

JMDE

Journal of MultiDisciplinary Evaluation

Number 2, February 2005

Part III

Editors

E. Jane Davidson & Michael Scriven

Associate Editors

Chris L. S. Coryn & Daniela C. Schröter

Assistant Editors

Thomaz Chianca

Nadini Persaud

John S. Risley

Regina Switalski Schinker

Lori Wingate

Brandon W. Youker

Webmaster

Dale Farland

Mission

—The news and thinking
of the profession and discipline of evaluation
in the world, for the world—

A peer-reviewed journal published in association with

The Interdisciplinary Doctoral Program in Evaluation
The Evaluation Center, Western Michigan University

Editorial Board

Katrina Bledsoe	Shawn Kana'iaupuni
Nicole Bowman	Ana Carolina Letichevsky
Robert Brinkerhoff	Mel Mark
Tina Christie	Masafumi Nagao
J. Bradley Cousins	Michael Quinn Patton
Lois-Ellen Datta	Patricia Rogers
Stewart Donaldson	Nick Smith
Gene Glass	Robert Stake
Richard Hake	James Stronge
John Hattie	Dan Stufflebeam
Rodney Hopson	Helen Timperley
Iraj Imam	Bob Williams

Table of Contents

PART III: Global Review—Publications

Summary of American Journal of Evaluation, Volume 25(4), 2004.	145
<i>Melvin M. Mark</i>	
New Directions for Evaluation.	150
<i>John S. Risley</i>	
Education Update.....	153
<i>Nadina Persaud</i>	
The Evaluation Exchange—Harvard Family Research Project	157
<i>Brandon W. Youker</i>	
Canadian Journal of Program Evaluation, Volume 19(2), Fall 2004	161
<i>Chris L. S. Coryn</i>	
Evaluation: The International Journal of Theory, Research and Practice, Volume 10(4), October 2004	164
<i>Daniela C. Schröter</i>	
Measurement: Interdisciplinary Research and Perspectives, Volume 1(1), 2003	168
<i>Chris L. S. Coryn</i>	

Summary of American Journal of Evaluation, Volume 25(4), 2004

Melvin M. Mark

The following is excerpted from the introduction to Volume 25, Issue 4 of the *American Journal of Evaluation*, by former *AJE* editor Dr. Melvin M. Mark. It is reprinted here with permission from Dr. Mark; *AJE*'s current editor, Dr. Robin Miller; and the American Evaluation Association (AEA). The *American Journal of Evaluation* is the official journal of the American Evaluation Association and is distributed to AEA members as part of their membership package. To learn more about AEA and how to receive *AJE*, please go to www.eval.org.

In the first paper, **Robert Orwin, Bernadette Campbell, Kevin Campbell, and Antoinette Krupski** examine the effect of the 1997 termination of the Social Security Administration's Disability Insurance and Supplemental Security Income benefits for persons diagnosed with drug or alcohol addiction. The paper describes and illustrates innovations and recent developments in quantitative methods for evaluation, including a combination of the interrupted time series with growth curve modeling; propensity scoring analyses; the use of alternative ways to estimate the counterfactual; and sensitivity analyses to empirically assess the plausibility of validity threats. Importantly, Orwin and his colleagues go further, carefully conducting and considering the implications of a set of post-hoc exploratory analyses. These analyses suggest a far more nuanced interpretation of the effects of benefit termination than did the primary, state-of-the-art tests. The paper is a valuable example of the principled examination of quantitative data that can move us beyond overall, global estimates of an intervention's average effects.

In the second paper, **Katherine Ryan** describes, illustrates, and critiques three approaches which fall under the broader umbrella of "democratic evaluation approaches". The three are the seminal democratic evaluation approach of MacDonald, the deliberative democratic evaluation approach of House and Howe, and the emerging notion of communicative evaluation advocated by Niemi and Kemmis. Ryan compares and contrasts these three approaches, in part by presenting for each a vignette describing a case in which the approach was implemented. Ryan goes beyond simply examining the three democratic evaluation approaches in the abstract. Instead, she considers the implications of these approaches in an environment in which educational accountability has been shaped by the No Child Left Behind legislation and related forces. In effect, Ryan asks how evaluators can contribute to more democratic forms of educational accountability.

