Friday, November 3, 1:50 P.M. to 3:20 P.M.

Sponsored by the Topical Interest Groups on International and Cross-cultural Evaluation & Teaching of Evaluation
Cross-cultural Considerations for Evaluation Training: Comparisons and Contrasts from US, Japanese and Australian Perspectives
Chair: Ross F Conner, University of California at Irvine

There is an increasing need for well-trained evaluators. This need is especially felt in countries outside the US that are experiencing a sharp increase in the interest in evaluation and the consequent need for evaluators.   Evaluation training, however, cannot simply be moved from the US context into a foreign context. Different cultural perspectives on evaluation itself, as well as on training, necessitate different training components and approaches. This session will present three perspectives, one from the US, one from Japan and another from Australia. Each presenter will describe the general evaluation training occurring in his culture. Then, discussants with experiences in cross-cultural evaluation and training will comment on these perspectives, following which the audience will be brought into the discussion. The goal of the session will be to distill significant lessons learned about the critical components of evaluation training that cut across cultures and those that are more culture-specific.

Panelists: *Graduate Training in Evaluation: Developing Skills and Sensitivities*, Ross F Conner, University of California at Irvine

Conner teaches graduate level program evaluation at the University of California Irvine. His US-based approach uses both didactic and experiential components, in order to address his belief that there are "skills" and "sensitivities" that evaluators must learn. The skills could be considered the "science" of evaluation; the "sensitivities" component is the "art" of evaluation. The evaluation skills are learned largely through textbooks and class lectures and discussions. The evaluation sensitivities are learned largely through students' experiences working with an on-going social program to develop a full evaluation plan. For his presentation, Conner will briefly describe the components of his course and provide examples of student project evaluation plans. He will compare and contrast his class with others in US, closing with a set of US-based evaluation training characteristics and components.

Adapting Evaluation Training to the Japanese Cultural Context, Masafumi Nagao, Hiroshima University

Nagao conducts research on program evaluation at the Center for the Study of International Cooperation in Education, Hiroshima University. Given the recent surge of interest in evaluation in Japan, his principal concern is how to elaborate programs of education and training for evaluators. There is as yet no systematic evaluation training in the country going beyond cost-benefit analysis, financial audit and administrative inspection. In the hierarchical, cultural environment of Japan, there is not much room for open and independent evaluation. However, in industry, continuous process review for technological and managerial improvement has become an accepted practice. In his presentation, Nagao will briefly describe how the growing interest in evaluation practice is leading to the societal recognition of the need for systematic evaluation training, discuss the importance of evaluation culture, and, then, explain how culturally-sensitive evaluation training programs may be formulated using, as examples, the schemes being elaborated.

A Dedicated Post Graduate Teaching Program in Evaluation: An Australian Example, John M Owen, University of Memphis

Owen jointly coordinates a Master of Assessment and Evaluation at the University of Melbourne. Within this program, students can take a major in evaluation; students typically take six courses and a thesis. Some courses are 'epistemologically led' such as Evaluation for Management and Development, while others are 'methodologically led', for example Qualitative Data Methods. There is a strong experiential flavour throughout, and theory is discussed taking into account recent practices of participants, staff and students. The term evaluation is used in a more eclectic fashion in Australia than in most countries. This is in response to the perspective that evaluative enquiry should affect all stages of policy/program decision making, and so it provides an umbrella for topics such as benchmarking, action

research and performance auditing. The presentation will elaborate on the issues discussed above as a basis for comparison with the situation in other countries.

Discussant: James R Sanders, Western Michigan University
              Jim Rugh, CARE

SESSION 439: MultiPaper   **Room:** Akaka

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

Mandated Collaboration: A Feminist Issue in Evaluation

Chair: Denise Seigart, Mansfield University
Presenters: *Lessons in Capacity Building: Requiring Grassroots Projects to Present Their Evaluations*, Michael Luxenberg & Anne Betzner, Professional Data Analysts Inc

Grassroots organizations often avoid evaluation because they do not see it as useful and the findings may challenge their assumptions. Two Minnesota state agencies contracted independent evaluators to overcome these barriers by building the capacity of 22 community-based sexual violence prevention projects to conduct outcome evaluation, and to create an evaluation community where projects are supported in sharing results. Evaluators used six methods of training and technical assistance to achieve this goal, the most important of which was requiring projects to present evaluation results at a statewide conference illustrating best practices in the field of sexual violence prevention. Despite substantial challenges, project staff were successful in completing their evaluations, and the vast majority reported using their results. This presentation discusses the benefits of instituting a forum for accountability in evaluation, effective methods of training and technical assistance, and conditions in which capacity building evaluation initiatives are most effective.

A Model of Empowerment Evaluation for Sexual Assault Service Delivery and Prevention Programs, Kelly Bennett, Freya Bonter, & Heather Dorey, Michigan Public Health Institute

This presentation is an overview of an empowerment evaluation model that provides technical assistance to sexual assault agencies to build their program evaluation capacity and expertise. Evaluating sexual assault prevention and victim service delivery programs poses many challenges, including limited resources, lack of evaluation experience, and the unique challenge of protecting the safety and confidentiality of rape survivors while maintaining scientific integrity. The long-standing values and philosophies of violence against women prevention and service agencies must also be respected. In this presentation, we will describe the Michigan Sexual Assault and Rape Prevention (SARP) Evaluation Project, which offers one model for addressing these multiple issues. The SARP Project produces step-by-step evaluation training workbooks, conducts evaluation trainings, and provides frequent, individualized technical assistance follow-up to agencies throughout Michigan. This presentation will describe how the SARP Team was effective in building the evaluation capacity of both sexual assault prevention and service delivery programs. Recommendations for planning similar empowerment evaluation projects will also be discussed.

Increasing Evaluation Capacity among Women's Centers: One Encouraging Result of 'Mandated Collaboration', Sharon M Wasco, Jennifer Camacho, Larry Bennett, Rebecca Campbell, Lisa Frohmann, Stephanie Riger & Paul Schewe, University of Illinois at Chicago

Rape crisis centers and domestic violence shelters have met funders' increasing requests for outcome evaluation of their service programs with some resistance. This resistance is grounded in philosophical opposition to what is often perceived to be government, or funder, regulation and surveillance; programs' desires to protect victimized clients and to not compromise the nature of crisis services; as well as practical restraints such as limited evaluation capacity within their organizations. This presentation will highlight how these issues became salient in the context of a collaborative evaluation project initiated by a state funder, facilitated by a team of feminist, university-based researchers, and ultimately implemented by approximately 80 sexual assault and domestic violence programs in varying community settings across the state of Illinois. Project results include a year's worth of standardized outcome data for services statewide as well as increased resources and technical skill for implementing program evaluation

among participating agencies.

Discussant: Sharon Rallis, University of Connecticut

SESSION 440: MultiPaper   **Room:** Koko
Sponsored by the Topical Interest Group on Assessment in Higher Education
Assessment Methods in Higher Education
Chair: Cynthia M Hewitt-Gervais, Florida Gulf Coast University
Presenters: *Why Survey Online? A Practical Look at Issues in the Use of the Internet for Surveys in Higher Education,* Christina S
                   Ballantyne, Murdoch University

Using the internet to conduct surveys has increased considerably in the last few years. Web-based survey forms are easy to construct and free survey generators are available on the internet. Murdoch University in Western Australia has been using online survey forms over the last two years, both as a 'stand alone' method of collecting student ratings of teachers, and as an option to try to increase the response rate in other surveys of students and recent graduates. This paper examines the benefits and pitfalls of using an online survey, including security issues, cost effectiveness and response rates. Are those respondents who choose the online option different from those who complete the more traditional paper form, and if so need this be a concern? How can we encourage more use of the online option with its savings in mail and data entry costs?

