Thursday, November 2, 12:00 Noon to 1:30 P.M.

Sponsored by the Topical Interest Group on Minority Issues in Evaluation
Building Evaluation and Research Capacity in Traditionally Underserved Communities
Chair: Susan J Rogers, Academy for Educational Development

The capacity of a community to evaluate their initiatives has historically been linked to the presence of researchers and evaluators in that community and the resources to carry out research/evaluation activities. As a result, several communities have traditionally been left out of the evaluation learning process, and, in fact, often served as target populations and exploited subjects of research and evaluation. This has led not only to a cultivated distrust by these communities toward research and evaluation but the disempowerment of people to effectively understand, learn from and document their own progress. The present panel will make presentations on several projects that are being conducted to confront this issue and build evaluation/research capacity among traditionally underserved communities. Evaluation training and the development of an evaluation resource database to serve underserved communities in the Mid-South Delta will be reviewed along with "Institute Without Walls," a program that provides curriculum and vehicles for the disadvantaged to pursue careers in evaluation and a Devolution Initiative to promote public participation in informing policy agendas and decisions. In addition, a presentation will be made regarding AEA's proposal for "Building Diversity Among the Evaluation Community."

Panelists: *Scholar/Practitioner Teams on the Devolution Initiative*, Ricardo A Millett, WK Kellogg Foundation

The Devolution Initiative is a program supported by the W.K. Kellogg Foundation that focuses on healthcare reform and welfare reform and the shift of federal responsibilities to the states. The Initiative recognizes that the success of such reform is based on an informed citizenry and the importance of empowering communities at the grassroots level to inform and implement policy. In order to foster this empowerment, the Initiative utilizes scholar/practitioner teams of mentors and researchers in the targeted mobilization states. They work to translate and make more useful national research data and information to support efforts to inform policy within a state context. They also generate new information from the community level to inform the debate around state/local interpretations of how best to effectively and efficiently implement welfare reform policies. In general the teams: 1) produce state and local research related to devolution to be used by the Initiative's state lead organization and their community organizations; 2) help communities translate and interpret research and evaluation findings that can be used to enhance the policy decision-making process; and 3) assist in creating connections and dialogues among community organizations, researchers, evaluators and policy analysts to support the state work.

The Institute Without Walls (IWW), Astrid Hendricks-Smith, WK Kellogg Foundation

Evaluations are generally conducted by university institutes, not-for-profit "think tanks", and consulting firms. Professional evaluators typically acquire a Masters, and often a Ph.D. from a major institution. Few from economically disadvantaged backgrounds have historically had access to such a career path. As a result, the market for evaluation professionals lacks diversity. Investments in "systems change" are often made in economically disadvantaged communities that include economically, culturally, and ethnically diverse populations. Evaluation professionals with experience, sensitivity, and knowledge about diverse populations could enrich the explanatory power and improve the rigor and relevance of those evaluations. The IWW is a program, sponsored by the W.K. Kellogg Foundation that supports Community Voices: Health Care for the Underserved Initiative, that strives to create model curriculum and institutional vehicles for people from disadvantaged backgrounds to pursue a career in evaluation. This presentation will describe the philosophy and structure of the IWW as well as the early implementation.

AEA's Initiative for Building Diversity among the Evaluation Community, Edith P Thomas, United States Agriculture Department; Donna M Mertens, Gallaudet University; and David M Chavis, Association for the Study and Development of Community

The purpose of the Initiative for Building Diversity is to improve the quality and effectiveness of evaluation by increasing the number and ethnically diverse evaluators in the evaluation profession and improving the capacity of evaluators to work across cultures. It builds upon prior efforts of the AEA including the development and work of a Diversity Committee, a Minority Issues Topical Interest Group (TIG), a Membership Recruitment and Retention Task Force, and efforts to establish a mentoring program. The present Initiative, however, recognizes that further work is required, particularly work that will be transformative of the organization. Although the presenters of the Building Diversity Initiative will explain the two phases of the initiative (to first develop a concrete plan and next, implement the plan) they will largely focus on the first phase of plan development. They will discuss the defined objectives for phase 1, the management and staffing structure for the initiative, the proposed tasks and deliverables, and potential funders of Phase 1.

Building Evaluation Capacity in the Mid-south Delta, Laurine Thomas, Susan J Rogers & Elvis Fraser, Academy for Educational Development

The Mid-South Delta Initiative (MSDI) is an economic empowerment project funded by the WK Kellogg Foundation seeking to impact a 55 county area of Arkansas, Mississippi and Louisiana. Based on the results of a needs assessment conducted in the area, it was determined that evaluation capacity building was a priority to allow for both an effective assessment of Kellogg-funded initiatives in the area as well as to overall foster evaluation expertise in the Delta. To assist with this need, the Academy for Educational Development (AED) was contracted to identify and work with local evaluation expertise. Two products that have been developed to further this end is an evaluation training and a MSDI Research, Evaluation, and Technical Assistance Database. The training covers the topics of participatory evaluation and the logic model and the database contains information on over 50 technical assistance providers in the Delta with a scope of expertise in research and evaluation. This presentation will review the overheads developed for the training, the steps that were taken to develop the database and review how the database information is being used by a local MSDI evaluation committee in the Delta.

Discussant: Rodney K Hopson, Duquesne University

SESSION 238: MultiPaper   Room: Akaka

Sponsored by the Topical Interest Group on Feminist Issues in Evaluation
On Epistemology, Caring and Connection
Chair: Kathryn A Bowen, Bowen's Evaluation and Consulting Services
Presenters: *The Human Element of Evaluation*, Michal Galin, Independent Consultant and Denice A Cassaro, Cornell University

While discussions of evaluation practice are important and serve to better define the act and field of evaluation, an element of ANY evaluation that is more ambiguous and less frequently discussed is the human element. Even as evaluators may define themselves with descriptors such as post-positivist, constructivist, or feminist, important questions of who the evaluator is, what values are held by the evaluator, how the evaluator's identity(ies) or perceived identity(ies) interact and affect evaluation settings, and how these facets of any evaluator impact decisions made in terms of methodology, methods, and the interpretation/analysis of evaluation results. While it is important for an evaluator to be self-reflective and self-aware, more discussion is needed to explore and find ways to implement strategies addressing the human element in evaluation. Our discussion will focus on some of the issues that can arise and strategies that incorporate us as human beings in the evaluation process.

Thinking Outside the Box: A Feminist Re-evaluation of Rape Resistance Strategies, Patricia D Rozee, California State University at Long Beach

The paper focuses on the expansion of social program evaluation though the use of qualitative data. The author will describe one feminist's journey toward qualitative methods from the "box" of quantitative training. The research discussed involves broadening the definition of rape resistance to include women's perceptions as detailed in narratives about their rape experience. Lessons shared include how to listen to women's voices without constructing pre-conceived (suryey-format) notions about their experience, how to reconcile existing literature with what the women are telling you, and the reconstruction of common notions of resistance.