In the third paper in this issue, **Christine Leow, Sue Marcus, Elaine Zanutto,** and **Robert Boruch** address the question of whether taking advanced courses in math and science improves performance on basic achievement tests. Leow and her colleagues use propensity score methods in an attempt to control for the biases that otherwise would result because of the systematic differences between students who take advanced courses and those who do not. The paper thus will be of interest to readers who would like to learn more about propensity score analyses. Perhaps of more interest, Leow and her colleagues illustrate a kind of sensitivity analyses, which allows them to examine how susceptible their findings are to what is called hidden bias, that is, the bias that might arise from background factors that are not controlled for in the analyses. Sensitivity analyses should be an important technique in the tool kit of quantitative evaluators, as a way of helping assess how much uncertainty one should ascribe to evaluation findings.

As one outcome of the evidence-based practice movement, there seems to be a growing trend whereby mandates, recommendations, or incentives are put into place in an effort to lead practitioners to use programs that have passed some evaluative threshold. But this trend raises several questions, among them: How do practitioners learn about so-called evidence-based programs? What are the processes by which they adopt such programs and eliminate their current programs? Are the evidence-based programs likely to be implemented with sufficient fidelity that one would expect good outcomes? **Tena St. Pierre** and **D. Lynne Kaltreider** address these and related questions, in a replicated case study investigating school adoption and implementation processes of an evidence-based substance abuse prevention program. The findings should be noteworthy to those interested in program implementation, in the way schools choose to adopt and adapt programs, and more generally in how mandates for evidence-based practice play out in real life.

Huilan Yang, Jianping Shen, Honggao Cao, and Charles Warfield address "multilevel evaluation," which arises, for example, when there are multiple site-level projects within a broader programs or, as in the example Yang and colleagues discuss, three levels: project, cluster, and initiative. The authors of this paper lay out a process to facilitate multilevel evaluation alignment, that is, to facilitate congruence, compatibility, and efficiency across the evaluations at the different levels. In one sense, the process can be seen as the application of sound evaluation planning in the multilevel program context. However, Yang and her colleagues argue that the literature on multisite evaluation demonstrates the need for an alignment model specifically focused on multilevel evaluations.

In the final paper in the *Articles* section, **Tricia Leakey, Kevin Lund, Karin Koga, and Karen Glanz** address an issue of considerable importance to those who

evaluate programs based in schools or, more generally, who work with participants who are minors: obtaining parental consent. Leakey and her colleagues describe a case from their own evaluation experience examining a smoking prevention program. They employed different consent procedures at different times, and describe their experiences in this article.

In the *Method Notes* section, **Henry May** addresses a classic and continuing concern for evaluators: How can we best communicate our results, especially statistical findings, to those who need to make sense of and use evaluation findings? May discusses and illustrates the use of three guidelines for formulating and presenting more meaningful statistics. These are understandability, interpretability, and comparability. May also offers several interesting and valuable examples for reporting a variety of statistics, both simple and complex, in more meaningful ways.

This issue includes an atypical contribution in the *Exemplars* section. In the past, this section has presented a series of interviews with evaluators who discuss a specific evaluation they had conducted. In those interviews the section editor, Jody Fitzpatrick, questioned the evaluator to understand more about the various choices he or she made throughout the evaluation, from the initial steps in planning, to the involvement of stakeholders, to the data collection methods and evaluation approaches employed, and to the steps taken to disseminate findings and facilitate use. With the naming of a new editor for the *Exemplars* section, I invited **Jody Fitzpatrick** to reflect on the numerous interviews she had conducted. Such an effort to "sum up" previous work in a section of *AJE* is not completely new. Two years ago, Michael Morris (2002), section editor of *Ethical Challenges*, invited Lois-ellin Datta (2002) and Nick Smith (2002) to examine previous commentators' responses to 10 ethical challenges Morris had previously posed in the section. As

was the case with the Datta and Smith reflections, Jody Fitzpatrick has provided a fascinating piece. In essence, she treats the interviews from *Exemplars* as a set of case studies, allowing her to examine similarities and differences across a set of evaluators in terms of such important characteristics as preferred evaluation role, the purpose of evaluation, the factors the evaluator used to organize and frame their work, the nature of stakeholder involvement, and method choices.

