

# **JMDE**

## **Journal of MultiDisciplinary Evaluation**

Number 3, October 2005

ISSN 1556-8180

### **Editors**

E. Jane Davidson & Michael Scriven

### **Associate Editors**

Chris L. S. Coryn & Daniela C. Schröter

### **Assistant Editors**

Thomaz Chianca

Nadini Persaud

Lori Wingate

Ryo Sasaki

Brandon W. Youker

### **Webmaster**

Joe Fee

## **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 II

### Global Review: Regions and Events

National and Regional Evaluation Networks.....	150
<i>IOCE</i>	
A Call to Action: The First International Congress of Qualitative Inquiry.....	155
<i>Chris L. S. Coryn, Daniela C. Schröter, &amp; Michael Scriven</i>	
Evaluation in Canada .....	166
<i>Chris L. S. Coryn</i>	
Evaluation in the People’s Republic of China .....	172
<i>Xuejin Lu and Donghai Xie</i>	
Evaluation in Germany: An Overview .....	180
<i>Gerlinde Struhkamp</i>	
Evaluation—Making it Real in Aotearoa New Zealand: Leading by Example, Leading by Association.....	195
<i>Pam Oliver, Kate McKegg, Geoff Stone, and Maggie Jakob-Hoff</i>	
A Review of the Chinese National Center for Science and Technology Evaluation .....	197
<i>Laura Pan Luo</i>	
Evaluation in Japan .....	200
<i>Ryo Sasaki</i>	

## **Global Review: Regions and Events**

### **National and Regional Evaluation Networks**

#### IOCE

The following list of national and regional evaluation networks was obtained from the International Organization for Cooperation in Evaluation (IOCE) at <http://ioce.net>. The IOCE is an organization for evaluation networks and societies that is committed to building a worldwide evaluation community.

#### **Evaluation Networks with Websites**

- **African Evaluation Association** [www.afrea.org/](http://www.afrea.org/)
- **American Evaluation Association** <http://www.eval.org/>
- **Australasian Evaluation Society** <http://www.aes.asn.au/>
- **Brazilian Evaluation Association** [www.avaliabrasil.org.br](http://www.avaliabrasil.org.br)
- **Canadian Evaluation Society** <http://www.evaluationcanada.ca/>
- **Danish Evaluation Society** <http://www.danskevalueringsselskab.dk>
- **Dutch Evaluation Society** <http://www.videnet.nl/>
- **European Evaluation Society** <http://www.europeanevaluation.org/>

- **Finnish Evaluation Society** <http://www.finnishevaluationsociety.net/>
- **French Evaluation Society** <http://www.sfe.asso.fr/>
- **German Evaluation Society** <http://www.degeval.de/>
- **International Program Evaluation Network** (Russia & Newly Independent States) <http://ipen21.org/ipen/>
- **Israeli Association for Program Evaluation** <http://www.iape.org.il>
- **Italian Evaluation Society** <http://www.valutazioneitaliana.it/>
- **Japan Evaluation Society** [http://www.idcj.or.jp/jes/index\\_english.htm](http://www.idcj.or.jp/jes/index_english.htm)
- **Latin American and Caribbean Programme for Strengthening the Regional Capacity for Evaluation of Rural Poverty Alleviation Projects** (PREVAL) <http://www.preval.org/>
- **Malaysian Evaluation Society** <http://www.mes.org.my>
- **Niger Network of Monitoring and Evaluation** (ReNSE) [www.pnud.ne/reuse/](http://www.pnud.ne/reuse/)
- **Polish Evaluation Society** <http://www.pte.org.pl/obszary/enginfo.htm>
- **Quebec Society for Program Evaluation** <http://www.sqep.ca>
- **South African Evaluation Network** (SAENet) [www.afrea.org/webs/southafrica/](http://www.afrea.org/webs/southafrica/)
- **Spanish Evaluation Society** <http://www.sociedadevaluacion.org/>
- **Swedish Evaluation Society** <http://www.svuf.nu>
- **Swiss Evaluation Society** <http://www.seval.ch/de/index.cfm>
- **Uganda Evaluation Association** (UEA) [www.ueas.org](http://www.ueas.org)

- **United Kingdom Evaluation Society** <http://www.evaluation.org.uk/>
- **Wallonian Society for Evaluation** (Belgium) [www.prospeval.org](http://www.prospeval.org)

## **National and Regional Evaluation Networks without Websites**

- **Bangladesh Evaluation Forum**, Syed Tamjid ur Rahman,  
[tamjidr@bangla.net](mailto:tamjidr@bangla.net)
- **Benin**, Maxime Dahoun, [mdahoun@yahoo.fr](mailto:mdahoun@yahoo.fr), or [francois-corneille.kedowide@iucn.org](mailto:francois-corneille.kedowide@iucn.org)
- **Botswana Evaluation Association**, Kathleen Letshabo,  
[letshabo@mopipi.ub.bw](mailto:letshabo@mopipi.ub.bw)
- **Burkina Faso M&E Network**, Marie-Michelle Ouedraogo,  
[mmouedraogo@unicef.org](mailto:mmouedraogo@unicef.org)
- **Burundi Evaluation Network**, Deogration Buzingo,  
[buzingdeo@yahoo.com](mailto:buzingdeo@yahoo.com)
- **Cameroon Development Evaluation Association** (CaDEA), Debazou Y. Yantio, [yantio@hotmail.com](mailto:yantio@hotmail.com)
- **Cape Verde**, Francisco Fernandes Tavares, [Francisco.Tavares@ine.gov.cv](mailto:Francisco.Tavares@ine.gov.cv)  
or [chicotavares@yahoo.com.br](mailto:chicotavares@yahoo.com.br)
- **Central American Evaluation Association**, Johanna Fernandez,  
[johannaf@cariari.ucr.ac.cr](mailto:johannaf@cariari.ucr.ac.cr)
- **China**, Chaoying Chen, [chenzhaoying@ncste.org](mailto:chenzhaoying@ncste.org)
- **Columbian Network for Monitoring and Evaluation**, Consuelo Ballesteros [consocds@colomsat.net.co](mailto:consocds@colomsat.net.co) or Daniel Gomez  
[dgomez@uniandes.edu.co](mailto:dgomez@uniandes.edu.co)