A Feminist Perspective on Assessment: An Epistemology of Caring and Connection, Pamela A Sandoval & Stacey
    Gray, Indiana University Northwest

This presentation discusses a feminist perspective on institutional and student assessment through the nine
principals of feminist assessment and the nine principles of Good Practice of Assessing Student Learning (American
Association for Higher Education).   We suggest that the education of all students will improve through connected and
caring assessment in contrast to sorting based on classroom and standardized tests. Caring and connection means
that students are encouraged and connected to others through improvement with a focus as found in feminist models
of assessment. Whether at the institutional or individual level, a feminist perspective on assessment illuminates
success through a process of caring that includes an interactive dialogue between the educator and the student, and
between the educator and other educators that illuminates what is done well as well as what needs improvement.

Discursive Evaluation: The Journey of an "Epistemorph" Toward an Interpretive Practice and Inquiry, Cynthia A
    Tananis, University of Pittsburgh

This paper focuses on the portrayal of my journey as an "epistemorph" in evaluation practice and associated inquiry.
As an "epistemorph" I lay out the key differences in moral and cognitive choices implicit in evaluation practice and
inquiry settings by exploring what I refer to as the "ologies" --- what I believe to be true (epistemology), valuable
(axiology) and to be real (ontology). The paper examines the positivist/functionalist historical and philosophical
framing of evaluation and offers a critique from a more interpretivist worldview that underpins a variety of more
currently suggested evaluation approaches. I present each perspective, and its implications in practice and inquiry,
through the heuristic of personas.

Discussant: Thomas A Schwandt, University of Illinois at Urbana-Champaign

---

SESSION 239: MultiPaper   **Room:** Koko

Sponsored by the Topical Interest Groups on Research Technology and Development Evaluation & State and Local Government
Issues in Evaluating Research Activities and Organizations
Chair: Irwin Feller, Pennsylvania State University
Presenters: *Research Evaluation: Controlling or Developing the Knowledge Production*, Finn Hansson, Copenhagen Business
              School

Research evaluation has - at least in Western Europe - for the last 10-15 years been used as a instrument in New
Public Management to control and steer public research and researchers. The perspective on research evaluation has
been retrospective and predominantly focused on research productivity as seen in publications and citations counts.
In a time with rapid changes in the organization of science (mode 2 science, triple helix) the traditional concept of
research evaluation is becoming outdated. This paper argues that research evaluation ought to be used as a
developing tool in the research organization - the concept of participatory evaluation seems to be most relevant to
facilitate knowledge production. The paper will present some preliminary results from a large study of research
management in Denmark of the use different types of research evaluations in private and public research
organizations.

Assessing Innovation in Research Organizations: A Comprehensive Framework , Gretchen B Jordan, Sandia National
    Laboratories

Innovation is one of the expected practices of a research organization as well as an expected outcome. In the last 25
years, many authors have described the aspects that they believe are the most important to classifying and evaluating
research organizations. This paper summarizes some of the work that has been done by experts in the fields of
science and technology management and organizational theory, and consolidates these views into a broad framework
for assessing research organizations. This new framework includes attributes internal to an organization such as the
qualities of individual researchers, resources, and characteristics of the organization. It includes defining the purpose
and products of the organization and the stage of innovation of the technologies involved. In addition, it incorporates

aspects of the external environment that also affect the organization's performance.

From Concept to Reality: Moving the Yardsticks on Performance Measurement – Key Roles for Evaluators, Robert Lahey, Public Service Commission of Canada; and Steve Montague & Jennifer Birch-Jones, Performance Management Network Inc

The raised level of consciousness for improved performance measurement has derived in large part on legislative requirements and/or explicit requirements to report on 'outcomes' achieved. Government and non-government agencies alike have devoted increasing attention and resources to the tasks of measuring and monitoring 'outcomes'. This has not however always translated into more effective performance assessment and reporting – the meeting or expectations, to date is quite mixed. The need to improve results-based performance measurement has raised many challenges in the government sphere: orienting the management culture to an outcomes approach rather than simply the traditional activity approach to measurement; methodological issues around attribution, i.e. measuring the extent to which the activities and outputs of a program have contributed to the outcomes achieved; and systems issues around the most cost-effective and feasible approach to gather needed data for monitoring and reporting purposes. Added conceptual and measurement challenges are introduced when addressing areas of policy research. The presenters will discuss what value-added role evaluators can contribute to their organization, building on their experience from the federal government of Canada, with particular examples from the policy and policy research area.

Improving Research Proposal Evaluation Processes by Modeling the Underlying Logic, Michael Obrecht, Medical Research Council of Canada

Most studies of the effectiveness of systems for assessing research grant applications do not question system logic. That is, analysts rarely examine the extent to which the desired outcome of the assessment process is clear and understood by all reviewers. Likewise, few authors examine the critical relationship between desired outcome and the criteria used for the selection, or the extent to which criteria have been defined, weighted and scaled. It may well be that thick fog is assumed to be land. Fairness is the usual focus in studies of systems for reviewing research proposals. The principal question is whether irrelevant variables influence the assessment process. Age, discipline, gender, language, prestige, race, affiliation and experience, may all be considered with respect to both applicant and reviewer. Systems must be fair, and studies of equity are essential. But unless the underlying logic is also examined, the likelihood of making significant improvements to assessment systems is greatly reduced.

SESSION 240: Panel   **Room:** Waimea

Sponsored by the Topical Interest Group on Teaching of Evaluation
Training Evaluators: What's Needed
Chair: Rita G O'Sullivan, University of North Carolina

The demand for well-trained evaluators is growing, and adequate preparation of evaluators is an important task facing the profession. Currently, vast number of individuals practice evaluation without any formal or even experiential training. At the same time, the nature of evaluation tasks can vary greatly. Those trained to perform one type of evaluation, may not be prepared to implement other types of evaluations. These different demands require that we train evaluators with different responsibilities in different ways. Panelists will share their experiences with evaluation training programs and identify existing gaps. Attendees will be encouraged to contribute to the session.

Panelists: *Promoting Collaborative Evaluation Training at the University of North Carolina*, Rita G O'Sullivan, University of North Carolina

Currently, evaluation training at the University of North Carolina occurs in two distinct ways. At the graduate level, master's and doctoral students may complete degrees with a specialization in educational program evaluation. In addition to formal evaluation training, a considerable amount of evaluation training occurs through collaborative evaluation projects that focus on enhancing the evaluation skills of project personnel. In both situations, the evaluation training that is provided barely begins to meet the apparent needs.

Development and Implementation of a Master's Degree Program in Project Management and Evaluation at the University of the West Indies, Farley Braithewaite, University of the West Indies

The Caribbean Territories over the years have depended heavily on donor agencies such as the Caribbean Development Bank and other international lending agencies for project funding. To this end, territories have had to depend to some extent, on exernal agents to assist in the management and evaluaiton of projects. The lack of well-trained Project Managers and Evaluators has signaled the need for a program in Project Mangement and Evaluation. The Faculty of Social Sciences at the University of the West Indies has recently responded to this need and developed a master's degree in Project Management and Evalaution. The purpose of this presentation is to share the structure of the master's degree program and report about training experiences.