Expanding Assessment Capacity in Higher Education, John S Klemanski, Lizabeth A Barclay & Kenneth A York,
        Oakland University

Universities are seeking alternative forms of assessment beyond class performance. The experiences of several university departments that have conducted focus groups, in part because of accreditation requirements, are detailed. These sessions sought student perspectives on curriculum, scheduling, career preparation, program services and possible reasons for student failure. The groups were also used as a teaching tool in a program evaluation and a research course. In some departments, this process was broadened to include separate sessions with undergraduate students and other stakeholder groups. In addition, a pilot study of an assessment center for outgoing students as an evaluation tool is also briefly discussed. Both of these techniques are qualitative methods, which provide information useful to program review, are best used in conjunction with other methods. The cost-effectiveness of the techniques will be discussed. Using the tools described can increase a department's assessment capacity at a very low cost.

*The Study of the Psychometric Properties of the Learning and Study Strategies Questionnaire (LASSI): Why
        Universities Should Not Use the LASSI*, Catherine E Moak, Sam Houston State University

More and more, students are coming into universities under-prepared. They lack the learning and study skills necessary to survive the academics of the university. An instrument that provides valid data is necessary to assess the needs of these students. The most commonly used instrument for assessing learning and study skills is the LASSI. Unfortunately, there is no established validity relating to scores from the LASSI instrument and the scales given by the author of the instrument. This research study will examine the construct, concurrent, and predictive validity of the LASSI data by comparing the results with the MSLQ and the SAMS data on similar constructs, by examining the predictability of the instrument with grade point average, and by conducting a factor analysis.

Interjudge Reliability of Ratings from Engineering Design Assessments, Michael S Trevisan, Larry J McKenzie &
        Denny C Davis, Washington State University

Enhancement of undergraduate engineering education continues to receive considerable focus and attention by the engineering community. This attention is reinforced by accreditation requirements calling for continual assessment and evaluation of student and program outcomes. A recent NSF-funded project developed an assessment system to evaluate engineering design competencies. The assessment system consists of short-constructed response, essay and performance assessments. This paper reports findings of a pilot study concerning the reliability of ratings obtained from application of scoring criteria used to assess student design competence. Intraclass correlations for

scores from two faculty raters judging student performance in two undergraduate programs were computed. Results show improved reliability after refinement of the scoring criteria. Implications for evaluation practice include the importance of clear criteria for judging student work and making decisions about program quality. Ways in which this process is being used to develop the evaluation capacity of engineering faculty are discussed.

SESSION 441: Panel   Room: Waimea
Sponsored by the Topical Interest Group on Quantitative Methods: Theory and Design
Measurement Issues in Evaluation: Conceptual and Methodological Issues with Longitudinal Designs
Chair: Katherine M McKnight, University of Arizona

Evaluators often assess program effectiveness via the measurement of change over time in an outcome(s) of interest. Yet as Rogosa (1991) reminds us, measuring change is seldom simple. Evaluators must be able to conclude that change in an outcome is not confounded with change in the measurement of that outcome. Therefore methods for analyzing change rely on the assumption of measurement invariance, i.e. the measurement model for an outcome remains constant over time. This assumption may not be tenable in some, perhaps many, evaluation situations. In this panel, we discuss issues related to the assumption of measurement invariance, providing examples and offering a variety of solutions.

Panelists: *The Use of Factor Scores in Growth Curve Analysis*, Lee B Sechrest & Opher Caspi, University of Arizona

Studies implementing growth curve methodology and analysis may have multiple dependent variables, suggesting the desirability of data reduction by factor analysis. Two options present themselves: (1) factor analyze the raw data, calculate factor scores, and then determine growth parameters for those factor scores; (2) determine growth parameters for each of the variables, and then factor analyze those growth parameters and calculate factor scores subsequently. This problem has been examined in the context of other kinds of data transformations, and the literature is reviewed. Additionally, results of simulation studies are presented.

Measurement Invariance and Factorial Stability: Making Sense Out of Seemingly Senseless Measures, Patrick E McKnight, University of Washington

Items on a test contribute differentially to a sum score based upon the item variance. Items with higher variance contribute more to the total score variance than items with lower variance. Although this point may seem rudimentary, it may be extremely important and usually ignored in measurement of people over time or even in cross-sections. As items change over time, the item's contributions to the common factor change. In addition, if evaluators rely on sum or total scores for either cross-sectional or longitudinal research, their results often will reflect unknown variations that usually occur at the item level rather than the factor level. The following paper and presentation will illustrate these points and discuss methods to resolve these problems.

Masking Variability Over Time: The Problem with Composite Scores, Katherine M McKnight, University of Arizona

To demonstrate program effectiveness, evaluators are often interested in establishing interindividual differences in change in an outcome of interest. Yet, as Nesselroade (1991) observed, interindividual differences can be confounded with other sources of variability, i.e., intraindividual variability and intraindividual change. A lack of awareness of these sources of variability can confound interindividual differences in change. In this paper, we will demonstrate the difficulties of distinguishing between these sources of variability when using a common scaling method in which items are weighted equally and summed to create a composite score for the outcome measure. Such a method can be insensitive to intraindividual variability, intraindividual change and interindividual change due to the focus on quantitative change at the composite score level and ignoring qualitative change due to differential item functioning within that composite. Strategies for addressing these issues will be presented and illustrated with examples.

Measurement Invariance in Evaluation: Conceptual and Methodological Issues, Souraya Sidani, University of Toronto

In evaluation studies, multi-item outcome measures should demonstrate measurement invariance, or similar patterns

of factor loading over occasions of measurement. Yet, measurement invariance may be difficult to maintain at the conceptual and methodological levels. Some interventions may result in qualitative change in the outcomes, which is difficult to examine under the notion of measurement invariance. Outcome measures that capture change are those most responsive to the interventions, and demonstrate low inter-individual variability in the total sample at pre-test and within each group at post-test, but increased variability in the total sample at post-test, due to group differences in the response to the interventions. Therefore, differences in the extent of inter-individual variability obtained for the total sample at each point in time is associated with variability in the factor loading patterns, which may be misinterpreted as indicating lack of measurement invariance. Strategies for addressing these issues will be presented.

SESSION 442: Demonstration   Room: Niihau
Sponsored by the Topical Interest Group on Program Theory and Theory-driven Evaluation
Understanding the Theory of Change Approach to Evaluation
Facilitator: Andrea A Anderson, Aspen Institute

Over the past five years, the 'theory of change' approach to evaluation has grown in popularity and in application. This approach has become popular because surfacing theories of change permits an evaluator to answer questions about program effects, as well as 'how?' and 'why?' a program reaches intended outcomes--questions often left in an unexamined 'black box' when we conduct program evaluations using traditional approaches. This demonstration session will show participants how the theory of change approach can be used in evaluations of comprehensive community initiatives (CCIs). Examples drawn from the CCI field will be used to show how theories of change have been developed, how to determine if a program theory meets four essential quality standards, how to make the 'crosswalk' from program theory to evaluation design, and how to make mid-course modifications to the theory of change without jeopardizing the ability to draw conclusions about a program's effectiveness.

SESSION 443: MultiPaper   Room: Hilo
Sponsored by the Topical Interest Group on Non-profit and Foundations Evaluation
Building Capacity in Non Profits through Planning and Evaluation
Chair: Donald Tobias, Cornell University

This Session will include four presentations that examine the planning process and its potential effect on non-profit organizations. The planning process common to all these projects is concept mapping. All of these papers use the concept mapping methodology to collect and evaluate input from stakeholders in addressing organization issues. Issues of engagement and representation in planning are informed by this collection of papers. The use of technology as a means of increasing both the quantity and quality of information would be discussed in this session. The opportunity to build on the planning database as a means of program evaluation is evident in all this projects. The question of who participates and how they participate will be examined in a variety of settings.

Presenters: *The Potential for New Technologies to Re-invigorate the Practice of Organizational Strategic Planning*, Jonathan A
                Gans, Cornell University

This presentation will review the state-of-the-art in models for strategic planning and propose a new model based on the Concept System practice methodology. Models for strategic planning have changed so little over the last 40 years that management theorists and practitioners alike have suggested that the process is counter-productive. Part of the reason for this stems from the fact that the management community continues to put new laces on an old shoe. Advances in communications technology and computer software applications have opened up tremendous possibilities for achieving the original goals of organizational strategic planning. One such approach is based on the Concept System model. Many of the limitations identified by critics of the present state-of-the-art in strategic planning will be discussed and contrasted with the potential of the Concept System model to address and overcome these limitations.