- **Egyptian Evaluation Society**, Ashraf Bakr, [picardm@care.org](mailto:picardm@care.org)
- **Eritrean National Evaluation Association**, Bissrat Ghebru, [bissratgk@asmara.uoa.edu.er](mailto:bissratgk@asmara.uoa.edu.er) or Woldeyesus Elisa, [dolab@eol.com.er](mailto:dolab@eol.com.er)
- **Ethiopian Evaluation Association**, Gizachew Bizayehu, [medac2@telecom.net.et](mailto:medac2@telecom.net.et)
- **Ghana Evaluation Network (GEN)**, Charles Nornoo, [cnornoo@internetghana.com](mailto:cnornoo@internetghana.com) or [bds@africanus.com](mailto:bds@africanus.com)
- **Ghana Evaluators Association**, [isodec@ghana.com](mailto:isodec@ghana.com)
- **Indian Evaluation Network**, Suresh Balakrishnan, [sbalakrishnan@vsnl.net](mailto:sbalakrishnan@vsnl.net)
- **Kenya Evaluation Association**, Gitonga Mburugu Nkanata, [gitonga35@avu.org](mailto:gitonga35@avu.org) or Karen Odhiambo, [karenodhiamboo@hotmail.com](mailto:karenodhiamboo@hotmail.com)
- **Korean Evaluation Association**, Sung Sam Oh, [edulove@kkucc.konkuk.ac.kr](mailto:edulove@kkucc.konkuk.ac.kr)
- **Madagascar**, Barbara Rakotoniaina, [Barbara.Rakotoniaina@caramail.com](mailto:Barbara.Rakotoniaina@caramail.com) or Dominique Wendling, [Aea.evaluation@netcourrier.com](mailto:Aea.evaluation@netcourrier.com) or [aea.evaluation@yahoo.fr](mailto:aea.evaluation@yahoo.fr)
- **Malawi Network of Evaluators**, John Kadzandira, [csrbasis@malawi.net](mailto:csrbasis@malawi.net) or [csr@malawi.net](mailto:csr@malawi.net)
- **Mauritanian M&E Network**, Ba Tall Oumoul, [oktconsult@yahoo.fr](mailto:oktconsult@yahoo.fr) or Mohammeden Fall, [mfall@unicef.org](mailto:mfall@unicef.org)
- **Namibia Monitoring Evaluation and Research Network**, Bob Hochobeb, [bhochobeb@unam.na](mailto:bhochobeb@unam.na)
- **Nepal M&E Forum**, Suman Rai, [srai@icimod.org.np](mailto:srai@icimod.org.np)

- **Nigeria**, Adam Suleiman, [adamsuleiman@yahoo.com](mailto:adamsuleiman@yahoo.com) (interested in establishing a network)
- **Perú Network for Monitoring and Evaluation**, Emma Rotondo, [erotondo@terra.com.pe](mailto:erotondo@terra.com.pe)
- **Red de evaluacion de America Latina y el Caribe. (ReLAC)**, [contacto\\_relac@yahoo.com](mailto:contacto_relac@yahoo.com)
- **Rwanda Network for Monitoring and Evaluation**, James Mugaju, [imungaju@unicef.org](mailto:imungaju@unicef.org) or Philippe Ngango Gafishi, [pgafishi@yahoo.fr](mailto:pgafishi@yahoo.fr)
- **Senegalese Network of M&E**, Eric d Muynck, [eric.de.muynck@undp.org](mailto:eric.de.muynck@undp.org)
- **Spanish Evaluation Society**, Carmen Vélez Méndez, [carmenvelez@idr.es](mailto:carmenvelez@idr.es) or Carlos Román del Río, [carlosroman@idr.es](mailto:carlosroman@idr.es)
- **Sri Lanka Evaluation Association**, Indra Tudawe, [sleva@sltnet.lk](mailto:sleva@sltnet.lk) or Ira Thabrews, [mrthab@dynaweb.lk](mailto:mrthab@dynaweb.lk)
- **Thailand Evaluation Network**, Rangsun Wiboonuppatum, [rangsun@hotmail.com](mailto:rangsun@hotmail.com)
- **Zambia Evaluation Association (ZEA)**, Greenwell Mukwavi, [gmukwavi@zamtel.zm](mailto:gmukwavi@zamtel.zm) or [twizamtc@zamnet.zm](mailto:twizamtc@zamnet.zm)
- **Zimbabwe Evaluation Society**, Mufunani Tungu Khosa, [mkhosa@mandel.co.zw](mailto:mkhosa@mandel.co.zw) or [emkhosa@ecoweb.co.zw](mailto:emkhosa@ecoweb.co.zw)

## **A Call to Action: The First International Congress of Qualitative Inquiry**

Chris L. S. Coryn, Daniela C. Schröter, & Michael Scriven

Around the globe governments are attempting to regulate interpretive inquiry by enforcing biomedical, evidence-based models of research. These regulatory activities raise basic philosophical, epistemological, political and pedagogical issues for scholarship and freedom of speech in the academy. Their effects are interdisciplinary. They cut across the fields of educational and policy research, the humanities, communications, health and social science, social welfare, business and law.