Evaluation Training Needs in the United States: Where Are We and Where Do We Need to Go? Arlen R Gullickson, Western Michigan University

A number of centers around the United States are actively engaged in training evaluators to fill the growing demand. The need for training, however, far out-distances the availability of training opportunities. This presentation reflects on the current state of training for evaluators and suggests ways that evalution training might be delivered.

Discussant: Paul Morgan, Caribbean Development Bank

SESSION 241: MultiPaper **Room:** Niihau
Sponsored by the Topical Interest Group on Theories of Evaluation
Issues in Evaluation Theory and Practice
Chair: Barbara Lee, University of South Florida
Presenters: *Evaluation: Connecting and Strengthening Communities of Practice*, Frank E Davis, Lesley College

This paper discusses the role and practice of evaluation from the perspective of the concept of communities of practice. In general, a community of practice is a group of individuals who are mutually engaged in a joint enterprise through a shared and developing repertoire of skills and knowledge. A process of social learning sustains the practice of such communities. Using this idea the paper will examine the growing field of evaluation and its practice. The paper will identify different communities of practice that appear to define differences in evaluation theory and an practice, as well as suggesting a broader notions of practice which may provide a more unifying vision of evaluation theory and practice.

Evaluation Theory and Practice: Limitations on Their Reciprocal Influence, Nick L Smith, Syracuse University

Although there is general agreement in evaluation that theory should guide practice, and that practice should improve theories, it is not clear how these interactions are supposed to happen. From a careful analysis of both (a) the nature of contemporary evaluation theory as it is defined, constructed, and used, and (b) the nature of the current practice of evaluation, however, we can identify a number of serious limitations on how theory and practice can reciprocally influence each other. This analysis explains how, for example, (1) disparate evaluation studies can equally claim to follow a given evaluation theory, (2) strictly speaking, no specific evaluation study can be said to validate or invalidate an evaluation theory, and (3) theories are not inherently methodologically prescriptive, in spite of theorists' rhetoric. Presenting this analysis and its conclusions can assist practitioners in a more thoughtful, grounded use of evaluation theories in designing their studies, and can prompt theorists to pursue more descriptive and evaluative studies of practice as a basis for improving their theories.

DELPHI Review of Seven Evaluation Issues, Robert E Stake, Cheryl Bullock & Merrill Chandler, University of Illinois at Urbana-Champaign

Using a DELPHI Technique (Adler & Ziglio, 1996) we refined multiple program evaluation issues, broadly categorized as: Bias, outcomes assessment, representing quality, and recommendations. Presentation of these complex issues after refinement will provide guidance (not currently found in the Standards or Guiding Principles) for problems and

conflicts frequently faced by evaluators. A sampling from both evaluation theorists and practitioners participated in our study. Theorists were contacted first, responses reviewed, issues refined, and sent to round two theorists. A similar procedure was used with practitioners. Specifically, we'll present results from our issue Does the evaluator have greater responsibility to keep from hurting the poor or social servants serving the poor than to keep from hurting more privileged parties? This was posed from the diverse perspective of evaluators' responsibilities to (1) the concept of truth in reporting and (2) social consequences of results. The iterative process, subsequent refinements, and interpretations will be presented.

SESSION 242: MultiPaper   **Room:** Hilo

Sponsored by the Topical Interest Group on Distance Education and Other Educational Technology
Issues in Evaluating Interactive Technologies in Education
Chair: D Michelle Hinn, National Center for Supercomputing Applications
Presenters: *Evaluating Attitudinal Change Over Time toward a College-level Online Science Course*, Kimberly A Bissell, University
                    of Arizona

College courses that are offered in an online format need to be evaluated to understand their level of appropriateness for different students and in different settings. An advanced collegiate science course was evaluated to assess students' attitudes toward the course throughout the semester. Continuous feedback was provided to the instructor so that improvement could be made throughout the duration of the course. Results from a longitudinal data analysis, which include predictors of attitudinal change over time, indicate that attempts to implement online courses should be approached with caution. Issues related to online data collection and evaluation of online education will be discussed.

Increasing Capacity in the Environment of Technology and Teacher Practice, Christopher B Coutts & Walter F
       Heinecke, University of Virginia

Increasing evaluation capacity in educational organizations is a complex endeavor. This paper investigates attempts by an internal evaluator to build and develop the evaluation capacity of a project aimed at integrating technology in teacher education. It documents the role of the internal evaluator in developing evaluation practice in an extremely dynamic environment of technology and teacher practice. Through this example, the paper discusses the political, normative and technical aspects of formative and summative program evaluation of complex programs.

Evaluating Technology Literacy/Innovation Programs: A Stages of Adoption Model, Ronald F Dugan, Jeffrey Kirk,
       Dean T Spaulding, Joseph M Baltrus & Dianna L Newman, State University of New York at Albany

At the current time millions of dollars and person hours are being poured into efforts to integrate technology into education curriculum and instruction. A major issue in the evaluation of technology supported curriculum programs is the documentation of adoption and integration. Several models provide plans or guidelines for adoption while a very few go so far as to suggest indicators of achievement. These few models tend to reflect outcome oriented descriptors, that is, what should be in place when full integration has been achieved. This lack of intermediate outcomes, reflective of stages of integration, makes evaluation difficult. As a consequence, there is a high need for an evaluation model that will address the developmental process of people, resources, and activities, with measurable indicators that can be used to document technology integration. The purpose of this paper is to present such a model along with indicators that have been shown to be successful in multiple settings.

Critical Issues in Evaluating the Effectiveness of Technology, Mark Hawkes, Dakota State University

Where once the emphasis was on building and implementing a technology infrastructure, today it is on evaluating the effectiveness of its use in schools and classrooms. The need for evaluation stems from experiences of the last three decades showing that despite the hundreds of research reports published on the use of technology in education, little of that research community's findings has had bearing on the practitioner's community of the classroom (Norris, Smalka, and Soloway, 1999). Of course, evolving technologies have done much to limit the utility of past evaluations, if not render them completely obsolete. As the technologies themselves have changed, so have our evaluation questions. When computer technologies were first applied to student learning, questions took the form of, "what kind

of computer-based activities improve student learning?" Those evaluations that treated technology as a discrete and isolated entity fortunately gave way to evaluations attempting to embed technology use in a larger process of school change. This discussion formatted presentation identifies the characteristics of evaluations that have consistently helped us identify the kinds of sustained and substantial inquiry we want in learners. The presentation also summarizes the national dialogue on technology evaluation to identify seven critical issues in evaluating the effectiveness of technology in education. These critical issues-informed by past and current evaluations-work on the premise that learning what works is just as important as learning what does not work.

Educational Interactive Video for High School Students: How Do the Teachers Perceive the Program? E Lea Witta, University of Central Florida

Teachers participating in an educational interactive video (ITV) program were surveyed each year for four years. The survey instrument contained 31 five-point Likert-type questions, 3 questions concerning specific comparisons between interactive video and traditional teaching methods, and open-ended questions. A total of 74 teachers responded over the four year time span. Although teachers report that interactive video teaching methods require more preparation and new methods, they also are reporting more comfort with the use of the technology required for interactive video. More troubling is the decreasing view of interactive video as a good addition to the curriculum. Teachers report that time-on-task and learning are the same in traditional and interactive video classes. They overwhelmingly agreed that they received support from the ITV project director, the remote principal, and other sources. Results and procedures are discussed.