Participant Designed Training for Non-profit Managers, Kristin Jackson, Cornell University

The increasing recognition of the important role of nonprofit organizations in our society has led to the creation of

training programs. The content of these training programs is typically organized around the functional areas of managers and/or board members. This paper will focus on the use of concept mapping to design and evaluate the curriculum of a nonprofit management institute. The use of technology provides an opportunity to better coordinate expectations of participants and presenters prior to the event and a more focused evaluation of the curricula post event. This paper examines the potential for the use of planning data set as a means of monitoring and identifying continuing training needs of participants. The engagement of participants to design and evaluate their training experience is in the best traditions of adult education and capacity building.

Increasing Participant Engagement through Planning and Evaluation, Hedwig E Lee, Cornell University

This presentation will focus on the use of concept mapping as a process for increased involvement of key stakeholders in a community program. The project is a university and non-profit partnership providing education and technical assistance regarding food security to targeted populations. Not unlike other mature community based interventions, this program has moved through a number of planning and evaluation cycles. The concern of the project managers is the reduced level of engagement in planning by stakeholders and the decreased motivation of those who remain engaged. This paper examines the degree to which the concept mapping methodology expands the degree of participation in planning. If participants perceive higher levels of ownership in the creation and management of planning data will they become more engaged? We believe this paper shows such increases and resulting improvement in the capacity of the project and community.

Collaborative Program Planning and Evaluation, Donald Tobias, Cornell University

This paper will focus on the use of technology to support a highly collaborative planning and evaluation process. The engaged university has become synonymous with an increase in activity linking academic and community interests. Implicit in these collaborative efforts is a move from "expert" driven models to more participatory approaches. Stakeholder involvement is demanded by many funders and is a viewed as basic to capacity building at the community level. How then can evaluators and project managers create a process that is both accessible and manageable? This paper suggests that mixed methods using technology and interpersonal communication can produce high levels of involvement at relatively modest costs. The specific project involves academic, government and non-profit partners in developing a statewide program in Economic Development. The database created for planning became the basis for a technology supported evaluation process.

Discussant: William Trochim, Cornell University

SESSION 444: Business Meeting   Room: Puna
Sponsored by the Topical Interest Group on Minority Issues in Evaluation
TIG Chairs: Ivy Jones Turner, Peace Games Inc
                   Hazel L Symonette, University of Wisconsin at Madison
Program Chair: James Earl Davis, University of Delaware

SESSION 445: MultiPaper   Room: Kohala
Sponsored by the Topical Interest Group on Evaluation of Services for Special Needs Populations
Impact of Impact Evaluations: Results for Consumers
Chair: University of British Columbia
Presenters: *Application of Evaluation Results by ASHA's Legislative Council Assemblies*, Herbert M Baum & Sarah C Slater,
                 American Speech-Language-Hearing Association

The American Speech-Language-Hearing Association (ASHA) represents the professions of speech-language pathology and audiology. In 1998, the Legislative Council--a 150-member governance group--restructured into two separate bodies to address issues relative to the individual professions. After each semi-annual meeting, assembly members provide feedback on the content, format and outcomes of their activities. This presentation will describe the mechanisms used to evaluate the meetings and discuss how the evaluation results have been applied to planning subsequent meetings.

Family Voices: An Evaluation of the Impact on Families of the Canuck Place Children's Hospice Program, John B Collins, Rose Steele & Karen Cook, University of British Columbia; Betty Davies, University of California at San Francisco; and Ingrid Pipke, Ingrid Pipke and Associates

Perhaps there is no family event so stressful as the death of a child. Since opening its doors in November 1995, Canuck Place is a free-standing, multi-service children's hospice that has served more than 250 families who have children with progressive, life-threatening illnesses or who have had a child die suddenly--often traumatically. The first of its kind in North America, Canuck Place underwent a participatory action research-based evaluation during its beginning years to identify and correct any problems early in its operational history. End-of-life care, respite services and bereavement support were each examined separately. Both qualitative and quantitative evaluation techniques with more than 270 individuals (ill children, parents, siblings, hospice staff, physicians, and volunteers) showed that the hospice program provided critically helpful services and support, but that there were also opportunities for improvement and greater program articulation and community outreach. Overheads and hand-outs will augment the paper presentation.

Longitudinal Evaluation of Preschool Special Education Services, L Dee Torre, MGT of America, Inc.

The presentation will address the design strategies and issues associated with seven year longitudinal study, following 3- and 4-year olds with disabilities from preschool to fourth grade. The purpose of this first statewide longitudinal study is to determine the long-term effect of preschool special education programs and services on the education achievement, emotional well-being, social adjustment and placement of school-age students through Grade 4 who received special education programs and services during their preschool years. The paper will be presented by the team of researchers from MGT of America, Inc conducting the study and a designated New York Department of Education official. The presentation will highlight data management issues, instrument development, technological implications over the seven year timeframe and other unique sustainability considerations for this special population. The State has a special interest in gathering information on integrated and separate service arrangements for all disability categories.

SESSION 446: Demonstration   **Room:** Kona

Sponsored by the Topical Interest Group on Qualitative Methods

Using Comparative Case Studies to Understand State-wide Systems Change in Substance Abuse Prevention Program Coordination

Chair: Beth L Green, Northwest Professional Consortium Inc
Presenters: Beth L Green, Northwest Professional Consortium Inc
             Scott W M Burrus, Northwest Professional Consortium Inc

Since 1998, the Center for Substance Abuse Prevention (CSAP) has funded 19 state incentive grants (SIGs) to help states improve the way prevention programs are planned and funded. By focusing on changing prevention service systems, this initiative hopes to create more enduring changes in the quality of prevention programs and their associated outcomes. However, evaluating systems change is a complex and difficult undertaking that cannot be readily accomplished using traditional experimental evaluation techniques. The proposed demonstration will show how evaluators can develop and utilize a comparative case study design to systematically describe and evaluate the implementation and outcomes of systems change programs. Evaluators will learn how to efficiently synthesize qualitative and quantitative data from multiple sources into comparative systems models, using NUD*IST to facilitate data synthesis. The use of logic models to facilitate theory-based comparative analysis will be highlighted, using data collected as part of Oregon's SIG project.

SESSION 447: Panel   **Room:** Ewa

Sponsored by the Topical Interest Group on Extension Education Evaluation

Tales from the Trenches: Real Life Examples Using Logic Models

Chair: Molly Engle, Oregon State University

Repeatedly, Extension Service faculty and staff are required to be accountable for the outcomes of their programs, yet, they are struggling with how to demonstrate accountability. When the topic of logic models comes up in conversations about evaluative inquiry, faculty look confused, befuddled, lost. This session will provide a "talking" opportunity for Extension Faculty to share real life examples of how they used logic models successfully, or not, as the case may be. The objective is to share, in an informal format, what worked, what didn't work, what was learned from the experience, what barriers were experienced, what would be done differently in the future. Audience members as well as panelists will be invited and encouraged to participate in this conversation.

Panelists: *"Seeing" Differences: Comparing Theory at Different Sites of the Financial Counseling Program*, Midge Smith, University of Maryland

What we did was practical, simple, quick, and made it very easy to see the differences in what was happening at the different sites. Not real high level theory (from academic viewpoint), but very useful for decision making.

A Logic Model for Planning and Reviewing "Quality" Outreach, Mary P Andrews, Michigan State University

Michigan State University, along with a number of other higher education institutions have been working on defining and integrating the Outreach function into fundamental scholarly activities of the University. Models for unit analysis and appraisal, as well as for project and personnel appraisal have been developed to assist groups in articulating the values and criteria expected in quality outreach. Each unit or situation can be different, but a core set of principles guide the evaluation process. The logic underlying the evaluation, can thus, be standardized, even though individual elements are individualized.

Logic Model Use in the Development and Evaluation of a Breast Cancer Education Program, Kathleen S Tajeu, Auburn University

Auburn University Extension Service has developed a breast cancer eduation program for use in communities. Logic model use in the development and evaluation helped us clarify issues and focus our planning.

