Scriven, Claremont Graduate University
E Jane Davidson, Claremont Graduate University

One of the ways in which evaluation differs fundamentally from applied social science research is in the way results are reported. Rather than simply presenting results on a range of variables and allowing the reader to draw his or her own conclusions, the evaluator often needs to pull all this information together into overall determinations of absolute or relative merit, worth, or significance. In this session, the presenters will demonstrate a range of methods for combining different kinds of data in order to grade or rank programs, personnel, and products (examples of each will be used). Methods include: (a) the basics of combining qualitative and quantitative data, (b) evaluation and synthesis of results on non-interval multiple-choice items, (c) the qualitative weight and sum (QWS) method, and (d) synthesis of results using program logic/theory-based evaluation.

SESSION 471: MultiPaper **Room:** Kauai

Reconsidering the Role of the Evaluator II

Chair: Katherine Ryan, University of Illinois at Urbana-Champaign

Historically, the evaluator's role was one of detachment, avoiding bias, with great faith placed in the methods selected for the evaluation. Weiss (1998) links some of the re-alignment of the role of the evaluator to program people to the constructivist notion of multiple perspectives and multiple realities. She suggests these new roles, e.g., "critical friend, co-investigator, facilitator, and problem-solver" signal different kinds of understandings about what evaluation is and should be.

What are the implications of these shifts for the role of the evaluator in theory and practice? In this double session, presenters and the audience will engage in dialogue to collectively reconsider the construct of evaluator role in evaluation theory, practice and society. Questions to be considered include whether the construct of role may mask as much as it reveals for contemporary evaluation practice and whether "role" adequately represents the processual nature of the evaluator in the evaluation context and society. Collectively, the audience and presenters will propose alternative conceptualizations that may more accurately portray who and what the evaluator is and does.

Presenters: *Roles as Masks for Evaluators,* Jennifer C Greene, University of Illinois at Urbana-Champaign

A role is donned and doffed, put on and taken off. Like an overcoat or a pair of snow boots, a role covers up the person underneath, her individuality and identity, as reflected in her stylish suit or her Birkenstock sandals with Gold-Toe socks to be played or And everyone does it. Everyday. Without shame.

But the construct of role may mask as much as it reveals for contemporary evaluation practice. For much that is important about this practice concerns (1) our relationships with one another and (2) the positioning of our selves and thus our work with respect to global crises of poverty, hunger, injustice and unequal distribution of resources. Forging relationships with others that are authentic, respectful, and caring requires not roles but meaningful engagement of self. And dialing in, connecting - or not - one's own evaluation work with global concerns and issues is a perversion of "scientific citizenship" if it's just a role. Scientific citizenship also invokes a commitment of self. This paper will use case examples to argue that at least some of our important work as evaluators is better captured as qualities of selfhood than operational functions of roles that can be put on and taken off, donned and doffed.

Reflections and Contributions on the Role of the Ethnographer in Evaluation, Rodney K Hopson, Duquesne University

The role of the ethnographer in evaluation is, at best, between amorphous and non-existent. Except for Fetterman's pioneering work over a decade ago (1984, 1989), little to no work pays attention to the ethnographer's role in evaluation or the connection between ethnography and evaluation. In keeping with the aim of the larger panel session, this paper contemplates how the role of the ethnographer in evaluation further contributes to the different kinds of understandings about what evaluation is and should be. The purpose of this paper is twofold. One purpose is to illustrate the role of the ethnographer in evaluation as depicted in two interventions, one in a community-based disease prevention program in Baltimore, MD and the other in an elementary school in Pittsburgh, PA. A second purpose of this paper is to identify a number of future directions the role of the ethnographer may play in evaluation. Rather than merely imply that the identity of the evaluator shapes (and is shaped by) the evaluation (Lindesmith, Strauss, and Denzin, 1999), the paper suggests that the ethnographer's contribution in evaluation has far-reaching implications to evaluation theory, practice, and society.

Title: The Evaluator's Role in the Transformative Context, Donna M Mertens, Gallaudet University

The role of the ethnographer in evaluation is, at best, between amorphous and non-existent. Except for Fetterman's pioneering work over a decade ago (1984, 1989), little to no work pays attention to the ethnographer's role in evaluation or the connection between ethnography and evaluation. In keeping with the aim of the larger panel session, this paper contemplates how the role of the ethnographer in evaluation further contributes to the different kinds of understandings about what evaluation is and should be. The purpose of this paper is twofold. One purpose is to illustrate the role of the ethnographer in evaluation as depicted in two interventions, one in a community-based disease prevention program in Baltimore, MD and the other in an elementary school in Pittsburgh, PA. A second purpose of this paper is to identify a number of future directions the role of the ethnographer may play in evaluation. Rather than merely imply that the identity of the evaluator shapes (and is shaped by) the evaluation (Lindesmith, Strauss, and Denzin, 1999), the paper suggests that the ethnographer's contribution in evaluation has far-reaching implications to evaluation theory, practice, and society.

Discussant: Jean A King, University of Minnesota