(Denzin, 2005a)

The First International Congress of Qualitative Inquiry, held at the University of Illinois at Urbana-Champaign from May 5-7, 2005, was assembled so that the international community of qualitative researchers could address the implications of attempts by federal funding agencies to “regulate scientific inquiry by defining what is good science” (Denzin, 2005b). The Congress was attended by more than 800 persons from more than 45 nations. More than 160 sessions consisting of more than 650 papers authored by more than 750 persons were presented. The complete Congress program, including session and paper abstracts, complete papers, and other information is available at <http://www.qi2005.org/>.

JMDE visited the conference to learn more about the ongoing debate regarding evidence-based science and policy and cutting-edge qualitative methodologies. Following are brief overviews of Congress panels and sessions attended.

## **Opening Keynote Addresses**

There were two opening keynote addresses introduced by Norman K. Denzin. The first of these was Janice Morse's "The Politics of Evidence." As Morse (2005) argued "evidence, by definition is definite, hard, indisputable, unchanging" and "yet, what counts as evidence, what we are willing to consider as evidence, and, most importantly, what we are willing to consider constitutes evidence, is fickle, irrational, and arbitrary." She went on to explain that the "criteria for defining evidence and the means by which it is accrued, is selected by passive agreement, often unchallenged, and supported by mainstream academia, policy makers and government" (Morse, 2005). For evidence-based research, the Cochran criterion has long been the standard for what is applicable and acceptable in research, resulting in the exclusion of qualitative research from funding. Morse then explained how the qualitative community responded throughout the years. Key strategies included for example appeals, the development of qualitative meta-analysis, and mixed methods approaches that demonstrate efficacy by using logic and common sense. Furthermore, Morse presented alternative methodologies including: (i) forensic designs for cases in which "near misses" are investigated, that is, the incident under investigation has not yet occurred and outcomes are hypothetical, thus, the converse to statistical significance and devoid of quantitative criteria; (ii) trials of interventions, that is, microanalysis of rare events that are experimental, but where outcomes are unknown; (iii) observations and precise micro-analytic observational descriptions, and (iv) simulations of high risk situations.

Linda Tuhiwai Smith presented the second keynote address, “On Tricky Ground: Researching the Native in the Age of Uncertainty,” in the form of stories from her own and other’s experiences. Smith illustrated the ‘tricky ground’ that fills the spaces “between research methodologies, ethical principles, institutional regulations and human subjects as individuals and as socially organized actors and communities” (Smith, 2005). She further asserted that “this ground is richly nuanced in terms of diverse interests through epistemological challenges to research, to its paradigms, practices and impacts” and “in this context—building on what indigenous communities have struggled for, tried to assert and have achieved—what is possible in the application of indigenous perspectives that examine the intersections of methods, ethics, institutions and communities” (Smith, 2005).

## **Plenary Sessions**

### ***Science, Etc.: From Bicycle Helmets to Dialogue Across Differences***

*Chair:* Elizabeth St. Pierre

*Panelists:* Michael J. Feuer, Lisa Towne, and Elizabeth St. Pierre

This plenary session was a friendly debate between Michael Feuer of the National Academy of Science (NAS) and Elizabeth St. Pierre of the University of Georgia.

Feuer started out the presentation by providing a brief description of the development and history of the National Academy of Sciences and the National Research Council and then devoted considerable time to defending the NAS and NRC in guiding and informing the federal government in “science policy” and “science-based policy.” Feuer claimed that science is only objective and independent if it is not paid for. Therefore, both the NRC and the NAS are independent of the government and must, if called for advice, be “faithful” to data,

to evidence. Interdisciplinary committees are invited to engage in a process of evidence-based consensus building which is to affect federal law and policy decision making. Keys to decision making include appreciation and understanding of standards of evidence as well as the appropriateness of the level of evidence, which if set to high, thwarts decision making.

The diversity of interests considered is reflected in reports published by the National Academy Press (see <http://www.nap.edu>). One NRC report specifically referred to “Advancing scientific research in education” (see executive summary [http://www.nap.edu/execsumm\\_pdf/11112.pdf](http://www.nap.edu/execsumm_pdf/11112.pdf)) and was build on the report on “Scientific research in education” which defined what “high-quality scientific inquiry” is or should be (see executive summary [http://www.nap.edu/execsumm\\_pdf/11112.pdf](http://www.nap.edu/execsumm_pdf/11112.pdf)).

St. Pierre was introduced as the “extreme postmodernist” and started out by referencing well-known postmodern theorists including Jacque Derrida, Judith Butler, and Michael Foucoult. St. Pierre found the NRC report offensive and stated that the government is “narrowing science” and “the current definition of science is grounded in positivism.” She also claimed that “science is not the same in all paradigms in terms of ontology, epistemology, and methodology” and that “the rage of causation is nothing more than an attempt at meaning making.” Evidence-based research was not only pointed out to being “dangerous” because it narrows science but because it is based on power, politics, and economy. St. Pierre emphasized that it is essential to consider epistemologies in science to understand the limitations research. For example, science could not capture lived reality; instead it is everywhere, does not have an identity, and is always in the making.