SESSION 243: Panel   **Room:** Puna
Sponsored by the Topical Interest Group on Quantitative Methods: Theory and Design
Power in Multivariate Analysis
Chair: Lee B Sechrest, University of Arizona

The concept of power in multivariate analysis is discussed first through an overview of power analysis and its use in hypothesis testing, then by examining some less frequently addressed issues concerning what can be done to optimize power, for example, improvements in design and reliability. Power analysis in structural equation modeling (SEM)is then examined along with a discussion of the interpretation of model fit in relationship to power. Finally, an overview of model testing and the implications of power in SEM is discussed.

Panelists: *Power Analysis Overview: Conceptual Foundations and Possible Extensions*, Ron Wright, University of Arizona

Power analyses are de rigueur in grant and editorial offices with good reason: They indicate that researchers' time and grantor's money will not be wasted and that negative findings can be taken seriously. The concept of power arose regarding null hypothesis testing, and its usefulness there has led to proposed extensions, for example, to structural equations modeling. It is nonobvious, however, how these extensions should work, since bedrock aspects of power analysis (effect size, "minimal clinical significance") are not common concepts in modeling. For example, one traditionally improves power by refining measures or increasing sample size. Refining measures shows promise, yet large n's will inflate chi-square, leading to the rejection of a proposed model, not its acceptance. This presentation will set the stage for other papers on the panel by reviewing the conceptual foundations of power analysis, and examining necessary modifications for useful extensions of this concept to modeling.

Issues in Power: Design and Reliability, Michele Walsh, LeCroy and Milligan Associates

Although there is considerable focus on increasing sample size in order to increase the probability of detecting an effect of an intervention, there is rarely consideration of other factors also potentially under the control of the researcher. We, as evaluators, can often improve on the precision of our estimates of effects even when sample size is limited. Data-based examples of how measurement error and inefficient designs can have an impact on results will be drawn from current evaluations. This paper will focus on providing practical ways of increasing power by improving the design of and measurement in evaluation studies.

Power in Structural Equation Modeling, Kathleen Insel & Aurelio Jose Figueredo, University of Arizona

Recent proposals for the calculation of power in structural equation modeling (SEM) have been made and focus on the number of structural parameters involved and the potential number of participants. However, the current algorithms fail to address effect size as reflected in the magnitude of the structural path coefficients. Yet effect size is often the aspect of power analysis of greatest importance to program evaluators.

What Are These Models We Want So Much To Test? Lee B Sechrest, University of Arizona

The term model is used loosely, perhaps especially so in the context of SEM. Three levels of meaning need to be distinguished. Conceptual models tend to be idealistic and flawless. Statistical models in SEM are considerably simplified and lacking in essential detail. Achieved models represented by results are almost always crude and often scarcely recognizable versions of the original conceptual model. The idea of power to test a model verges on the quixotic. Closing the gap between conceptual and statistical models by moving toward greater specificity of parameter values, point estimates rather than simply directional predictions, would be a great advance, even though quite difficult at present. Using population rather than sample estimates in statistical testing would also be an advance. The greatest advance by far, however, would be to improve the quality of the materials, the data, going into the construction of our models.

SESSION 244: MultiPaper   Room: Kohala

Sponsored by the Topical Interest Group on Collaborative, Participatory & Empowerment Evaluation
Filling Your Toolbox: Collaborative, Participatory and Empowerment Methodologies
Chair: Victor J Friedman, Ruppin Institute
Presenters: *From Logic Models to Modeling Logic-in-use: An Action Science Approach to Participatory Self-assessment*, Daniel V Folkman, University of Wisconsin at Milwaukee and Sheryl Gotts, Milwaukee Public Schools

A logic model shows the causal relationship between major program variables and desired outcomes. The assumption being: if the logic model is valid and the program is shown to be effective, then anyone applying similar program activities in a similar fashion and context should produce similar results. Action Science presents the evaluator with a dilemma. It states that people espouse one theory and act on another. Accordingly, Action Science challenges the evaluator to reveal the logic-in-use that guides the every day behavior of those implementing the program with the proviso that what they say they do is not necessarily what they do. In other words, the logic model is an espoused theory about relationships between variables, which may have little, if any, connection to the theory-in-use that governs program implementation. The proposed paper outlines the basic tenets of Action Science and the implications for evaluation. It provides a set of tools that can be used by evaluators to facilitate a participatory inquiry among program staff into their own logic-in-use. A case study is included that looks at how a team of educators and social workers design and implement highly effective programs in a Milwaukee, central city middle school. The case study demonstrates how Action Science tools are used in guiding inquiry and the impact it has on the practitioners themselves as well as the implications for continuous program improvement and replication elsewhere.

Getting to Outcomes: Methods and Tools for Planning, Evaluation, and Accountability, Pamela S Imm & Abraham H Wandersman, University of South Carolina; Matthew Chinman, Yale University; and Shakeh Kaftarian, Center for Substance Abuse Prevention

This paper describes a user-friendly model of results-based accountability called, Getting To Outcomes: Methods and Tools for Planning, Evaluation, and Accountability (GTO). In the GTO system, accountability is defined as the systematic inclusion of critical elements of program planning, implementation, and evaluation in order to achieve results. This empowerment evaluation tool was designed to help practitioners plan, implement, and evaluate their programs to achieve results. GTO is based on 10 accountability questions which includes the following elements of effective programming: needs and resource assessment, clarifying goals, selecting science and best practices programs, assuring program fit, organizational capacities, effective planning, quality implementation, outcome evaluation, continuous quality improvement, and sustainability. By linking all of these program elements systematically, there is a higher likelihood of achieving the desired outcomes and demonstrating accountability. GTO

was written for various audiences, including practitioners, program directors, and evaluators. The presentation of this paper will focus on the presentation of the GTO model, the theoretical roots of the system, and methods for implementation for a variety of audiences.

A Multilevel, Mixed Methodology Approach to Evaluating Comprehensive Community Initiatives, Sanjeev Sridharan, Aracelis Holguin-Pena & Heidi Vaughn, Caliber Associates

We present a multilevel, mixed methodology design that is being implemented to evaluate a comprehensive community initiative intended to reduce juvenile crimes. Our multilevel design integrates qualitative narrative-based methods (using event structure analysis) with a focus on network structures and a pathways-of-change approach (using structural equation models). The longitudinal research design focuses on the linkages between community-level events, community contexts and community- and system- level outcomes. Our evaluation treats the comprehensive community initiative as a "learning system." Fundamental to the learning system is a dynamic view of program planning and implementation: programs evolve as a result of "lessons learned." We compare a dynamic "learning systems" based view of comprehensive community initiatives with more static views of programs. One of the innovative features of our analysis is our focus on "cracks" in the planning networks.