Designing Programs of Practical Education, John A Michael, US Department of Agriculture

Will discuss experience of using logic models to design non-formal educational programs, working in collaboration with State Cooperative Extension Services.

Logic Models in Planning State Strengthening Projects in Delaware, Leslie J Cooksy & Patricia T Nelson, University of Delaware

This presentation will describe the introduction of a logic model approach to extension agents responsible for implementing and evaluating Delaware's State Strengthening projects. Then, the use of the logic model approach in the project evaluations will be discussed. The presentation will conclude with an outline of the strengths and weaknesses of the logic model approach in this situation

Building Evaluator Capacity Using Logic Models to Teach, Ellen Taylor Powell, University of Wisconsin

Discussion will focus on our experience in building capacity in using the logic model for improved program planning and evaluation and impact at the organizational level. With a common language and simplified model to better plan, describe and evaluate their programs, the "lights are going on" for many extension faculty and staff. I will share the logic model we are using, our training process and activities and the lessons we are learning as we build this capacity at the individual and organizational level.

Discussant: Molly Engle, Oregon State University

SESSION 448: MultiPaper   **Room:** Honolulu

Evaluation of Family Support and Family-focused Services
Chair: Robert L Fischer, Families First
Presenters: *The Family Resolutions Project: Methodological Considerations and Complications*, Josephine Allen, Christopher
              Cihlar & Monica Casares, Cornell University

The Family Resolutions Project is an initiative of the New York Office of Children and Family Services aimed at ensuring children's safety while preserving families, reducing reliance on government and improving family self-sufficiency. Through a variety of tactics, the project seeks to encourage families to create their own strategies for solving their difficulties. The program began in 1997 and seven counties were involved when the evaluation commenced. These counties were given broad discretion in determining how they implemented the project. Because this program was designed with the strengths based approach as a guiding principle, an evaluation had to be both participatory and able to account for the nuances that defined the program in the several counties. This paper examines the methodological considerations that arose with an approach that strove to include the voices of not only different stakeholders but different stakeholders that experienced the program in seven unique organizational systems.

Assessing Client Change in Family Counseling: Monitoring Service through Outcome Measurement, Robert L Fischer,
    Families First

Family counseling services are a mainstay in the field of community-based behavioral health. This paper draws together what is known about such inquiries and illustrates how data collection can be embedded in the provision of counseling services, such that credible results are generated without disrupting treatment. This paper examines the experience of a large multi-service family service agency, in developing and implementing a client session self-rating instrument, to both monitor client feedback and improve agency-counseling services. This approach emphasizes the assessment of treatment outcome based on the client's perspective of his or her own condition. Data were collected from clients in a family-counseling program at the beginning of each session of counseling on a scale of day-to-day functioning and emotional coping. Data were compiled on all cases upon termination of the counseling episode. Data will be presented on clients who completed counseling during a two-year period.

National Evaluation of Family Support Programs: Results of a Meta-analysis, Jean I Layzer & Barbara D Goodson,
    Abt Associates Inc

As part of the five-year, congressionally mandated National Evaluation of Family Support Programs, conducted for the US Department of Health and Human Services, a meta-analysis was conducted to assess the effects of family support programs and services on children and families. A new hierarchical coding system, developed to capture multiple program and study characteristics, was applied to more than 650 studies. From these, 325 experimental and quasi-experimental studies of 173 programs were selected for inclusion in the analysis. Single-predictor analyses and multivariate modelling techniques were used to test the relationship between selected program and treatment characteristics (predictors) and effect sizes. The study reports the impact of family support programs and services on outcomes in several domains, including: child cognitive development and school achievement; child social-emotional development; child health; parenting attitudes, knowledge and behavior; parent educational achievement and family economic well-being; and parent mental and physical health.

Evaluation of an Investment Club as a Human Service: Assessing Social and Financial Assets of Women in Recovery,
    Frank A Sansone, University of West Florida

Reports on the implementation of an investment club model for assets development with women receiving aftercare services from the Women's Transition Center (WTC). Social Security reforms propose Individual Development Accounts despite the poor financial literacy of Americans. This interim report presents first year results of a group of women in recovery who participate in an investment club to build financial assets, including wealth, financial literacy, and social assets as defined by an enhanced sense of empowerment, social support, and social capital. Club outcome measures are as follows: social assets of social support measured by the Interpersonal Support Evaluation

List and the complementary Support Network List; social capital measured by memberships and voting behavior; financial asset gains in knowledge measured by a modified Financial Literacy Test from the JumpStart Coalition, and the value of tangible financial assets (home ownership, savings, and stocks). Implications for human services and professional development of evaluators are discussed.

SESSION 449: Roundtables   **Room:** Kahuku
Capacity Building Tools and Tactics
(This session includes two 45-minute rotations of roundtables. The Host will ask the tables to rotate at 2:35.)
Host: Libby Cooper, Charities Evaluation Services

Roundtable A (First Rotation): Building Capacity through Technology Use
Presenters: *Using Technology to Build Capacity in the New Millenium*, Margaret M Lubke, Debra Cheney & Catherine Callow
            Elwell, Utah State University

With increasing demands from both government and the public for self-assessment and accountability in education, monitoring has become an important specialization in the field of evaluation, often incorporating management information systems (MIS; Patton, 1997). However, as Patton suggests, designing MISs to be useful--and then actually getting them used--is tricky. The IDEA-SIS MIS was designed and developed with extensive cooperation and input from key stakeholders including special educators, administrators, school psychologists, and university faculty from special education and evaluation. IDEA-SIS brings technology, evaluation, and monitoring together by
* supporting timely self-assessment procedures,
* providing a prompting mechanism for insuring compliance to regulations,
* increasing the flexibility to customize reports and other communications for monitoring and federal reporting,
* building capacity by increasing the amount of data that can be managed and reducing the level of expertise required to manage it,
* improving the data's reliability and accessibility, and
* establishing an on-going focus on evaluation.

Using Technology as a Tool for Increasing Evaluation Capacity: A Step Forward or Backward? Patricia B Keith,
      Abraham H Wanderman, Jessica Snell-Johns & Melanie Dye, University of South Carolina

The purpose of this roundtable is to facilitate a beginning dialog among evaluators who are currently using various kinds of technology (e.g., Internet, fax, e-mail, conference calls) during their projects. We will bring to the table a brief overview of how the field appears to be using technology, briefly share our field experiences using various kinds of technology, and spend most of the time facilitating among our colleagues, an open discussion of the potential benefits or potential problems that technology offers for those interested in increasing their evaluation capacity.

Roundtable B (First Rotation): Examining Impact
Presenters: *Coming to Term: The Evolution of a National Evaluation Strategy*, Judy Watson, Health Canada and Gail V
            Barrington, Barrington Research Group Inc

The Canada Prenatal Nutrition Program (CPNP), which is funded by Health Canada in collaboration with a variety of community partners, developed an evaluation strategy more than five years ago. This paper provides an overview of that strategy and then focuses on the impacts which have resulted from the implementation of the evaluation. The community development process employed in the design has resulted in outstanding participation rates of and the development of a large database of information about high risk women and their infants which is now being accessed by many stakeholders. It has build capacity both in the field and the national office, attracted significant additional funding from a number of sources, forged links with other government departments and non-government organizations, influenced program practice and provided a benchmark for other large-scale evaluations.

Building Evaluation Capacity in Brazilian Programs for Children and Youth at Risk: Impact of an Impact Evaluation,
      Thereza Penna-Firme , CESGRANRIO Foundation; Vathsala I Stone, New York University at Buffalo; and Juan
      A Tijiboy, Federal University of Rio Grande do Sul

The paper describes how an impact evaluation study funded by UNICEF triggered an ongoing capacity building effort in Brazilian programs for Children and Youth At Risk. It presents the innovative evaluation methodology the authors developed as part of this study as well as its repercussions in Brazil. By virtue of meeting the program evaluation standards (Joint Committee, 1994) of utility, feasibility, propriety and accuracy, this methodology held an appeal to its stakeholders, among them the policy makers (UNICEF) and the program participants (coordinators and street educators). UNICEF commissioned a capacity building project for these programs in distance learning mode, for which the authors transformed their new methodology into seven training modules. The paper describes these modules, their non-traditional approach and accessible style, and also their use in various community programs in Brazil. The paper also presents a meta-evaluative analysis of the impact study, conducted to enhance its value for stakeholders.