***Monsters of Evidence: Qualitative Research and the Globalization of Audit Culture***

*Chair:* Patti Lather

*Panelists:* Patti Lather, Lis Hojgaard, Dorte Marie Sondergaard, Ian Stronach, Harry Torrance, and Phil Hodkinson

In this session, presenters from Denmark, the United Kingdom, and the United States reflected on evidence-based research under different cultural traditions. The Scandinavian presenters described the arrival of evidence-based research in Scandinavia and called for elaboration and redefinition of the term “evidence.” Evidence-based research is perceived as one single method, which not only limits the questions asked but also the answers elucidated, thus leading to a knowledge gap.

Stronach focused on the gap between rules/regulations and reality, leading to circularity and suppressed nucleation of research, the “either/or;” while Torrance discussed the shift in locus of control, questioning who defines and controls research in society. Torrance claimed that the managerial audit culture hurts the quality of research, which is evaluated and judged based on its management rather than on its intrinsic value.

Hodkinson discussed the return of positivism, specifically in the United States with regard to educational research. He pointed out that learning is a contested social construct and that acquisition views dominate learning. However, acquisition perspectives view learning as an outcome, leading to the neglect of the learning process. Moreover, evidence-based research would only view measurable outcomes as significant. The application of post-positivist objectives to learning would result in the following paradox issues: (i) there is no independent variable,

(ii) noise matters, and (iii) objectivity is biased. Evidence-based research would not include the methods that bring truth.

### ***‘Scientifically Based Research’ and Qualitative Research Methodologies***

*Chair:* Katherine Ryan

*Panelists:* Yvonna S. Lincoln, Earnest House, Julianne Cheek, Frederick Erickson, Nicholas Burbules, and Ian Stronach

Each of the presentations in this plenary session focused on differing aspects of scientifically based research and qualitative methodologies.

Burbules attempted to look “Beyond Method,” and emphasized that researchers need to clarify (i) value propositions, (ii) the field they are from including outspoken critics of that field, and to accept (iii) consequences of their research. This includes an understanding of cultures of inquiry and epistemological virtues. Epistemological virtues involve intellectual and moral qualities. For example, tolerance of alternative methodological and ethical approaches to research are the underlying necessities for objectivity. Fallibalism on the other hand, is the virtue that researchers leave room for failure and admit it when they experience it, thus, fallibalism functions as a change initiator and agent. Questions posted at the end included: where do epistemological questions come from? What good are methods without epistemological virtues? And how do epistemological virtues generate debates?

Katherine Ryan’s presentation emphasized the old and new scientism and argued that “evidence is not evident.” Moreover, she asserted that the reemergence of positivism can be attributed to the audit culture.

Lincoln discussed qualitative methodology and social justice. She illuminated five recent trends in the social science community: (1) there is more openness regarding social justice, (2) qualitative methods are deployed to collect the construction of marginal groups, (3) there is willingness to utilize opinions of marginal groups to pose research questions, planning, and conducting research, (4) there are active advocates for the poor and other marginal groups, and (5) false neutrality is abandoned. Moreover, Lincoln provided three suggestions regarding the qualitative/quantitative debate: (i) be available to discuss and be tolerant of different and alternative methodologies, (ii) senior staff should team up with junior staff to thwart anxiety of doing qualitative research prior to tenure, and (iii) colonize them.

House provided an overview of 40 years of (policy) evaluation and pointed out developments in the perception of causation, from regularity based causation to complexity of causation. Moreover he constructed an analogy in the current evidence-based debate to the existing neo-fundamentalism prevalent in the United States. The golden standard for causation provides researchers with only one source of truth that is described Campbell and Stanley. Therefore, research is limited in accessibility, prophetic in its vision about the future, and closed to other ideas. Moreover, the fundamentalism is marked by listening to only those who share the same ideas, by rejecting others, and by persuasion through coercion. The methodological fundamentalism would be marked by blacklisting those who can(not) do research, by a shift from the Cochran to the Campbell regime, and by not listening to others.

## **Special Featured Panels**

### ***Why Measurement Fails***

*Presenters:* Jaber Gubrium and James Holstein

Does measurement fail? Gubrium and Holstein suggest that it does. In fact, the presenters argued that (1) measurement can't capture interactions, (2) freezes context, (3) reifies meaning, and (4) requires fixed variables. Grounded in the sociological literature (e.g., Mayhew, Znaniecki, Whyte, Cicourel, and Rose, among others) and exemplified through discourse analysis of court conversations, the authors assert that the issue of measurement's failures needs to be revisited because the concerns have not been resolved, that we live in a "measurement society," in which applied concerns ignore the issue of the seriousness of empirical reality. Their most compelling argument, however, was that measurement fails to account for differing meaning for different groups of persons and does little to account for context, which is defined differently for qualitative and quantitative researchers.