Finally, after too long a delay, the *Book Review* section reappears. **Shirley Copeland** reviews a recent book by Martha Feldman, Jeannine Bell, and Michelle Berger on the process of gaining access and qualitative research. Thanks to Shirley for an informative review.

References

- Datta, L-e. (2002). The case of the uncertain bridge. *American Journal of Evaluation*, 23, 187-197.
- Morris, M. (2002). Ethical challenges. *American Journal of Evaluation*, 23, 183-185.
- Smith, N. L. (2002). An analysis of ethical challenges in evaluation. *American Journal of Evaluation*, 23, 199-206.
- The Oral History Project Team (2003). The oral history of evaluation Part I. Reflections on the chance to work with great people: An interview with William Shadish. *American Journal of Evaluation*, 24, 261-272.
- The Oral History Project Team (2003). The oral history of evaluation Part II. An interview with Lois-ellin Datta. *American Journal of Evaluation*, 25, 243-253.

New Directions for Evaluation

John S. Risley

The two most recent issues of *New Directions for Evaluation* each cover international perspectives in the field. The Fall 2004 issue (Rugg, Peersman, and Carael) addressed “Global Advances in HIV/AIDS Monitoring and Evaluation” while the Winter 2004 issue (Russon and Russon) concerned “International Perspectives on Evaluation Standards.”

The Fall issue covers a wide range of topics in HIV/AIDS monitoring and evaluation including political influences, international perspectives focusing on the roles of the United Nations and the World Bank, and specific program evaluation experiences.

While this issue deals mostly with subjects specific to HIV/AIDS prevention and treatment it does offer some insight into evaluation questions with a wider impact. These questions are identified nicely by Michael Quinn Patton in his overview chapter “A Microcosm of the Global Challenges Facing the Field: Commentary on HIV/AIDS Monitoring and Evaluation.” Patton identifies issues touched on by the various authors that are seen in many evaluation contexts, such as the denial of problems despite compelling evidence, the use of evaluation for accountability vs. program improvement, and selective use of evaluation findings.

The three main critiques Patton offers are: 1) the sense that the authors are overwhelmed by numbers and fail to include stories of real people affected by HIV/AIDS, 2) the “deeply entrenched mechanistic linearity” (p. 168) in evaluation, and 3) the acceptance of unrealistic goals. He argues for including stories of real

people along with the reporting of data so that the data doesn't take on "an abstract life of their own." (p.168) He criticizes the "input-activities-output-outcome-impact" framework presented in one chapter as the "basic organizing framework" endorsed by all agencies "to organize the data required to monitor program progress." (p. 37) Patton cites Uganda and Brazil as two successful cases of countries greatly reducing their HIV/AIDS infection rates through "complex, dynamic systems change." (p. 169) In situations such as these "complex systems change mapping and networking models hold more promise than do traditional linear-logic models." (p. 169) Patton also contends that evaluators should not merely accept the program goals when evaluating a program. Specifically, he says overly optimistic goals, like those set by the United Nations regarding HIV/AIDS, should be questioned by evaluators.

The Winter issue reviews the development of evaluation standards in the United States, Western Europe, Africa, Australasia, and at some large international nongovernmental organizations. Craig Russon (co-editor of the issue with Gabrielle Russon) provides an overview of the development of national- and regional-level evaluation standards in the years since the Joint Committee's Program Evaluation Standards were adopted in 1994. He notes that the Joint Committee Standards were influential on all standards that followed, acting as either a "point of departure" or as an example of what some national and regional groups "did *not* want their standards to be." (Russon, p. 90)

One such instance is addressed by Doug Fraser in his review of the experience of the Australasian Evaluation Society's (AES) ongoing process of developing a policy on standards. Fraser recounts how the Joint Committee Standards were the starting points but they "depended on a number of fundamental preconditions or assumptions that did not necessarily hold true" in the environment of Australia and

New Zealand. (p. 71) The Program Evaluation Standards concentrated on “risks that were internal to the evaluation itself: risks of evaluators’ overreaching themselves, overlooking key aspects of their task, exercising bias, behaving unethically, or failing to apply an appropriate range and quality of techniques.” (p. 71)

AES members saw the risks and threats they wished to address as being external to the process of evaluation. These risks and threats concern how evaluation is managed, planned, supported and used. Many of these issues are controlled by those who fund and use evaluation, therefore any standards should address these audiences, not simply practicing evaluators.