Roundtable C (First Rotation): Tapping into Untapped Resources for Capacity Building
Presenters: *Increasing Evaluation Capacity: Using Retired Teachers as a Resource in Educational Evaluations*, Patricia A Muller, Kim K Metcalf & Nicole M Stacey, Indiana Center for Evaluation

Securing adequate resources is a common problem in educational evaluation. This lack of capacity is exacerbated in evaluations that apply ethnographic or qualitative approaches requiring data collection through personal contact and visitation. The unique needs of a diverse range of evaluations led us to develop a substantial cadre of well-trained, flexible, and available data collection personnel consisting of retired teachers. These individuals provide relatively inexpensive, highly dependable personnel for on-site data collection, including: (1) proctoring standardized achievement tests, (2) conducting on-site interviews, focus groups, and observations, and (3) assisting in longitudinal case studies. When provided with proper training and support, these individuals are an efficacious, cost-effective, commited addition to projects. Advantages include: availability and flexibility of schedule, education level, understanding of school environments, and experience with children, teachers, and administrators. This session overviews approaches to identifying, recruiting, preparing, and using these individuals for large-scale and complex evaluations.

Roundtable D (First Rotation): Building the Profession by Increasing the Presence of Evaluators of Color
Presenters: *Increasing Evaluation Capacity by Increasing Evaluation Diversity: Expanding the Pool, Presence, and Participation of Evaluators of Color*, David A Sul, University of California at Berkeley

The evaluation profession, like many others, is in the process of coming to terms with issues of diversity. A growing number of evaluators are concerned with questions such as who conducts evaluations and the impact of evaluations on communities of color. A strategic effort to increase the proportion of evaluators of color first will require a structured plan. Further, nurturing students of color in evaluation and evaluation-related fields as they pursue evaluation careers should adhere to this plan. Rather than emphasizing a singular short-term approach for developing a pool of diverse evaluators, a multi-stage, long-term strategy for increasing the diversity of the evaluation profession is presented. Included is a discussion concerning the ownership of the problem, responsibility for keeping diversity issues at the forefront of evaluations, and the proper placement of diversity issues within the evaluation profession.

Strategies to Increase Evaluation Capacity and Engagement among Students of Color, Lea Williams, Columbia University

This paper discusses four important strategies to help create a professional environment that is attractive, supportive and validating to students of color. These strategies are:
1) Recruitment: establish a marketing campaign that is inclusive of students of color;
2) Preparation: give students of color the preparation and skills they need to build a knowledge base--this includes book knowledge and practical knowledge;
3) Supports: identify networks of people (i.e., trailblazers and peers) and create financial means so students of color can maneuver through obstacles and periods of transition successfully; and
4) Sustainment: maintain a clear infrastructure (i.e., professional development and career ladders) that promotes feelings of sustainment among students and professionals of color in the field of evaluation.

These four strategies will help increase evaluation capacity among students of color and within the evaluation profession as a whole.

Roundtable E (First Rotation): Increasing Evaluation Capacity through Stakeholder-based Evaluation

Presenters: *Increasing Evaluation Capacity for HIV Supportive Housing Providers: Using a Peer Evaluation Model to Assess Standards of Care*, Rachel A Kogan & Stephen Fleischer, AIDS Housing Corporation

Using a peer model for evaluating compliance with standards of care for HIV supportive housing programs represents the culmination of years of collaborative work among stakeholders as varied as the housing providers, the Massachusetts Department of Public Health (MDPH), residents of housing programs, and AIDS Housing Corporation (AHC), a local technical assistance provider. The peer model provides a unique, collegial framework for evaluation practice: evaluation teams are comprised of program residents, staff, peer evaluators, and a facilitator; over two months, the group is empowered to evaluate the housing program's compliance with standards they helped create. Animated discussions within evaluation teams are full of innovative, creative solutions for providing best practices in HIV supportive housing. The result? An evaluation process that providers truly own, and a practical method of assessing contract compliance for MDPH. This presentation will describe the process of developing a peer evaluation system, the benefits and challenges, and lessons learned.

Enhancing Capacity for Self-evaluation: Structural and Procedural Considerations, Charles L Usher & Judith B Wildfire, University of North Carolina at Chapel Hill

Many evaluators now view their professional role as encompassing a responsibility to help their clients enhance their capacity to do evaluation and to use evaluative information.   Too often, however, our conception of evaluation as a time-limited event--a "project," an "engagement," or a "report"--compromises our efforts in building capacity. Longer term success (i.e., well after the evaluator ceases to be involved) requires efforts to put into place structures and processes that are (1) responsive to critical information needs and (2) embody principles of self-evaluation and accountability. This presentation will summarize efforts to build state and local self-evaluation teams in the child welfare systems of ten states and some of the nation's largest cities. It will address issues of membership and ownership, the forms of data found to be useful, and the flow of information among participants. The presentation will draw on experience in several initiatives sponsored by national foundations.

Roundtable A (Second Rotation): Capacity Building for Evaluating Justice Programs

Presenters: *Increasing Evaluation Capacity with a Justice Data Warehouse*, Suzanne B Riles, Multnomah County Health Department

How can evaluators get better cross-agency justice data to learn 'what works' to stop criminal behavior? In 1999, the Public Safety Coordinating Council of Multnomah County in Portland, Oregon built a data warehouse and decision support system to access justice data from multiple sources. This data warehouse effort is both information technology for storing related databases, and a dynamic collaboration of attentive data proprietors. This presentation will 1) describe history and issues in the Multnomah County collaboration, 2) demonstrate the structure of the decision support system built to access the justice data warehouse, and 3) share results of efforts made by evaluators to use the new system for program evaluation.

Capacity-building in the Community Justice Setting: The Role of Program Evaluation in Portland, Oregon, Scott S Keir & Charlene Rhyne, Department of Community Justice of Multnomah County

The presentation will present key steps to increasing program evaluation capacity using the Multnomah County (Portland, OR) Department of Community Justice (DCJ) as a case study. The 'evolution' of the relationship between DCJ management teams, Program Evaluation (PE) Unit and the Information Services (IS) Units will be explored. The department-wide data collection systems of the juvenile side (Juvenile Information Network - JIN) and the adult side (Supervision Program Information Network - SPIN), will be highlighted as tools that permit the collection of and reporting on key indicators of departmental performance. Several reports that were developed with key input from the Program Evaluation unit will be presented and a discussion of how they are utilized by the department will be

addressed.

Roundtable B (Second Rotation): Working Together for Evaluation Capacity Building
Presenters: *Mutual Evaluation Capacity Building: An Evolving Model*, Pauline E Brooks, Cassandra Malry, Wendy Chen & Tenzing Donyo, The California Endowment

This paper presents an experimental model used by a large foundation that funds health programs for traditionally underserved and multi-cultural populations. This model is one mechanism toward developing more integrated, collaborative, services for grantees and prospective grantees in underserved and culturally diverse communities. Foundation program staff and foundation evaluation staff work together in pairs, as one team, to provide technical assistance both on- and off-site to grantees and prospective grantees. During such sessions, program staff work on the programming capacity of grantees. Coordinating with the program partner, evaluation staff simultaneously integrate evaluation training throughout the session. This model yields several benefits: Evaluation, program, and grantees learn together from each other, in real time. The capacity-building discussion/training stays grounded and focused. These three-way interactions promote clarity, common understandings within a real context, and the integration of evaluation thinking into the grantee's program.

Enhancing Evaluation Capacity through Collaboration among Service Providers, Phyllis M Thomas, Mid-continent Research for Education and Learning

For several years, The Eisenhower Regional Consortia for Mathematics and Science Education and Eisenhower National Clearinghouse have been collaborating on evaluation in order to provide a national picture of Consortia work. At the same time, participating organizations have also benefitted from enhanced capacity for conducting evaluations and enhanced support for the evaluator's work. This presentation will describe the nature of the collaboration (i.e.,exchanging information, altering activities, sharing resources, enhancing capacity for mutual benefit to achieve a common purpose (Himmelman, 1994) as well as the benefits, challenges, and lessons learned from this experience.