## **General Sessions**

### ***Mixed and Mixed-Up Methods***

*Chair:* Ian E. Baptiste

*Presenters:* Ian E. Baptistse, Ljiljana Vuletic, Michel Ferrari, Marina Micari, Susanna Calkins, Melissa Luna, Greg Light, and C. Mimi Harvey

Unfortunately, only three of the eight presenters showed-up for this session; Ian E. Baptiste, Marina Micari, and Susanna Calkins. Baptiste's paper titled "Mixed and Mixed-Up Methods: Reconceptualizing Mixed-Methods Design" was an expose on what "constitutes a method." That is, the author argued that a procedure qualifies

as a method once it incorporates some strategy or strategies for collecting words or numbers and that words are qualitative whereas numbers are quantitative. Moreover, Baptiste argued that research has four analytic interests, each with corresponding methods. These were:

1. Identify and measures associations—with the corresponding methods being correlational studies and quasi-experiments
2. Explore phenomena—with the corresponding methods being qualitative research methods
3. Establish cause—with the primary methods being experiments and quasi-experiments
4. Describe frequency distributions—with the corresponding method being surveys

Micari and Calkins presented “Achieving Accountability in Education: Phenomenography as Research-Based Evaluation,” in which they described an evaluation which employed phenomenography in addition to a variety of other methods to evaluate an education program. Phenomenography was described as “the empirical study of the limited number of qualitatively different ways in which we experience, conceptualize, perceive, or apprehend various phenomena.”

### ***IRBs and the Politics of Informed Consent***

*Chair:* Gaile S. Cannella

*Presenters:* R. Wiles, G. Crow, S. Heath, V. Charles, Stephen J. Sills, Bart W.

Miles, Amy E. Blank, Barbara F. Sharf, M. Carolyn Clark, and Marco Marzano

Wiles, Crow, Heath, and Charles presented “Research Ethics and Regulations in the UK: The Case of Informed Consent.” The authors conducted research of

researchers regarding the increased enforcement of regulated informed consent in the UK, and how their subjects positioned themselves in relation to these issues.

Sills and Miles discussed their study “Investigating Visual Researchers’ Experiences with Institutional Review Boards.” The authors conducted survey research with qualitative, visual researchers in academic institutions and found that researchers’ experiences with IRBs varied widely in terms of perceived quality and satisfaction with the IRB process.

Blank, a doctoral student in a traditionally quantitative department, discussed the process of attaining IRB approval for her dissertation research in “The IRB’s Role in Ethnography of Vulnerable Populations: Protection of the Subject or Protection of the Paradigm?”

“The Dark Side of Truth(s): Ethical Quandries in Accessing and Reporting Qualitative Analysis of Life Stories” presented by Sharf and Clark discussed their research in female prison populations. The authors presented a number of difficulties in their research as it related to ethics and IRBs. Primarily, the authors struggled with their research subjects revealing information with the portent for creating ethical dilemmas. Furthermore, the authors argued that IRBs do not meet the needs of qualitative researchers and are stuck in the positivist, medical model frame of mind regarding ethics and research.

Marzano discussed “Towards Ethical Globalization? Freedom of Research and Moral Constraints in Qualitative Research,” in which he shared his experiences conducting ethnographic research in a hospital. This research required that the researcher “go undercover,” that is, he dressed and acted as a medical professional in order to conduct research on medical professionals.

## **The Second International Congress of Qualitative Inquiry**

The Second International Congress of Qualitative Inquiry is scheduled to take place from May 4-7, 2006 at the University of Illinois at Urbana-Champaign. Additional information is available at <http://www.c4qi.org/qi2006.html>.

### **References**

Denzin, N. K. (2005a). The first international congress of qualitative inquiry.

Available at <http://www.qi2005.org/DenzinICQI.pdf>

Denzin, N. K. (2005b). Welcome from the director. *First international congress of qualitative inquiry: Official program, panel abstracts, individual abstracts, and general information*. Urbana-Champaign, IL: University of Illinois at Urbana-Champaign.

Morse, J. M. (2005). The politics of evidence. Abstract available at

<http://www.qi2005.org/plenaries.html>

Smith, L. T. (2005). On tricky ground: Researching the native in the age of

uncertainty. Abstract available at <http://www.qi2005.org/plenaries.html>

## Evaluation in Canada

Chris L. S. Coryn

The Canadian provinces continue to be a source of evaluation-related activities and events such as Evaluation 2005, Beaulac, Goodine, and Aubry's work on a report card of homelessness in Ottawa, the 2005 International Program for Development Evaluation jointly sponsored by the World Bank Group and Carleton University, and the Canadian Evaluation Society Student Case Competition and Paper Contest, to name but a few. For those interested in detailed information on these and other Canadian evaluation news and events please visit the [Canadian Evaluation Society Website](#).

### **Evaluation 2005: The Joint American Evaluation Association/Canadian Evaluation Society Conference**

By early accounts the upcoming joint conference—[Evaluation 2005](#)—sponsored by the American Evaluation Association and the Canadian Evaluation Society to be held in Toronto, Ontario, Canada from October 24-October 30, 2005 promises to be a great success. A recent news release from the Canadian Evaluation Society indicated that

A total of 1,206 proposals were submitted from representatives of 43 countries. Some 879 proposals were from United States representatives, 200 from Canada and 127 from other countries.

About 17% of proposals are from Canadians; this compares to about 3% in the past 3 years of proposals to the AEA annual conference (which were not joint conferences with the CES). Overall, there are also 50% more proposals submitted to the 2005 joint conference than there have been in the 3 most recent years of AEA conferences. All of this to say that there will be a lot to select from and that the content of Evaluation 2005 will certainly be of very high caliber.