Fraser recounts how the AES has long had a practitioner code of ethics but the process of developing a set of standards for evaluation stalled in 2001 owing to many factors. However, an Ethics and Standards Committee did prepare a draft set of standards for the society’s 2001 conference. This draft included six categories: transparency, utility, practicality, cost-effectiveness, ethics, and accuracy/quality/comprehensiveness. (p. 77) Fraser notes the prominence of transparency in this draft as contrasted with the Joint Committee Standards.

References

- Rugg, D., Peersman, G., & Carael, M. (Eds.). (2004). Global advances in HIV/AIDS monitoring and evaluation. *New Directions for Evaluation, 103*.
- Russon, C. & Russon, G. (Eds.). (2004). International perspectives on evaluation standards. *New Directions for Evaluation, 104*.

Education Update

Nadini Persaud

Websites

The World of Education <http://www.educationworld.net/> user friendly website provides links to jobs in education, world facts, education forums, a library, a web directory and a bookstore. The “world facts” link is particularly informative; it provides a database including every country on the globe. Once a country is selected, the visitor can get access to a country map and brief profiles on the geography, people, government, economy, communications, transportation, military and transnational issues for each country. The “forum link” directs the viewer to the Education American Network and Education Canada Network where various workshops on teacher-to-teacher and lessons, plans and curricula, can be found.

The [Internet Public Library Website](#) provides links to subject areas in Arts and Humanities, Business, Computers, Education, Entertainment, Health, Government, Regional, Science and Technology and Social Sciences. It also provides links to a Ready Reference database (almanacs, calendars, dictionaries) and a Reading Room (books, magazines and newspapers). The “regional link” directs the viewer to databases on history and travel and tourism by Continents/Region.

Journal Articles

Journal of Teacher Education (Volume 56: 2005 and Volume 55: 2004) has a number of interesting articles including:

Integrating Technology into Teacher Education: A Critical Framework for Implementing Reform by Valerie Otero, Dominic Peressini, Kirsten Anderson Meymaris, Pamela Ford, Tabitha Garvin, Danielle Harlow, Michelle Reidel, Bryan Waite, and Carolyn Mears [\[PDF\]](#)

The article “Integrating Technology Into Teacher Education: A Critical Framework For Implementing Reform” discusses the challenges of integrating technology into teacher education. According to the authors, teachers must be skilled in technology applications and knowledgeable about using technology in order to enhance and extend student learning. In this article, the authors present a model for technological change and also describe a critical framework to facilitate discourse among education faculty from which understandings of why, when and how to use technology emerge. Implicit in this model for technological change is a strategy for sustainability. The authors conclude that a shared vision about the role of technology in teacher education has not yet emerged in the field of education.

Teaching Under High-Stakes Testing: Dilemmas And Decisions Of A Teacher Educator by Rosemary E. Sutton

The article “Teaching Under High-Stakes Testing: Dilemmas And Decisions Of A Teacher Educator” reviews how an experienced teacher was forced to change her teaching strategies as a result of the introduction of the PRAXIS II: Principles of Learning and Teaching (PLT) tests, which are now mandated in Ohio. According to Sutton (2004), many of her students failed the PRAXIS II test when it was first introduced, because they were not good standardized test takers. She explains that as a result of the unsatisfactory pass rate on the PRAXIS II test, the Dean’s Office in the College of Education and Human Services encouraged the faculty to take the tests. Sutton (2004) notes that it was only after taking the test, that she realized that

she would have to alter her method of assessment, content and teaching strategies, in her education psychology courses to prepare students for the PRAXIS II tests. She concludes by noting that the implementation of PRAXIS II increased collaboration among faculty teaching education psychology at her University. Now instructors meet regularly to choose a common text book, share resources and discuss topics that should be incorporated into the educational psychology courses offered by the University.