Roundtable C (Second Rotation): Beyond Our Borders: Expanding the Field
Presenters: *Promoting Greater Evaluation Capacity within a Global Business Environment*, Jerry F Luebke, Thomas O Schrader & N Jane Moeller, Arthur Andersen

Patton (1996) argues that evaluation can serve many purposes including, but not limited to, formative and summative evaluation. Within Patton's framework of evaluation practice, this paper presents three case studies that illustrate how different evaluation practices have been used to support various organizational initiatives within a global professional services firm. The first study describes a training evaluation effort of an internal executive development program and how developmental evaluation information has been used by the program's decision-makers. The second study describes how evaluation practice was used for knowledge generating efforts within the context of a large-scale internal change management project. The third study describes a developmental evaluation approach that was used in helping the organization's HR leadership to make decisions as they worked to improve the organization's career development program as well as employee satisfaction and retention. An overall synthesis of the case studies will be provided and related to promoting evaluation capacity within a business environment.

Roundtable D (Second Rotation): In the Beginning: Starting Out in the Field of Evaluation
Presenters: *The Novice Evaluator*, Ronald F Dugan & Dianna L Newman, State University of New York at Albany

As graduate students, beginning evaluators are often assigned to projects having little or no experience in the practical application of evaluation. This unsettling experience lasts only for a short time, however, when the graduate student realizes that the stakeholders from the projects he or she is evaluating are often as naive to the intricacies of evaluation as the graduate student is. The success of such an endeavor depends on several factors including: expert guidance from an experienced professional evaluator, willingness of the graduate student to learn not only from the expert, but from the stakeholders as well, and a commitment on the part of stakeholders to learning about and accepting the validity of evaluation. With these factors present in an evaluation by a novice graduate student, the end result is a rewarding experience for the graduate student, and a relevant evaluation for the stakeholder to utilize to its

fullest capacity.

Attracting and Keeping Students of Color in the Evaluation Profession: Ideas from Other Fields, Vasuki Rethinam &
    Miriam Mulsow, Texas Tech University

Diversity needs to be considered in planning, developing, implementing, and evaluating culturally competent programs. Although we need more evaluators who fully understand various cultures, there is dearth of research on minority recruitment to the field of program evaluation. Published studies focus primarily on ways to prepare women and minorities for educational research and leadership positions. Based on a literature review from other fields in which minority recruitment has been implemented and Rethinam's personal experiences, the authors will discuss ways to increase participation of minority students in program evaluation. Strategies include ways to:
- increase awareness of program evaluation courses and what program evaluation is.
- encourage departments to require program evaluation classes.
- provide examples of successful evaluators of color in educational materials.
- include cultural components in other courses.
- use focus groups to develop culturally competent courses.
- design selection criteria and recruitment procedures to reach minority students

Roundtable E (Second Rotation): Increasing Capacity To Evaluate Community Development: What Are We Going to Measure?
Presenters: Susan S Jakes, North Carolina State University
    Sherry C Betts, University of Arizona
    Lisa Cassidy, North Carolina State University

The Community outcome area of the multi-state Children, Youth, and Families At-Risk Evaluation Collaboration has been given the task of creating evaluation resources to evaluate community development efforts to benefit children, youth, and families. This group has struggled with the difficult question of how to increase the capacity of local programmers to evaluate the community development efforts before there are measurable results for children, youth, and families. The issue that has surfaced most persistently during this work will be the subject of the proposed think tank:: "In the evaluation of community development efforts, there is frequently confusion regarding WHAT is to be evaluated: structural, process or social/ cultural changes." This session will divide participants into groups to discuss current community development frameworks and propose possible revisions or alternatives. The session will also create a participant list that evaluators can use in networking around these issues.

SESSION 450: Panel   **Room:** Oahu
Sponsored by the Topical Interest Group on Evaluation Use
Promoting Evaluation Capacity within Organizations: Lessons Learned from Internal Evaluation Efforts
Chair: Jennifer W Martineau, Center for Creative Leadership

Evaluation, when conducted by an "objective observer", pits one "side" against "another" in the quest for continuous improvement of products, programs, and services. Without malicious intent to do so, evaluation creates a group that is "evaluating" and a group that is "evaluated". Therefore, it is often a difficult endeavor in almost every type of organization. This difficulty can be heightened when the "objective observer" is a member of the same organization - internal evaluation. However, internal evaluation is beneficial for many reasons, including costs, availability of staff, and existing knowledge of the initiative(s) being evaluated. The presentations in this session will demonstrate the importance and challenges of internal evaluation in a variety of settings: for-profit; not-for-profit human services; not-for-profit research and education; and not-for-profit education organizations. Panelists will discuss issues such as:
* challenges in implementing high-quality;
* effective internal evaluation studies;
* strategies for overcoming these challenges; and
* the use of internal evaluation research in ongoing improvement of initiatives.

Panelists: *Strategies for Promoting Evaluation Capacity in Social Service Programs*, Ann Minnett, Salesmanship Club Youth and
        Family Centers Inc

The Salesmanship Club Youth and Family Centers, Inc. conducts internal evaluation studies of our social services
with the goal of developing programs, refining practice, and sharing what we learn with others. The greatest challenge
in implementing the studies and learning from the results is gaining practitioner "buy in" -- a tricky task because
nonprofit sector professionals typically fail to appreciate evaluation's relevance to serving clients, and most are
unfamiliar with (if not intimidated by) evaluation methods. We have overcome these barriers to evaluation capacity by
collaborating with professionals to weave the evaluation process into their practices. Participants help design and
implement the studies, which helps them understand and use the results. This presentation will describe a model of
our collaborative process and provide examples of how we engage program staff to evaluate residential treatment,
family therapy, and restructuring a community school in an urban impoverished neighborhood.

Tales from the Front: The Highs and Lows of Internal Evaluation, Jennifer W Martineau, Center for Creative
        Leadership

While most training and development organizations utilize some type of evaluation methodology, many are faced with
a myriad of challenges to conducting well-integrated, high-quality evaluations on an ongoing basis. The reasons
include: lack of financial resources to fund evaluation research; small evaluation staff; trainer discomfort and distrust
of evaluations; and a sincere desire to avoid asking program participants to "take more tests". However, a training and
development organization cannot survive, nor thrive, without a clear picture of the effectiveness of its services.
Therefore, internal evaluation is a critical and necessary component of the work in these organizations, and the
responsibility lies with the internal evaluator to educate the organization regarding its importance and value. This
presentation will discuss:
* the benefits and pitfalls of internal evaluation
* approaches taken to gain acceptance, internally, for evaluation and its findings
* the organizational barriers to carrying out high-quality internal evaluation
* the strategies used to overcome these barriers

Internal Evaluation and Organizational Learning: Pitfalls and Possibilities, Rosalie T Torres, Developmental Studies
        Center

This presentation will discuss the challenges encountered and the successes experienced by an internal evaluation
unit explicitly charged with facilitating organizational learning in an educational non-profit. The presentation will begin
with a brief overview of the role of evaluative inquiry in organizational learning. To illustrate challenges and successes
the presenter will compare and contrast specific evaluation efforts across different projects within the organization.
This analysis will describe the background and mediating factors for each project. Further, viewing each evaluation
effort as a "problem set," the presentation will address interventions and next steps for furthering the goal of learning
within projects. The presentation will conclude with a discussion of how efforts to analyze across-project learning are
contributing to organizational learning.

Balancing Profits and Learning: Internal Evaluation in For-profit Organizations, Jennifer D Dewey, North Central
        Regional Educational Laboratory

For-profit organizations face a critical challenge to internal evaluation: "chargeability". In the quest to generate profits,
consultants must spend the majority of their hours on billable, project work whose purpose is to produce revenue for
the organization. This often leaves little time for the internal assessment of its own functions. This paper will examine
how internal evaluation has been carried out in a large, for-profit consulting firm within the context of:
* Opportunities for evaluation
* Who is designated to conduct internal evaluations
* Deciding priorities for evaluation
* How evaluation information is acquired and analyzed
* Effective methods of communicating evaluation results

* Turning results into action on the part of the organization
* The role of the evaluation consultant in the internal evaluation process
Balancing the profitability goal with the need for the internal assessment of organizational effectiveness, efficiency, economy, and performance and how that assessment can have a positive impact on organizations, is discussed.