(Canadian Evaluation Society, 2005a)

## **The Alliance to End Homelessness**

Earlier this year the Center for Research in Community Services at the University of Ottawa released the [Report Card Methodology and Indicators: Development of the Report Card of Homelessness in Ottawa](#) (Beaulac, Goodine, & Aubry, 2004) prepared for the Alliance to End Homelessness in Ottawa. The report is divided into two parts; Part I—A Review of the Literature and Part II—Indicators and Canadian Report Card. Based on a review of relevant literature, this overview of the methodological aspects on the development of report cards was undertaken as the preliminary work for the development of the report card on homelessness in Ottawa. The purpose of this report is to provide a brief overview of the literature on report card methodology, including the history and current status of report cards, the purposes and processes of developing and formulating report cards, the dissemination and translation of report cards, and suggestions for the Ottawa report card on homelessness in light of the findings uncovered in the literature review.

## **2005 International Program for Development Evaluation**

The fifth annual International Program for Development Evaluation: Building Skills to Evaluate Development Interventions is designed to meet the professional development needs of mid-level evaluation and audit professionals working in

developed and developing nations, development agencies, and non-government organizations. The program was jointly sponsored by The World Bank Operations Evaluation Department and Carlton University's Faculty of Public Affairs and Management and was held at Carlton University, Ottawa, Canada from June 13 through July 8, 2005. It offered a two-week core course consisting of 80 hours of instruction in essential tools and techniques, current lessons from the field, expert guidance, and practice in developing evaluation plans and designs. The core course curriculum was followed by two-weeks of 26 free-standing workshops on various topics and themes specific to development evaluation. For additional information please visit the [International Program for Development Evaluation Training Website](#). Fees ranged from US \$2,132 through US \$9,952 and room and board was available (included in some fee schedules).

### **Canadian Evaluation Society Student Case Competition**

The final round of the annual CES Case Competition for 2005 was held on May 14, 2005 at Carleton University. The final round teams were [Right Approach Consulting](#) (University of Ottawa, Education), [QuickStar Consulting](#) (University of Waterloo, Applied Health Sciences) and [Transformations](#) (Georgian College, Research Analyst Program).

The teams had five hours to prepare an evaluation case before presenting it to the judging panel and audience (Canadian Evaluation Society, 2005b). Teams were each given thirty minutes for a presentation, followed by a ten minute question period for the judges. This year's judging panel featured evaluation experts from both the public and private sector who donated their time and effort to adjudicate both rounds of the competition. The 2005 judges were Marc L. Johnson, Consultant, Research and Evaluation; Susan Morris, Chief, Evaluation, Natural

Sciences and Engineering Research Council of Canada; and Martine Perrault, Consultant Manager, Goss Gilroy Inc (Canadian Evaluation Society, 2005b).

For the first time in the history of the CES Case Competition the judges announced a tie, between QuickStar and Transformations for the 2005 competition.

For additional information on the annual CES Case Competition please see Coryn (2004) or visit the [CES Case Competition Website](#).

### **Canadian Evaluation Society Student Paper Competition**

Each year the CES conducts a student paper contest. The contest is intended to provide exposure to promising Canadian students who study or have an interest in evaluation. Awards are granted for the best paper written by a post-secondary student in the field of evaluation. The winner of the 2005 CES student paper competition was Michelle Anderson-Draper, Faculty of Agriculture, Forestry and Home Economics, University of Alberta. Her paper, titled “Understanding cultural competence by evaluating “Breaking the silence: A project to generate critical knowledge about family violence within immigrant communities,” examined

...the concept of cultural competence for evaluators by presenting the evaluation of “Breaking the silence: A project to generate critical knowledge about family violence within immigrant communities” as a case study. Using data from monthly facilitated discussions, findings indicate participants furthered their knowledge about the issue of family violence and received information to assist them in their work with immigrant families. Constructs from the Social Cognitive theory and the PRECEDE-PROCEED model provide the framework for the planning, implementation and evaluation of this project. Experiences of the internal evaluator in relation to cultural competency are explored.

(Canadian Evaluation Society, 2005c)

The CES 2005 student paper competition honorable mention went to Kelly Skinner, Health Studies and Gerontology, University of Waterloo, for her paper titled “Developing a tool to measure knowledge exchange outcomes.” The paper

...describes measures to assess outcomes of efforts to encourage use of better practices in chronic disease prevention (CDP). A CDP better practices model (Moyer et al., 2002) consists of knowledge synthesis, knowledge exchange (dissemination / adoption) and evaluation stages. Better practices are required at each stage. No previous knowledge syntheses of tools and models for evaluating the efficiency and effectiveness of the dissemination/exchange strategies were found. This project developed a usable model and specific scales to assess knowledge exchange efforts for best practices in type 2 diabetes prevention. The model can be adapted to other areas of population health.

(Canadian Evaluation Society, 2005d)

For additional information on the annual CES Student Paper Competition please see Coryn (2004) or visit the [Student Competitions](#) section of the CES Website.

## **References**

Beaulac, J., Goodine, L., & Aubry, T. (2004). Report card methodology and indicators: Development of the report card on homelessness in Ottawa.

Available at

[http://www.evaluationcanada.ca/distribution/200408\\_beaulac\\_julie\\_goodine\\_laura\\_aubry\\_tim.pdf](http://www.evaluationcanada.ca/distribution/200408_beaulac_julie_goodine_laura_aubry_tim.pdf)

Canadian Evaluation Society (2005a). A flurry of proposals made to Evaluation 2005. Available at

[http://www.evaluationcanada.ca/site.cgi?s=1&ss=1&\\_lang=an&num=00502](http://www.evaluationcanada.ca/site.cgi?s=1&ss=1&_lang=an&num=00502)

Canadian Evaluation Society (2005b). 2005 CES case competition final round.