Other interesting articles include

- Taking Stock in 2005: Getting Beyond the Horse Race by Marilyn Cochran-Smith [\[PDF\]](#)
- The Effect of Perceived Learner Advantages on Teachers' Beliefs About Critical-Thinking Activities by Edward Warburton and Bruce Torff [\[PDF\]](#)
- Shifting from Developmental to Postmodern Practices in Early Childhood Teacher Education by Sharon Ryan and Susan Grieshaber [\[PDF\]](#)
- Preservice Teachers Becoming Agents of Change: Pedagogical Implications for Action Research by Jeremy N. Price and Linda Valli [\[PDF\]](#)
- "Nadie Me Dijó [Nobody Told Me]": Language Policy Negotiation and Implications for Teacher Education by Manka M. Varghese and Tom Stritikus [\[PDF\]](#)
- Comparing PDS and Campus-Based Preservice Teacher Preparation: Is PDS-Based Preparation Really Better? by D. Scott Ridley, Sally Hurwitz, Mary Ruth Davis Hackett, and Kari Knutson Miller [\[PDF\]](#).

References

- Ford, P., Garvin, T., Harlow, D., Mears, C., Meymaris, K. A., Otero, V., Peressini, D., Reidel, M., Waite, B. (2005). Integrating technology into teacher education: A critical framework for implementing reform. *Journal of Teacher Education*, 56(1): 8-23.
- Sutton, R. E. (2004). Teaching under high-stakes testing: Dilemmas and decisions of a teacher educator. *Journal of Teacher Education*, 55(5): 463-475.

The Evaluation Exchange—Harvard Family Research Project

Brandon W. Youker

[Harvard Family Research Project](#) (HFRP) was founded by the Harvard Graduate School of Education in 1983. The HFRP aims to help strengthen family, school, and community partnerships of early childhood care and education; promote evaluation and accountability; and offer professional development to those who work with children and/or their families. The project has aided philanthropies, policymakers, and practitioners by collecting, analyzing, and synthesizing research and information.

HFRP's Goals:

- Develop, test, and communicate methods that promote continuous improvement and accountability
- Promote diversity, program and system complexity, and outcomes measurement and attainment through evaluation practices
- Expand and strengthen the professional development base of those who work directly with children and families
- Provide policymakers, practitioners, and foundations with research and information to guide them as they fund new strategies and strengthen existing initiatives

HFRP strives to reach its goals through providing:

- Knowledge Development
- Training and Professional Development
- Technical Assistance
- Continuous Learning and Dialogue

HFRP has two categories for research:

- a) Family-school-community partnerships
- b) Strategy consulting and evaluation

HFRP-partial list of funders:

- Carnegie Corporation of New York
- The Annie E. Casey Foundation
- The Ford Foundation
- The Heinz Endowments
- The W.K. Kellogg Foundation
- John D. & Catherine T. MacArthur Foundation
- The Charles Stewart Mott Foundation
- The Pew Charitable Trusts
- The Rockefeller Foundation

HFRP has an evaluation periodical, *The Evaluation Exchange*. The journal, published 3 or 4 times a year, addresses issues that program evaluations frequently encounter. The Evaluation Exchange emphasizes innovative methods and approaches to evaluation, emerging trends in practice, and practical applications of evaluation theory. It is designed as an ongoing discussion medium among evaluators, program practitioners, funders and policymakers. The journal is divided in to 5 sections; (1) Theory & Practice; (2) Promising Practices; (3) Spotlight; (4) Evaluations to Watch, and (5) Beyond Basic Training. Journal subscriptions are free and contributions are encouraged.