Representative Case Analysis: An Empowerment Methodology, Jan A Talley, Resource Development Institute

Representative Case Analysis is appropriate for single organism research. The methodology is used to examine individual case data more effectively by including the subjects in the evaluation process. The subject is a co-investigator and knows the purpose of the study and has every chance to comment on the accuracy of the data and the conclusions of the investigator. An application of this design will be presented in the form of a program evaluation of the 4H ASAP program. A Teen Mentor developed a questionnaire to determine what were the effects of 4H on the Teen Mentors' lives. The teen interviewed 18 other teens to answer the survey questions and then helped to analyze the responses. Scores on rating scales were graphed and conclusions were offered. The Representative Case Analysis approach will be examined and lessons learned will be offered to expand the choices of analyses available for empowerment evaluation.

Discussant: Abraham H Wandersman, University of South Carolina

SESSION 245: MultiPaper   **Room:** Kona
Sponsored by the Topical Interest Group on Extension Education Evaluation
Expanding Evaluation Methodologies
Chair: Richard L Poling, University of Tennessee at Knoxville
Presenters: *Evaluation of a Regional Internet Extension Inservice Training: Implications for Evaluation Practice*, Robert M Lippert & Rama Radhakrishna, Clemson University; Owen Plank, University of Georgia; and Charles C Mitchell, Auburn University

In the past three years, three regional Internet inservice trainings have been offered to county Extension agents from several southeastern states. The trainings included an on-line pretest and posttest in addition to participant reaction measures. Subsequently, a fourth 2-week Internet training was offered to over 100 county Extension agents from five states (Alabama, Georgia, South Carolina, North Carolina and Virginia). Ten specialists representing these states contributed to the Web development and Internet discussions. The latest training differed from the previous trainings by including a Kolb learning style instrument and questions about demographics in addition to the participants' reaction measures and a pretest and posttest. A strong majority of the participants have accepted training via the Internet as comparable to that of face-to-face training. The past four Internet trainings have provided justification to conclude that Internet inservice training is a viable and effective alternative to traditional training and reflects the technological changes occurring in Extension. Findings also indicate that the participants increased their knowledge level from pretest to posttest by reading the Web material and engaging in Listserv discussions. The presentation will show the process involved in developing, implementing and evaluating Internet inservice training. The audience will likely find useful information to implement in their own states. The correlation of pretest and posttest scores to the demographic data and the Kolb Learning Style Instrument will provide a basis for better understanding of the process

and the outcomes of Internet training. From an administrative and staff development perspective, the results will show the extent to which travel time and expenses can be minimized by offering training in this format. The success of multi-state collaboration in this context will be evident.

Using Success Markers To Identify Outcomes of Hard-to-measure Behaviors, S Kay Rockwell & Susan Williams, University of Nebraska

Success markers, identified actions/behaviors which indicate successful accomplishment of an outcome, were proposed by Barry Kibel in 1999. Success markers document program staff contributions in helping "clients" achieve hard-to-reach, longer-term sustained outcomes in developmental/transformational activities. A group of Nebraska Extension faculty are piloting the use of success markers to identify behavior changes in Extension programming with "at-risk" audiences. The paper will describe how to determine success markers for program objectives, where success markers are being tested in Extension programming, and how youth are using success markers to judge their own behavioral growth. An illustration will describe how success markers are being used to assess positive lifestyle changes, or growth, in a program focusing on improving the social and economic situation for culturally and economically disadvantaged youth. The illustration will help participants understand how success markers can be used alone, or as a preliminary step in a more complex evaluation.

Handling Response Shift Bias in Extension Evaluation, F Richard Rohs, University of Georgia

This paper investigates the effects of response shift bias on outcomes using a self-report measure in extension leadership development programs. While participants in this study rated themselves as having a "high" level of leadership skill at the end of the program, significant differences were found between their self-report ratings using the pretest/posttest and the then/posttest approach. The degree of response shift (then/posttest pre/post comparison) was also significant. The findings from this evaluation together with other evaluations cited suggest than when employing self-report measures, the then/post approach provides a less conservative and more accurate means of assessing leadership skill development than would the traditional pretest/posttest approach. Additional evaluation techniques will also be offered for controlling response shift effects.

SESSION 246: MultiPaper   Room: Ewa
Sponsored by the Topical Interest Group on Health Evaluation
Evaluating Implementation of Prevention Programs
Chair: Liesel A Ritchie, Mississippi State University
Presenters: *Correlates of Community Implementation of Science-based Prevention Practice*, John S Briney, Michael W Arthur,
              David Hawkins & Richard F Catalano, University of Washington

Researchers at the Social Development Research Group (SDRG) have developed methods for evaluating community implementation of science-based prevention programming. This model of science-based prevention programming includes the following 4 steps: 1) assessing risk and protective factors associated with adolescent drug abuse and antisocial behavior, 2) prioritizing and targeting areas of greatest prevention need, 3) selecting and implementing research-based prevention that address the identified needs and 4) monitoring the impact of these strategies on targeted risk and protective factors and outcomes over time. This presentation will report findings from a study of community-wide implementation of science-based prevention programming. Data from telephone interviews with community leaders and prevention experts in 41 communities across 7 states have been analyzed to assess implementation of this science-based prevention framework. Factors hypothesized to influence adoption and implementation of this framework, such as prevention training and community readiness factors, have also been studied. Implications of these results for the dissemination of science-based prevention practice will be discussed.

Characterizing Community Implementation of Science-based Prevention Programming, Renita R Glaser, Michael W Arthur, J David Hawkins, Richard F Catalano, & Sunah Kim, University of Washington

An emerging challenge for the field of prevention is to translate science-based knowledge into widespread practice. However, despite advances in prevention science, little is known about the processes required to transfer this

knowledge base to community members in order to promote effective prevention programming on a broad scale (Biglan, 1995; Mitchell, Stevenson, & Florin, 1993). In order to study the processes involved in translating prevention science into community prevention practice, methods for measuring the implementation of science-based prevention strategies are needed. This presentation will introduce a theory of and methods for assessing community implementation of a science-based prevention model developed by the Social Development Research Group (SDRG). Work on this presentation comes out of a larger research project studying the diffusion of science-based prevention practice funded by NIDA, CSAP, OJJDP and the Department of Education.

Evaluating the Mississippi Tobacco Pilot Prevention Program: Benchmarking for the Future in the Nation's First Tobacco Settlement State, Liesel A Ritchie, Mississippi State University

In 1999, the Mississippi State Department of Health contracted with the Social Science Research Center (SSRC) at Mississippi State University to conduct the Overall Evaluation of the Mississippi Tobacco Pilot Program (MTPP). As one of five components of the evaluation, the Overall Component was designed to capture the social climate in which tobacco control and prevention efforts are taking place.   Among the elements of the Overall Evaluation are: The 1999 Mississippi Social Climate Survey of Tobacco Use and Tobacco Control; examination of extant data; visualization of program data; development of a logic model for the MTPP; and electronically facilitated group discussions with program managers. Evaluators from the SSRC will present an overview of the evaluation design and findings, the processes and challenges of conducting an evaluation in this politically charged environment, and the future of evaluations of tobacco use and control in Mississippi.