Available at

[http://www.evaluationcanada.ca/site.cgi?s=1&ss=1&\\_lang=en&num=00525](http://www.evaluationcanada.ca/site.cgi?s=1&ss=1&_lang=en&num=00525)

Canadian Evaluation Society (2005c). Winner, student essay award, 2005.

Available at

[http://www.evaluationcanada.ca/site.cgi?s=4&ss=5&\\_lang=en&prixn=Anderson-Draper&prixp=Michelle&code\\_de\\_type=5&annee=2005](http://www.evaluationcanada.ca/site.cgi?s=4&ss=5&_lang=en&prixn=Anderson-Draper&prixp=Michelle&code_de_type=5&annee=2005)

Canadian Evaluation Society (2005d). Honorable mention, student essay award, 2005. Available at

[http://www.evaluationcanada.ca/site.cgi?s=4&ss=5&\\_lang=en&prixn=Skinner&prixp=Kelly&code\\_de\\_type=5&annee=2005](http://www.evaluationcanada.ca/site.cgi?s=4&ss=5&_lang=en&prixn=Skinner&prixp=Kelly&code_de_type=5&annee=2005)

Coryn, C. L. S. (2004). The state of evaluation in Canada. *Journal of MultiDisciplinary Evaluation*, 1, 55-68.

## Evaluation in the People's Republic of China

Xuejin Lu and Donghai Xie

Original evaluation practice in China can be dated back to the ancient time of 2200 B.C when the Chinese used essay examinations to help select civil service employees or to choose the most talented learner to serve in the civic administration ( Drummond, 2003). Modern evaluation practice is still mostly succeeding the early practice that government-sponsored evaluation plays a decisive role in evaluating all kinds of national development activities. Evaluation conducted by NGO (non-government organization) has not yet exercised any influence on the current evaluation practice. In the past two decades, China has seen rapid growth of a significant number of government-sponsored evaluation organizations established and a lot of evaluation activities conducted, suggesting that the important role of evaluation in national development has been highly recognized. The evaluation defined as providing information for decision making (Cronbach, 1963; Stufflebeam, et al., 1971) has been well accepted by various evaluation organizations. Deng Nan, vice-minister of the Ministry of Science and Technology (*People's Daily*, November 1, 1999) said that evaluation system can be of great help to the government and can function in the following four aspects: 1) improving the decision making process; 2) enhancing the macro-level management of technology; 3) promoting innovation in the science and technology management system; 4) and reinforcing the authority of the making and implementation of the national science plan. However, according to Bao, Zhang

and Li (2002), the conduct of an evaluation and the utilization of evaluation results are governed by principles characteristic of the administration, and also affected by the cultural characteristics. It is not easy to give a comprehensive description of the current evaluation practice in China, for the evaluation organizations are independent of each other and operate their duties closely related to their field respectively. But a brief introduction of some of the evaluation organizations and evaluation activities in recent China can be informative to people working in the evaluation field.

## **Evaluation Organizations**

The National Center for Science and Technology Evaluation of China (NCSTEC) was set up by the Ministry of Science and Technology of China in 1994. According to Bao, Zhang and Li (2002), NCSTEC is a specialized agency with responsibility of the evaluation of government-sponsored Science and Technology (S&T) projects. NCSTEC is the leading organization in the field. It plays an important role in providing objective and impartial evaluation to government departments, enterprises and other investment organizations for decision-making related to S&T development. Since the establishment of the Centre, it has conducted many evaluations of major scientific research programs, large high-tech projects, and ventures in high-tech development zones. The key activities conducted by the Centre include:

- evaluations of government-sponsored S&T projects;
- evaluations of S&T policies;
- performance measurement for government-sponsored research institutes; providing services to enterprises and investment companies in the fields of

S&T project evaluations;

- enhancing relationships with international organizations, government departments and other non-government agencies;
- helping to build the capacity of local S&T evaluation agencies.

NCSTE administers an APEC S&T Evaluation Forum Website <http://www.apecevalu.org> for evaluation discussions. The Evaluation Forum aims at promoting the evaluation capacity development, sharing the theories and experience, exchanging information interactively and understanding each other effectively. Evaluation Forum includes main columns as following:  
*What's new*: for the events announcement;

- ✓ *Forums*: for the discussion and sharing the viewpoints, information and knowledge;
- ✓ *Research*: for the collecting and issuing research work both in theory and practice;
- ✓ *Report*: for publication of evaluation reports and so on;
- ✓ *Questionnaire*: for collecting the information and answers to enhance and update the web content continuously.

Higher Education Evaluation Center was established by the Ministry of Education in 2004 (*People's Daily*, October 27, 2004). According to the Ministry of Education, China's economy as well as higher education has developed rapidly in recent years, and higher education must have a professional evaluation system that accords with the economic development. The evaluation center has the following functions.

- The evaluation center will conduct an evaluation of the teaching quality in China's higher education institutions every five years.
- The evaluation center will evaluate the quality of teaching in nearly 2,000 college and universities in the country;
- Rather than ranking universities, the evaluation aims to publicize the teaching process.
- The evaluation center will produce evaluation reports on the evaluated with one of four grades: excellent, good, qualified and unqualified.
- The center will use some traditional appraisal methods, such as measuring graduate employment rate, and will review data on teaching status submitted by each university and overrule any fraudulent data.
- Education departments will join hands with industry associations to evaluate professional education at the universities, adding that the mechanism combining professional evaluation, certification and certificate granting will ensure quality education.
- Apart from improved assessment, the evaluation center also will distinguish itself in terms of capital and evaluation standard and a special fund will be set up to pay for the evaluation process;
- With this large-scale, recurring evaluation practice, China will establish institutionalized evaluation systems to upgrade China's education level.
- Evaluation of key universities will involve foreign experts.