Examples of evaluation-related articles in the most recent journal publication (Volume X, No.4, Winter 2004/2005):

- “Improving Parental Involvement: Evaluating Treatment Effects in the Fast Track Program.”
- “Parental Involvement and Secondary School Student Educational Outcomes: A Meta-Analysis.”
- “Blending Evaluation Traditions: The Talent Development Model.
- “What Matters in Family Support Evaluation?”
- “Learning from Parents Through Reflective Evaluation Practice.”
- “Ongoing Evaluations of Programs in Parent Leadership and Family Involvement.”
- “Promoting Quality Outcome Measurement: A Home-Visitation Case.”

Past journal issues of particular relevance to evaluators:

Vol. X, No. 3, Fall 2004	“Harnessing Technology for Evaluation”
Vol. X, No. 2, Summer 04	“Early Childhood Programs and Evaluation”
Vol. X, No. 1, Spring 04	“Evaluating Out-of-School Time Program Quality”
Vol. IX, No. 4, Winter 03/04	“Reflecting on the Past and Future of Evaluation”
Vol. IV, No. 2, 1998	“Evaluation in the 21 st Century”
Vol. 1, No. 2, 1995	“Participatory Evaluations”

Canadian Journal of Program Evaluation, Volume 19(2), Fall 2004

Chris L. S. Coryn

The *Canadian Journal of Program Evaluation's* (CJPE) most recent issue contains 7 articles (5 in English and 2 in French), a research and practice note, and 3 book reviews.

The first article in this issue—*The Role of the Evaluator in a Political World*—is by Ernie House, and loosely based upon his keynote speech from last year's CES conference (*Evaluation 2004*) in Saskatoon, Canada. House focuses his attention on the political struggles faced by evaluators under the current Bush administration in the United States as well as the difficulties in balancing often conflicting, highly variable values and interest.

Following House's contribution is *Using Multi-Site Core Evaluation to Provide "Scientific" Evidence* by Frances Lawrenz and Douglas Huffman. The authors describe how to incorporate other evaluative purposes into evaluations in which the client's central concern is effectiveness (usually goal-based evaluation requiring experimental designs, for example). Lawrenz and Huffman argue that the prevalent standards of "scientific" evidence (such as the U.S. Department of Education's priority on "scientifically-based evaluation methods" (p. 18)—i.e., RCTs) have not been proven superior to other approaches. The authors demonstrate, through a case study, that scientific rigor can be maintained while including other evaluation approaches (in this case a participatory, collaborative approach) and purposes (to not only evaluate "what" happened, but "how").

Mary Sehl's piece entitled *Stakeholder Involvement in a Government-Funded Outcome Evaluation: Lessons Learned from the Front Line* describes involving project stakeholders in the planning and decision making process and the strengths and limitations associated with this approach.

Le Benchmarking Et L'amélioration Continue by Marthe Hurteau describes the benchmarking process and its use to identify relevant performance indicators. Hurteau uses this article to explore the "current preoccupation in evaluation literature: the impact of organizational context on the evaluation process" (p. 57).

L'évaluation des Technologies de la Santé: Comment L'introduire dans les Hôpitaux Universitaires du Québec? by Oliver Sossa and Pascale Lehoux describes the implementation of health technology in Quebec university teaching health centers and the structures that facilitate its development.

J. Bradley Cousins, Swee C. Goh, Shannon Clark, and Linda E. Lee explore the "conceptual interconnections and linkages among developments in the domains of evaluation utilization, evaluation capacity building, and organizational learning" (p. 99) in their article entitled *Integrating Evaluative Inquiry into the Organizational Culture: A Review and Synthesis of the Knowledge Base*. The implications for future research efforts and practice are also discussed.

In *The Analysis of Focus Groups in Published Research Articles* Geoffrey S. Wiggins critically assesses the use of analytic methods employed in published literature from several disciplines. The author found that fewer than half of these research articles utilized systematic analytic techniques (emergent or pre-ordinate) in assessing focus group transcripts. Even fewer utilize measures of reliability when analyzing transcripts.

The *Research and Practice Note* in this issue of CJPE is a piece by Allison Nichols, entitled *Pre- and Post-Scenarios: Assessing Learning and Behavior Outcomes in Training Settings*. Nichols frames her discussion of documenting training outcomes in learning and behavior using two case study examples.