SESSION 247: Panel   **Room:** Honolulu
Sponsored by the Topical Interest Group on Alcohol, Drug Abuse, and Mental Health
Building the Capacity of States and Communities to Implement and Evaluate Systems Change To Prevent Substance Abuse: Three Examples from CSAP's State Incentive Grant Program
Chair: Dave Robbins, Center for Substance Abuse Prevention

Many federal grants provide funding to state agencies with the expectation that they will produce outcomes at multiple levels, including the state, community, and program. This is the case with the Center for Substance Abuse Prevention's State Incentive Grant program, a three-year, $200 million component of the White House National Drug Strategy. The program funds 21 state governor's offices to coordinate and leverage federal and state funding for substance abuse prevention, develop statewide comprehensive prevention plans, and implement science-based prevention programs within communities. The evaluator's challenge is to assess the impact of these statewide and community systems change initiatives and to build the capacity of states and communities to evaluate the outcomes of these efforts. This panel will show how evaluations of state-level programs can demonstrate systems change and program outcomes at multiple levels, including changes within communities. Participants will learn how these unique program evaluations can be conducted.

Panelists: *Examining the Role of Community Partnerships in Statewide and Community Systems Change*, Margaret K Gwaltney, COSMOS Corporation and Dave Robbins, Center for Substance Abuse Prevention

Over the last decade, there has been tremendous growth in the formation of community partnerships that address a variety of complex social and health problems. Most of these initiatives have originated at the community level and have been guided by local concerns. In CSAP's State Incentive Grant program, community partnerships not only address local needs but also are given incentives to incorporate state-level priorities in their activities -- for example, implementation of science-based prevention programs and other priorities expressed in state prevention plans. States have elected to fund community partnerships as a way of furthering their state systems change efforts and achieving state and community-level prevention outcomes. This presentation will describe the results of an assessment of how community partnerships have helped states achieve systems change in substance abuse prevention. The assessment is based on interviews of state-level officials and partnership directors.

Assessing Organizational Dynamics and Other Characteristics of Community Coalitions, Robert L Flewelling,
    Research Triangle Institute

Coordinators of Vermont's SIG-funded community coalitions have been surveyed annually during the three-year SIG initiative. Individual coalition members were surveyed early in the second year of funding. These surveys were designed to gather information on attributes believed to be important to effective coalition functioning and coalitions' ability to produce desired intermediate and long-term objectives. Constructs that were assessed include coalition membership and structure, inclusiveness, level of participation by coalition members, the coalition's decision making process, and perceived strengths and limitations of the coalition. The information obtained from coalition members enriches findings from the coordinator surveys, and provides a measure of the level of agreement in perceptions of coalition coordinators and members. Information used in these assessments can be used to better understand the nature and the dynamics of local coalitions, identify and address problems, and examine relationships between coalition characteristics, state and local systems changes, and other outcomes of interest.

Evaluating Community Coalition Efforts in Reducing Adolescent Substance Use Rates, Patty Martin, Research
    Services of Greenbush

In October 1997, the state of Kansas received grant funding from the Center for Substance Abuse Prevention (CSAP) as part of the State Incentive Cooperative Agreement initiative. The purpose of the grant was to change state and local community prevention systems in order to more substantially impact adolescent substance abuse rates. Kansas elected to implement the community initiative through community coalition plans focused on local needs. The evaluation of the community initiative in Kansas has been challenged not only to measure coalition efforts, but also to measure immediate, intermediate, and long-term changes in the community environment. This presentation will describe the evaluation process used by the state of Kansas, distribute copies of tools used in the evaluation, and discuss the relationship between coalition efforts and community level change in reducing adolescent substance use rates as captured by the evaluation.

Outcomes from the National Cross-site Evaluation of the Community Coalitions Program, Ping Yu, COSMOS
    Corporation and Shakeh J Kaftarian, Department of Health and Human Services

Outcomes from the national cross-site evaluation of the Center for Substance Abuse Prevention's Community Coalitions Program are presented. The primary hypothesis guiding the evaluation is that coalition communities lead to a reduction in substance abuse and related health problems over time relative to their matched comparison communities. The multi-year, treatment-control evaluation design employs a large sample of communities (85 coalitions communities and 85 matched comparison communities) and utilizes time-series data from archival sources of public health indicators. A random-coefficients model was developed to assess changes in substance abuse and related health outcomes extracted from archival sources over a seven-year period. Archival data on substance abuse-related hospitalizations, fatal crashes, and crimes were of interest. The presentation reports results of the analyses from baseline and implementation data from coalition communities and their matched comparison communities. The presenters discuss both main outcome analyses and subgroup analyses of coalition outcomes by prevention dosage, coalition organizational features, and baseline drug and poverty conditions. The presentation should enable the participants to understand the features of one of the federal government's major substance abuse prevention initiatives. The audience also will learn how to use multi-year data from a variety of archival sources to assess the impact of prevention activities in a national cross-site evaluation setting.

SESSION 248: Panel   **Room:** Kahuku
Sponsored by the Topical Interest Group on Collaborative, Participatory & Empowerment Evaluation
Walking the Talk: Evaluation Coaches and Community Collaboratives
Chair :   Arlene B Andrews, University of South Carolina

The Institute for Families in Society (IFS) is an interdisciplinary research unit of the University of South Carolina.

Much of the work of IFS involves helping community-based organizations develop and evaluate collaborative activities designed to address the needs of vulnerable populations and enhance the capacities of communities to find answers to local concerns. Using an empowerment evaluation approach, evaluation teams work as coaches to train community participants to be results-driven while monitoring efforts and quality. This panel will present a discussion on the collaborative work of IFS, highlighting three aspects:
* The principles of empowerment evaluation that provide the framework for working with community collaboratives;
* The mirror image challenges of collaboration for community-based organizations and evaluation coaches; and
* The findings of a multi-site, multi-method evaluation of qualities of successful collaborations.

Panelists: *Empowerment Evaluation: A Framework for Building Community Collaborations*, Patricia Motes, University of South Carolina

Using an empowerment evaluation approach, IFS works with community organizations to strengthen organizational capacity, monitor performance, and evaluate results. IFS uses an evaluation team that works as coaches for community members trained to be results-driven while monitoring efforts and quality. IFS is providing consultation to more than 60 nonprofit organizations, 8 community coalitions with over 40 organizational members, several public hospitals, schools, and a statewide effort to create 46 county-level partnerships. State and federal grants and private foundations fund these services. The work of the evaluation team is based on principles of empowerment evaluation that include:
* Using evaluative information for routine program improvements and policy decisions;
* Stakeholder involvement (i.e., consumers, staff, board, funders);
* Inclusive processes (inclusive participation and decision-making throughout);
* Collaboration among stakeholders;
* Building capacity for program development and program evaluation;
* Valuing diversity as integral to program success; and
* Sharing of lessons learned.