Discussant: Arnold J Love, Independent Consultant

SESSION 451: Panel   Room: Waialua
Sponsored by the Topical Interest Group on Pre-K-12 Educational Evaluation
Measuring Educational Outcomes
Chair: Alan Ginsburg, US Department of Education
Panelists: *Estimating the Value Added of Vocational Education Using New State Data*, Marsha Silverberg, United States
              Department of Education

The National Assessment of Vocational Education (NAVE) is a Cngressionally-mandated evaluation of vocational/technical education in the United States. Many of the NAVE studies rely on national longitudinal surveys of students.   Unfortunately, these surveys are rarely current, and small sample sizes usually preclude detailed analysis by vocational program area or type, student course-taking patterns, or smaller subgroups (e.g., students with disabilities or limited English proficiency). To overcome these drawbacks, NAVE is undertaking a special study to rigorously estimate the value added of vocational education using student data from several states. States increasingly have electronic files with student information at the secondary and postsecondary levels that can be linked together and merged with data on labor market outcomes. These data have multiple cohorts, are longitudinal , and include test scores and demographic characteristics that allow seasoned researchers to account for potential selection effects.   Use of these state data systems will contribute to research on vocational education and to the field more generally.

Impact of Upward Bound on College Access and Other Outcomes, David Goodwin, US Department of Education

The Department of Education has conducted a multi-year longitudinal evaluation of the impact of the Upward Bound program using a nationally representative sample of 2800 eligible applicants who were randomly assigned to treatment and control groups.   A five year longitudinal survey has followed this sample throughout high school and the initial years of college. This study represents the first use by the U.S. Department of Education of an experimental design to evaluate a major ongoing operational program. Upward Bound Program is the largest federal program designed to help disadvantaged high school students prepare for and gain access to postsecondary education. Upward Bound projects, usually operated by a college, offer extensive academic instruction as well as counseling, mentoring, and other support services to high school students. Students meet throughout the school year and participate in an intensive residential summer program.

21st Century Community Learning Centers Evaluation: Challenges of Implementing Random Assignment, Amid Differences in Treatment Intensity and Duration, Elizabeth Warner, United States Department of Education

The 21st Century Community Learning Centers (CCLC) is a federally funded after school grant program. It was designed to meet an increasing need for quality, after-school programming for school-age children. These Centers provide academic enrichment and support; recreational, artistic, sports, and cultural activities; opportunities for children to participate in community service projects; nutritional and health services; access to technology and telecommunications; and they promote parent involvement and lifelong learning that can directly and indirectly benefit their children. A national evaluation of the 21st CCLC program, begun in 1999, will assess student impacts by using a random assignment design in elementary school sites and a matched comparison design in middle school sites. Outcomes range from in-school academic achievement, behavior, and to out-of-school behavior and safety. Data collection will include surveys of students, their parents and teachers, and collection of school and program records for each student in the sample. In addition, a pre- and post- reading achievement test will be administered to elementary school students in grades 3 through 5. Aside from addressing the usual challenges of implementing random assignment, this evaluation will also face differences in treatment intensity and duration across program

participants.

The Longitudinal Evaluation of School Change and Performance, Daphne Hardcastle, United States Department of
Education

The Longitudinal Evaluation of School Change and Performance (LESCP) is designed to examine the impact of
standards-based reform on student achievement in 71 high-poverty Title I elementary schools. The evaluation has
collected information on classroom practices, curriculum content, and professional development as well as schoolwide
reforms and parent involvement. An independent assessment was administered to both a longitudinal and cross-
sectional cohort of students over 3 school years. Achievement results on State/locally-required assessments will also
be analyzed. Building on its preliminary findings, the study is currently attempting to determine: how teachers
translated their understanding of standards, curriculum, and assessment into their classroom practice; what factors
led to ineffective classroom practices; and what combinations of state and district policies, school conditions, and
classroom practices were associated with greater or lesser student gains.

SESSION 452: Meta-demonstration   **Room:** Waianae
Sponsored by the Topical Interest Group on Cluster, Multi-site and Multi-level Evaluation
The Application of Innovative New Technologies and Methods to Cluster and Multi-site Evaluation
Chair   Jeffrey Greene, Third Sector Strategies
Presenters: *Using Communication Network Analysis as an Evaluation Tool*, R Sam Larson, Applied Research

Many large-scale funding initiatives involving multiple grantees include conferences or meetings where grantees are
convened by the funder. One purpose of such meetings is for grantees to share new ideas and information. As
evaluators for such initiatives, how can we determine if communication networks are developing among grantees?
Based on an evaluation involving 20 institutions and more than 60 participants, this paper illustrates how
communication network analysis techniques can measure and display (a) the emergence and stabilization of
networks, (b) the centrality of the actors in the network, and (c) the differences in interactions among and between
groups. This paper discusses procedures used to collect data from respondents. It presents findings using network
pictures and time series analysis graphs. It concludes with suggestions on how to use these findings to inform the
content of future meetings and to illustrate to board member the outcomes of network conferences.

Technical Assistance, Training and Technology: Tools for Building Evaluation Capacity, Healther L MacGillivary, The
Colorado Trust and Jim Adams-Berger, OMNI Institute

Prevention Evaluation Partners (PEP) is an evaluation system developed on the assumption that training and
technical assistance will enable local service providers to build their capacity for self evaluation. Over the past five
year, PEP has been developed and implemented by OMNI Institute for Colorado's Alcohol and Drug Abuse Division.
PEP includes a computerized management information system for tracking services, a CD-ROM software program to
assist providers in designing and conducting outcome evaluation, and a narrative prevention stories component
describing program successes. This paper will discuss the training and technical assistance necessary to implement
a technologically advanced evaluation system for over 100 programs. The challenges, strengths, and future
directions of this system will also be shared.

Evaluating Quality of Life Programs in the US Navy: Comparing Dissimilar Programs on a Common Metric, Michael J
Schwerin, Zannette Uriell & Paul Michael, Navy Personnel Research Studies and Technology

Quality of Life (QOL) is a complex concept and one that represents a significant investment by the U.S. Navy.
Decision-makers need data that will help them determine if this broad spectrum of QOL programs are meeting their
objectives and having a positive impact on QOL and retention. Program evaluation surveys for each of 12 QOL
programs were constructed, focusing on program Reasons for Being -- statements that headquarters-level QOL
program managers feel reflect the primary intended purpose(s) of QOL programs. Data were collected in three Navy
Fleet concentration areas: Hampton Roads Virginia, San Diego California, and Yokosuka Japan. A composite
program rating (mean score of the program objective measures and participants overall rating of program quality) was

formed. The relationship between program impact (as measured by the composite program rating) on global QOL and impact on retention intent was examined. Findings as well as advantages/disadvantages of this approach will be discussed.

Discussant: Astrid Hendricks-Smith, WK Kellogg Foundation

SESSION 453: Panel   **Room:** Molokai
Sponsored by the Topical Interest Group on Alcohol, Drug Abuse, and Mental Health
Persistent Effects of Substance Abuse Treatment Studies
Chair: Candace Macken, Center for Substance Abuse Treatment

PETS (Persistent Effects of Treatment Studies) is a coordinated family of studies funded by the Center for Substance Abuse Treatment (CSAT) to assess the longer-term outcome course of people who receive substance abuse treatment. Underlying PETS is a recognition that substance abuse disorders, while variable in their manifestations, are often chronic and prone to relapse. PETS studies in Cuyahoga County, Ohio, and Chicago, Illinois, were developed by extending ongoing shorter-term research projects. These studies are tracking adult samples (1,200 to 1,400 clients) for 3 to 5 years, and will use of administrative information system data to supplement client interviews. PETS staff are also conducting secondary analysis of previously collected longer-term outcome data, and are developing a study of adolescents. This 5-session panel will present early findings from the PETS studies and from analysis of previously collected longitudinal data. The papers will illustrate analytic methods appropriate for longer-term outcome studies.