The final section of this issue of CJPE is devoted to reviews of three books. The first is *Organizational Assessment: A Framework for Improving Performance* (Lusthaus, Adrien, Anderson, Carden, & Montalvan, 2002) reviewed by Stephen M. Morabito. The second, *Evaluating Social Programs and Problems: Visions for the New Millennium* (Donaldson & Scriven, Eds., 2003) is reviewed by Jennifer Carey. The final review, of the book entitled *Institutionalizing IMPACT Orientation: Building a Performance Management Approach that Enhances the Impact Orientation of Research Organizations* (Smith & Sutherland, 2002) is reviewed by Ronald Mackay.

**Evaluation: The International Journal of Theory, Research and Practice,
Volume 10(4), October 2004**

Daniela C. Schröter

The most recent issue of *Evaluation* contains six articles, one contribution to *A Visit to the World of Practice*, and *News from the Community* respectively.

Widmer and Neuenschwander discuss in their article—*Embedding Evaluation in the Swiss Federal Administration: Purpose, Institutional Design, and Utilization*—how evaluation is embedded within the Swiss political system and conclude that currently used evaluation measures in the government can be improved through purposeful differentiation of evaluation types. They first summarize four purposes of evaluation—accountability, improvement, basic knowledge, and strategy; secondly, five uses of evaluation—instrumental, conceptual, interactive, legitimating, and tactical; and thirdly, two predominant institutional designs in which evaluation is implemented—centralized and decentralized. Thereafter, Widmer and Neuenschwander demonstrate how evaluation is embedded within the different federal agencies of the Swiss government. Key findings of their study include that (i) accountability and improvement were the most relevant purposes in these organizational contexts; (ii) evaluation findings were most commonly utilized instrumentally, followed by legitimizing and interactive uses; (iii) the institutional design was of little or no relevance; and (iv) unanticipated blends of purpose and utilization existed.

In the second article—*Utilizing Evaluation Evidence to Enhance Professional Practice*—Helen Simons criticizes the current politically favored approach to

evaluation, namely evidence-based evaluation. She states that this approach to elucidating evidence “fails to recognize the holistic nature of professional practice and disregards the complexity of professional decision making and action” (p.410). Qualitative forms of knowledge generation for evaluative purposes would enhance the quality of evaluation and increase the utilization of evaluation findings.

In *The Meaning Assigned to Evaluation by Project Staff: Analysis from the Project-management Perspective in the Field of Social Welfare and Healthcare in Finland*, Seppänen-Järvelä examines how evaluation is understood by project staff and management and how it influences the work environment. Seppänen-Järvelä concludes that there is a need to update the current knowledge of project staff and management about evaluation and to promote and enforce evaluation culture and capacity building within organizations in the Social Welfare and Healthcare sector in Finland.

Oakley, Strange, Stephenson, Forrest, and Monteiro’s article—*Evaluating Processes: A Case Study of a Randomized Controlled Trial of Sex Education*—exemplifies the application of RCTs to evaluate how processes and outcomes are interrelated. The authors conclude that ultimately the choice of design and quantitative or qualitative approaches is context-dependent and related to the questions asked.

Following is McNamara and O’Hara’s article *Trusting the Teacher: Evaluating Educational Innovation*, in which the authors claim that the role of the external evaluator should be switched to that of the educating consultant for the teacher. The case for self-evaluation of the practitioner is supported by the argument that external evaluation would often fail to support improvement of the evaluand. To support “sound educational values,” (p. 472) evaluators should function as

facilitators and consultants in conducting the research and enhance the credibility of self-evaluation by meta-evaluating the internal self-evaluation processes.

The last article, authored by Hanberger and Schild, discusses *Strategies to Evaluate a University-Industry Knowledge-exchange Programme*. The authors consider two management-oriented approaches to program evaluation (program theory evaluation and outcome analysis) and two non-management oriented approaches (policy discourse analysis and qualitative network analysis). They conclude that different evaluation methods stress the values of different stakeholder groups. An integration of various methods is necessary, reduces the bias toward one stakeholder group, and increases the validity in contexts where multiple stakeholder groups are present. In situations with only one target group and few stakeholders, a combination of various evaluation approaches would not be as essential.