The Mirror Images of Collaboratives and Evaluation Teams, Anita G Floyd, University of South Carolina

The Institute for Families in Society (IFS) in an interdisciplinary research unit of the University of South Carolina. Much of the work of IFS involves helping community-based groups develop and evaluate collaborative activities designed to address the needs of vulnerable populations and enhance the capacities of communities to find answers to local concerns. Through our experience with community-based collaboratives, we have identified several challenges to collaboration that are manifested within the Institute's own faculty and staff in their efforts to collaborate. These challenges include opportunities and motivation for participation, leadership development, role confusion, conflict resolution, and sustainability. This presentation will discuss these issues as mirror images for community-based groups and their evaluation teams. The presenters will seek suggestions from other evaluators for strategies that can help evaluation teams implement and practice the guidelines for collaboration that are the bases for their evaluations of community-based collaboratives.

Characteristics of Successful Collaboratives: Findings from a Multi-site, Multi-Method Evaluation, Judith A Whiting-Fickling, University of South Carolina

Many current funding initiatives encourage the establishment of community-based collaboratives to address the needs of vulnerable populations and to enhance the capacities of communities to find answers to local concerns. These collaboratives possess similar and dissimilar characteristics, engage in different strategies to maintain participation, offer varying opportunities for leadership development, and are confronted by and deal with conflict of many kinds. The Institute for Families in Society (IFS) will present the findings of a multi-site, multi-method evaluation to identify qualities of successful collaboratives. This presentation will highlight the findings of the evaluation that provide a basis for recommendations to emerging or reorganizing collaboratives.

SESSION 249: Panel   **Room:** Oahu
Sponsored by the Topical Interest Group on Cluster, Multi-site and Multi-level Evaluation

Lessons Learned in Conducting Multi-site Evaluations in the Health Resources and Services Administration
Chair: Roger B Straw, Health Resources and Services Administration

It is well understood that factors such as fidelity of "program" implementation, evaluation design, choice of measures, and quality of data collection affect the ability to interpret the results of an evaluation study. In a multi-site evaluation, there are a variety of additional challenges that face evaluators. These challenges can include differences across sites in how the program is defined, in client populations, in organizational and community contexts, and in a variety of other ways. In addition, the way in which the multi-site evaluation is planned and conducted can have important implications for the interpretability of the findings. Each of the presenters on this panel will describe a HRSA multi-site evaluation, talk about the specific challenges in that situation, and offer ideas for dealing the challenges.

Panelists: *Multi-site Evaluation Issues: The Story of 34 Ryan White CARE ACT Special Projects of National Significance (SPNS) Grantees and Their Evaluation and Technical Assistance Center (ETAC)*, Barbara Aranda-Naranjo, Health Resources and Services Administration and Peter Messeri, Columbia University

Thirty-four Ryan White CARE ACT Special Projects of National Significance (SPNS) grantees were funded under the HIV Integrated Service Delivery Models (ISDM) Initiative. This initiative was designed to define and evaluate formal linkages and integration of mental health, substance abuse treatment, rehabilitation and/or critical HIV services with HIV ambulatory medical care in new and existing innovative projects. The initiative also funded the establishment of the SPNS Evaluation and Technical Assistance Center (ETAC) to coordinate the implementation of multi-site evaluations within groups of similar projects. A comprehensive and multi-method approach that combined the use of both quantitative and qualitative evaluation methods was implemented across sites. The paper will discuss the multi-site evaluation process (identifying the workgroups, development of common data elements/database/codebooks, instruments, design and implementation of project's local evaluation activities) and lessons learned.

Interpreting Multi-site Study Results: Do Ancillary Services Improve Primary Care for People Living with HIV (PLWH)? Richard Conviser, Health Resources and Services Administration and Moses Pounds, HIV Aids Bureau

Eight Ryan White CARE Act grantee sites, using existing data sets, examined whether ancillary services contribute to entry into and retention in primary care for PLWH. The sites varied in location, populations served, ancillary services offered onsite, and types of information available in their data sets. Most sites had clinical databases; several had surveys of populations served. The populations studied varied with respect to gender distribution; HIV risk factor; proportions with public, private, and no insurance; and stage of HIV illness. Analytic methods also varied across the sites, as did findings. The paper will discuss how to interpret differences in findings as attributable to population differences (e.g, poor minority injection drug users vs. privately insured gay white men), setting differences (urban and rural patients have different transportation needs), differences in the way ancillary services are delivered, or differences in analytic methods.

Th*e National Evaluation of Healthy Start: Challenges in Interpreting Findings from a Multi-site Community-based Public Health Program*, Karen Thiel Raykovich, HIV Aids Bureau; and Barbara L Devaney, Lorenzo Moreno & Embry M Howell, Mathematical Policy Research Inc

The evaluation of a fifteen site, community-based initiatiave to reduce infant mortality presented a series of challenges which impacted the design of the evaluation, the actual conduct of the evaluation, all phases of data analysis, and, ultimately, the interpretation of findings. The evaluators were succesful in the application of a multi-pronged set of evaluation methodologies to a diverse group of grantees. Comparable data were generated in the process analysis of program implementation and state vital record data were used to analyze program outcomes, However, program design features, as well as grantees' difficulties with client level data collection to document service utilization, affected the ability of the national evaluators to attribute health status outcomes to particular program components. How these challenges were addressed both analytically and interpretively will be the focus of this presentation.

SESSION 250: Skill-building Workshop   **Room:** Waialua

Sponsored by the Topical Interest Group on Pre-K-12 Educational Evaluation
Building Capacity for Evaluating Performance
Presenter: Macke Raymond , University of Rochester

Even the finest study of program performance is at risk of failure if the remaining elements of the decision making process are not attuned to the creation support and use of program performance assessments. Drawing on experience in the public, private and philanthropic sectors, this session will present an overview of the critical points in the life cycle of program evaluation that can hobble good program assessment practice. The necessary components of a successful system of program accountability and performance assessment are developed through group participation exercises. Attendees will obtain user-ready materials that they can employ immediately to better support the development of informed decisions that incorporate available study results in a constructive fashion.

SESSION 251: MultiPaper   **Room:** Waianae
Examining the Relationships between Evaluator and Evaluand
Chair: Catherine C Dunlop, Simon Fraser University
Presenters: *Inoculation Strategies for Successful Evaluations: Using Negotiation Theory To Imbue New Habits of Practice*,
                    Catherine C Dunlop, Simon Fraser University

Self-reflective evaluation practitioners are mindful of potential threats to the integrity of the evaluation process and product. These threats are often related to the following: tension between a flexible design and limited resources, confusion over access to information and who needs to know what and why, and varying expectations around the utilization and application of evaluation results. Negotiation theory provides useful strategies that can help evaluation practitioners be prepared for these threats and can create future immunity through new habits of practice.