Panelists: *Predicting the Extent of Treatment Involvement by Substance Abusers: A Two-year Longitudinal Analysis*, Peter J
             Leahy, Richard C Stephens & Shoba Nair, University of Akron

The extent of treatment involvement among 750 substance abuse clients in Cleveland, Ohio is examined. Data were collected on the Central Intake Assessment Instrument - Cleveland, a self-report instrument measuring substance abuse, demographics, family issues, criminal involvement, housing, mental health and social services. Re-interviews occurred at 6-, 12- and 24-months post baseline. Client treatment histories are linked from the Alcohol and Drug Addiction Services Board of Cuyahoga County (Ohio) Management Information System. The two-year length of the follow-up design and the inclusion of MIS data make this study unique in the substance abuse field. Independent variables include type and extent of substance use, previous treatment experience, level of care and relevant demographics. Multivariate techniques are used to construct the level of treatment received. Hierarchical linear modeling is used to predict the extent of treatment involvement.

Longitudinal Treatment Effects among Cocaine Users: A Growth Curve Modeling Approach, Yih-Ing Hser & M
             Douglas Anglin, University of California at Los Angeles; and Chih-Ping Chou, University of Southern California

This study examined longitudinal treatment effects among cocaine users. Face-to-face interviews were conducted with 371 cocaine users screened from arrestees in jails and from patients in sexually transmitted disease clinics and emergency rooms. Using a natural history instrument, drug use and other behaviors were retrospectively recalled in chronological order starting one year before the initiation of drug use until the time of interview. A series of growth curve models were applied. Results showed that use of cocaine and other drugs (heroin, marijuana) and criminal activities deceased dramatically among 250 subjects upon their entry into treatment, while no such changes were observed among the 121 subjects who had never been in treatment. However, drug use and criminal involvement gradually increased as patients dropped out of the treatment. Modeling results confirmed treatment effects in reducing relapse to cocaine among treated subjects across time (odds ratios ranged 1.94 to 4.85).

Impact of Continuing Care Services on Substance Use Outcomes in a Two-year Follow-up Study, James R McKay,
             DeltaMetrics and Peter J Leahy, University of Akron

This paper examines the effectiveness of continuing care, or "aftercare," in 750 substance abuse clients followed for 24 months post intake to treatment. All clients were treated at one of 20 participating agencies of the Alcohol and Drug

Addiction Services Board of Cuyahoga County. Services data were obtained from the county MIS and outcome data were collected at 6-, 12-, and 24-month follow-ups. "Continuing care" was defined as outpatient treatment received after a more intensive initial phase of care. Data analyses will (1) number of individual, group, medical, family, and crisis management continuing care services received by various client groups; (2) relative effectiveness of these services; (3) relative impact of formal continuing care treatment, self-help participation, and supportive housing; (4) optimal combinations of primary and continuing care interventions; and (5) identification of performance indicators during the initial phase of care that predict response to continuing care.

Cumulative Effect of Substance Abuse Treatments on Long-term Outcomes, Christy K Scott, Mark A Foss & Michael L Dennis, Chestnut Health Systems Inc

Prior treatment experience has been found to be related to increased retention in the index treatment episode, an increased likelihood of readmission following the indexed episode, and reduction in substance use (Marsh et al., 1990; Hser et al., 1997, 1998). This study examines the relationship between the number of and total length of time in prior treatment and (1) retention in treatment, (2) number and length of treatment readmissions, and (3) change in participants' intermediate and long-term drug use, while controlling for participant characteristics, pattern of drug use prior to intake, and the index treatment episode's level-of-care. Intermediate (6-months) and longer-term (24-months) post-intake outcomes on 1,328 participants across six levels-of-care are examined. Measures of treatment experience between intake and the 6-month follow-up are also included as a predictor of longer-term outcomes. The outcome variables are analyzed separately using a generalized linear model.

Counselor Effectiveness and Substance Abuse Treatment Outcomes, Garrett E Moran, Stephen Messer & Bill Luckey, Westat; and Richard C Stephens, University of Akron

Previous studies have rarely examined factors associated with counselor effectiveness in the field of substance abuse. The purpose of this study is to identify and characterize "effective" substance abuse counselors, defined on the basis of client outcomes. We will examine counselor attributes associated with effectiveness, accounting for baseline client differences and subsequent treatment episodes. The 750 clients in the study were interviewed at baseline, 6-, 12-, and 24 months in Cuyahoga County, Ohio. The County's drug and alcohol information system will complement client interviews by providing data on counselor characteristics and client-counselor pairings. Client change over the two-year period will be computed using latent trajectory analysis (LTA). LTA provides a slope parameter indicating the amount and rate of change in an outcome (here, substance use). To analyze the relationships between counselor/client characteristics and effective counselor status, a series of simultaneous multiple logistic regression analyses will be conducted.

SESSION 454: MultiPaper   **Room:** Kauai
Reconsidering the Role of the Evaluator I
(The second half of the session will be held in the same room in the next time slot.)
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: *Toward a Classification of Different Evaluator Roles*, Melvin M Mark, Pennsylvania State University

This paper describes an interim attempt to classify previously identified evaluator roles in terms of a small set of central factors. Evaluator roles may differ, for instance, based on the preferred function/purpose of the evaluation, the organizational relationship between the evaluator and the evaluated, and assumptions about objectivity/advocacy. Use of a simple (and admittedly incomplete and imperfect) classification system provides a launching point for addressing some interesting questions: What are the theoretically possible evaluation roles that have generally been ignored, and why? Do certain clusters of roles naturally fit together, and represent different styles, subspecialties, or domains of practice? The literature on self, a close cousin of the concept of identity, which is also addressed in this session, includes the notion of self-complexity, which stimulates other interesting issues in relation to a classification system for evaluator roles: To what extent does evaluation role-complexity influence evaluators' identity? What are the advantages and disadvantages of role complexity, for the field and for an individual evaluator?

Revisiting the Role of the Evaluator, Lizanne DeStefano & Katherine Ryan, University of Illinois at Urbana-Champaign

The concept of the 'role of the evaluator' is central to the theory and practice of evaluation. Historically the evaluator's role was one of detachment, avoiding bias, with great faith placed in the methods selected for the evaluation (Weiss, 1998). Nevertheless, the nature of evaluation and the evaluator has shifted with the interpretive turn in social science. What are the implications of this shift? Weiss (1998) suggests the contructivist notion of there are new roles for the the evaluator: "critical friend, co-investigator, facilitator, and problem-solver"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. In this paper, we critically examine the construct of the role of the evaluator through an examination of historical and current evaluation theories. The questions to be addressed include the following. How has the role of the evaluator developed? How has this role shifted? Is the notion of role the best concept to capture who the evaluator is in evaluation theory, practice and society? In addition to the construct of evaluator identity, are there other alternative conceptualizations that might be considered?

Performing Evaluation, Norman K Denzin, University of Illinois at Urbana-Champaign

Evaluation, A Democratic Institution, Gary T Henry, Georgia State University

The recent performance turn in social and cultural theory (Turner, 1986; Conquergood, 1995) signals the death of traditional role theory (Biddle and Thomas, 1966) in the social sciences. . I examine the implications of the performance turn for the practice of evaluation. I then outline a performance-based model of evaluation practice. The negative consequences of the performance turn for role theory are summarized by Giddens (1979). They involve three problems: (1) the perpetuation of an action/structure dualism; (2) an erroneous linkage of roles to models of normative consensus;   (3) an over-emphasis on the primacy of the normative in social theory, thereby ignoring the performative features of social life. If social life is a situated performance, then a performance-based theory of evaluation practice examines how stakeholders and evaluators enact gendered cultural meanings in the evaluation arena. The meanings of evaluation performances are constantly changing, every performance is locally situated. Stories and narratives grow out of these performances. Performance based evaluation practice collects and performs these stories. The consequences of this commitment for evaluation theory are discussed.