In *A Visit to the World of Practice*, Farrall and Gadd address *Evaluating Crime Fears: A Research Note on a Pilot Study to Improve the Measurement of the 'Fear of Crime' as a Performance Indicator*. Instruments intended to assess fear of crime are criticized for their poor design and neglect of crucial research concerns such as frequency and intensity. The authors suggest survey questions to be incorporated to improve instruments measuring fear of crime.

In *News from the Community*, Nicoletta Stame, President of the EES, reports on the 6th EES conference entitled *Governance, Democracy and Evaluation*. These issues include: (i) evaluation as a tool for democratic government, (ii) the question of an European evaluation identity, (iii) European standards for evaluation, (iv) relationships among evaluation networks and associations, and (v) training, education, and professional development of evaluation in Europe. For example, the

EES considers the establishment of European evaluation standards but does not want to conflict or overthrow features unique to individual national characteristics.

The journal concludes with translations of the article's abstracts into French and an *Annual Index* of articles.

**Measurement: Interdisciplinary Research and Perspectives, Volume 1(1),
2003**

Chris L. S. Coryn

Measurement: Interdisciplinary Research and Perspectives is a relatively new journal devoted to the interdisciplinary study of measurement in the human sciences and is intended to represent a broad range of disciplines and perspectives including psychometrics, ethnography, social theory, psychology, economics, education, linguistics, sociology, policy studies, history, and law. Each issue is devoted to a single, provocative focus article followed by commentaries and a rejoinder article. Further information can be found at <http://bearcenter.berkeley.edu/measurement/>. Presently eight issues are available covering objectivity and trust, standards-based testing, and certification testing, for example. The inaugural issue—Volume 1(1), 2003—was sent to us by the editors (Mark Wilson at the University of California, Berkeley; Paul De Boeck at K. U. Leuven, Belgium; and Pamela Moss at the University of Michigan) to encourage becoming involved in for example, debating a “focus paper” or participating in a commentary.

The focus article of the inaugural issue is *On the Structure of Educational Assessments* by Robert J. Mislevy, Linda S. Steinberg, and Russell G. Almond. This article describes a framework for assessment that makes explicit the interrelations among substantive arguments, assessment designs, and operational processes. This framework, called “evidence-centered” assessment design (ECD)

“...entails the development, construction, and arrangement of specialized information elements, or assessment design objects, into specifications that embody the substantive arguments that underlies an assessment” (p. 4).

The authors illustrate their ideas with examples from language testing and the article is presented parallel to the stages of the ECD design process: (1) domain analysis, (2) domain modeling, (3) conceptual assessment framework, and (4) operational assessment—the “four-process delivery system.”

In the first stage of ECD design, domain analysis, information about the domain is used to organize beliefs, theories, research, subject-matter expertise, instructional materials, and exemplars. In the second stage, domain modeling, the information gathered in stage 1 is organized into three paradigms; proficiency, evidence, and tasks. The third task, developing a conceptual assessment framework (CAF), specifies the technical details necessary for implementing the assessment; specifications, operational requirements, statistical models, and rubrics, for example. The final stage, the four-process delivery system, consists of four principal components: (1) activity selection, (2) presentation, (3) evidence identification—task-level scoring, and (4) evidence accumulation—test-level scoring.

Mislevy, Steinberg, and Almond’s intricate approach emphasizes measurement models which incorporate the relationship between “assessment purposes, substantive experience and theory, statistical models and task authoring schemas, and the elements and processes of operational models” (p. 56).

Eight commentaries follow the focus article which range from the limitations of Bayesian models in assessment (Earl Hunt) to critiques and comments from a

psychometric perspective (Cees A. W. Glas) and finally to a framework for shifting from principle to practice (Richard K. Wagner). The issue concludes with Mislevy, Steinberg, and Almond's rejoinder in which the authors address the themes of critique presented in the commentaries; constructivist and situative learning perspectives, model generalizability, user and statistical models, and implementation.