The Evaluator and Organizational Self-evaluation: Harmonious Pairing or Uneasy Partnership? Joan A Heffernan,
       Cornell University

What are the opportunities and challenges faced by evaluators in coaching an organization through a sustainable evaluation process? In what ways can an evaluator impact the process and outcome? Should this be a goal or concern? This paper presents two differing scenarios of organizations' efforts to evaluate an ongoing program; it compares and contrasts the role of the evaluator in each context. Is the evaluator's point of entry, into an evaluation, a critical issue? This paper focuses on the two different evaluation experiences and contexts: one, a student-oriented evaluation of an undergraduate curriculum at a large educational institution; the other, a consumer evaluation of a small, outreach program offered by a community-based organization. What are the challenges to reliability and validity inherent in the process of self-evaluation, and can these issues be addressed and overcome? How might a proliferation of organizational self-evaluations impact the evaluation profession? These are some of the issues discussed in this paper, depicting the author's experiences in two widely disparate evaluation settings with diverse processes, cultures, contexts, and participants.

Resolving Conflict When an Adversarial Relationship between Clients and Evaluators Exists, Mari R Kemis, Janet
       Padgitt & Ching-Chun Shih, Iowa State University

The relationship between a client and an evaluator is critical to the productive use of evaluative information. This roundtable will explore situations in which clients and evaluators have conflicting perceptions about the role of evaluators and/or the purpose of evaluation. Specifically, participants will discuss situations where the client is threatened by the evaluation or attributes undue power to the evaluator, as well as situations where the client doesn't understand or has conflicting opinions about the role of the evaluator or how evaluation results can be used to strengthen a project or program. Examples based on actual evaluation experiences will highlight adversarial relationships between evaluator and client. Roundtable participants will have the opportunity to share their stories and ideas for resolving these problems constructively.

The Dreaded Evaluator: Fostering Cooperation from Program Stakeholders, Anne E Marshall, University of California
       at Los Angeles

Most evaluators have witnessed the heightened tension in program administrators when the word evaluation is mentioned. When introduced to an evaluator, the predominant reaction on the part of program administrators is one of masked fear and loathing. An evaluator is perceived from the outset as the paper pusher whose goal is to find fault with the program or how it is being managed. When program administrators perceive the evaluator and evaluation process negatively, the utilization of findings is greatly limited. Dread of the evaluator's activities can also impede cooperation necessary for obtaining information for the evaluation. Additional complications may arise when the evaluator is also involved in the program's implementation. This paper addresses the issue of "evaluator-dread" and the importance of re-defining evaluation and the role of the evaluator for the stakeholders at the outset of the evaluation.

SESSION 252: Panel   **Room:** Molokai
Sponsored by the Topical Interest Group on International and Cross-cultural Evaluation
Building Evaluation Capacity in Russia
Chair: Molly Hageboeck, Management Systems International

Presenters will examine questions about the evaluation capacity of Russian professionals and Russian organizations. The objective evaluation of programs is a new concept for government and for Russia's emerging NGO sector. In 1997, an entrepreneurial group of Russian professionals took note of the dependence of international organizations on foreign expertise in evaluation and translated their interest in acquiring evaluation expertise into a Certificate Program in Evaluation which MSI developed and presented, with donor funding. A second class will have completed the Certificate Program by the end of May. In September 2000 the first all Russia conference on evaluation will be held in Siberia. It will be a stocktaking exercise for NGO organizations. Presenters will discuss conclusions from Siberia as well as current surveys that focus on (a) demand for evaluation in Russia among local and donor organizations and (b) evaluation experiences of graduates of the 1997 Certificate Program.

Panelists: *Transferring Evaluation Skills to Russia – The Impact of Practical Training*, Richard N Blue, Management Systems
                   International

This presentation will focus on the Certificate Program in Evaluation that was developed for Russian nationals. Two classes have completed this course. Dr. Blue will also report out survey results concerning the evaluation experiences described by graduates of the 1997 program. He will also explore the question of how well the Certificate Program's requirement for students to complete an evaluation of an NGO program or project as part of the course serves to 'jump start' their capacity to work in this field.

Building the Demand for Monitoring and Evaluation with Russia's NGO Sector, Alexander Borovikh & Ekaterina
     Greshnova, Center for Non-governmental Organizations and World Learning

This presentation will focus on the demand for monitoring and evaluation in the emerging NGO sector in Russia. Initially, the demand for monitoring and evaluation came from international donors and was linked to grant program funding. After gaining an understanding of monitoring and evaluation, Russian NGOs have become interested in these practices and in how they can benefit NGOs directly, as well as inform funding sources about progress and performance. Organizational commitment --more than individual skills in monitoring and evaluation -- is the key to full integration of these practices into Russia's NGO sector.

The Wider Market for Skilled Russian Evaluators, Molly Hageboeck, Management Systems International

This presentation expands the arena in which a demand for evaluation skills exists in Russia. Findings concerning the demand for evaluation skills beyond the NGO sector will be presented, e.g., among international donor organizations, within government, etc., will be presented. Findings concerning the interest of organizations that monitoring and evaluation in using Russian evaluation specialists will also be examined. The issue of what it takes for a country to build a credible evaluation capacity will be raised in this presentation and used to open the session for broader participation -- by selected discussants and members of the audience.

Discussant: Marcia McGraw-Olive, The Eurasia Foundation

SESSION 253: Business Meeting and Presentation   **Room:** Maui
Sponsored by the Topical Interest Group on Independent Consulting
Resiliency and the Independent Evaluator: Protective Factors for the Business of Evaluation
TIG Chair: Melanie A Hwalek, SPEC Associates
Program Chair: Sally L Bond, The Program Evaluation Group
Facilitators: Joann Hoffman, Hoffman Clark and Associates
                  Jane Reisman, The Evaluation Forum

Two independent evaluators will lead an exploration of the protective factors that make it more likely for an independent evaluator to experience positive outcomes. This highly interactive (and hopefully humorous) session will focus on three dimensions of the independent evaluator consulting practice: client management, financial management, and balancing your work and personal/family life. A self assessment tool will be distributed and analyzed during the session. A logic model will NOT be developed during this session because our theory is that anyone with sound reasoning would not become an independent evaluation consultant. Rationale for alternative format: This session will be both instructive and therapeutic. We intend for it to provide an opportunity for independent evaluators to learn from one another -- using a tongue-in-cheek format for exploring ways to enhance independent evaluator resiliency.

SESSION 254: Debate   **Room:** Kauai
Sponsored by the Topical Interest Group on Evaluation Use
Evaluation Use: Was It Good For You? A Debate
Moderators: Valerie J Caracelli, US General Accounting Office
                  Jean A King, University of Minnesota
Debaters: Michael Quinn Patton, The Union Institute
                  Gary T Henry, Georgia State University

Should or shouldn't evaluation use be a defining goal of evaluation? At Evaluation '99, Gary Henry challenged the emphasis placed on use as a defining goal of evaluation. Does utilization-focused evaluation have unrecognized and potentially negative ramifications for practice? Does it place too much emphasis on program improvement? Does a focus on use lead evaluation resources away from supporting democratic institutional processes leading to social betterment? Michael Quinn Patton picks up the gauntlet to defend his position that the primary goal of any evaluation is intended use by intended users. This debate between the experts is intended to sharpen our understanding about key issues that pertain to use and evaluation practice.