It is learned that this is the first time China has set up a specialized education evaluation center, although the assessment on the teaching of higher educations

was initiated in 1994. By the end of 2003, the Ministry of Education had conducted evaluations on 296 universities, with 16 graded excellent and 192 qualified.

China Information Technology Security Certification Center (CNITSEC; <http://www.itsec.gov.cn>) was originally established in 1997. It is a Chinese Government's authority designed to fulfill national IT security certification responsibilities. In accordance with China laws of product quality certification and IT security management, CNITSEC operates and maintains National Evaluation and Certification Scheme for IT Security. CNITSEC is China's only authorized information technology security certification organization. It is also the only national certification center in China to adopt the international GB/T 18336 idt ISO 15408 standard to test, evaluate and certify information security products, systems and Web services. CNITSEC has the main functions as follows:

- test, evaluation and certification for infosec product and technology;
- evaluation and certification for information systems security;
- evaluation and certification for qualification of IT security service providers;
- evaluation and certification for information security professionals.

The National Center for Safety Evaluation of Drug (NCSED) was set up in China in June, 2002. (*Xinhua News Agency*, June 21, 2002). NCSED is the first drug safety evaluation center that opens in China. The purpose of the Center is to ensure the safety of medicines and it is intended to meet the requirements of the Good Laboratory Practice for Non-clinical Laboratory Studies. The Center was funded by the Chinese government with equipment and technological assistance provided by the Japanese government and the Japan International Cooperation Agency.

## Evaluation Activities

China to Establish Intellectual Property Rights Evaluation System (*People's Daily*, April 19, 2000). A senior official of China's State Intellectual Property Rights Bureau said that a complete intellectual property rights evaluation system will help speed up the commercialization of intellectual property in the domestic market. The evaluation of intellectual property rights is a product of the market economy as well as an important aspect in the commercialization of intellectual property rights.

Evaluation System to Improve City Environment (*Xinhua News Agency*, October 23, 2003). The Chinese government will institute an evaluation system for the natural and living environments of its cities and towns in the hope to harmonize the economic and social development in a sustainable way. According to Wang Guangtao, Minister for Construction, the new system would be designed to evaluate the conditions of natural and living environments including water and gas supply, sewage and trash treatment, drainage system, city greenbelt, biological diversity, heating system, energy, public transport and cultural relic protection.

China's Land Evaluation Open to Public Scrutiny (*People's Daily*, May 31, 2004). Information on over 200 real estate appraisal institutions and over 21,000 land appraisers can now be found in an online information system, as a move to clear away under-the-table practice in land transactions, according to the Ministry of Land and Resources (MLR). The land evaluation sector has been a major social concern in recent years along with the country's economic boom and scandals of illegal land transactions in the burgeoning real estate industry have frequented the media reports. "The system marks that China's land evaluation has begun to be conducted fully in the sunlight," said the MLR in a statement.

China GLP Standard Safety Evaluation Center to be Set Up (*People's Daily*, August 6, 2004). GLP means “good laboratory practice”. It is special management regulations formulated, specially aiming at the medical safety. Currently, no single GLP laboratory in the country reaches the international standards completely. Under this circumstance, no international “pass” is available when China exports its new medicine. Therefore, it is far from mutual recognition among GLP organizations from China and the other countries. After China’s entry into WTO, it is a dispensable foundation for the medicine industry to have a high-level evaluation organization meeting the international standards.

Patent evaluation System designed in Shanghai (*Xinhua News Agency*, May 4, 2004.). A system that can evaluate a patent and give a fair price to it has been designed successfully in Shanghai, and approved by experts. The system, designed by the Shanghai intellectual property right service center and the Shanghai Lixin asset evaluation company, can store numerous data and patent cases and is equipped with special software for evaluation of patents.

## References

- Bao, Y., Zhang, J., & Li, X. (2002). Evaluating government-sponsored science and technology projects in China. *Evaluation Journal of Australasia*, 2(1), 16-19.
- Cronbach, L. J. (1963). Course improvement through evaluation. *Teacher College Record*, 64, 672-683.
- Drummond, R. J. (2003). *Appraisal procedures for counselors and helping professionals*. Upper Saddle River, New Jersey: PEARSON.

Stufflebeam, D. L., Foley, W. J., Gephart, W. J., Guba, E. G., Hammond, R. L., Merriman, H. O., & Provus, M. M. (1971). *Educational evaluation and decision-making*. Itasca, IL: Peacock.

## **Evaluation in Germany: An Overview**

Gerlinde Struhkamp

### **Historical Perspective**

Even though often considered being a “late-starter” concerning evaluation, the beginnings for systematic inquiry into impacts of governmental programs in Germany parallel the developments in the U.S. where evaluation became particularly prominent during the 1960s and from then on. As a modern democracy, too, German government and public administration were concerned with the effects of their actions, be it in the form of laws or programs or other measures of public intervention. So, for example, in 1970 the federal law was passed that there had to be “success controls” (“Erfolgskontrollen”) for governmental measures. Indeed, this law caused a leap in the market for such “success control” studies, though for the most part not academics but commercial research and consultancy firms succeeded to produce the lion’s share of the evaluation research funding (Wollmann, 1997). The relative absence of the academic world regarding evaluations sustained for another 20 years. That does not mean, however, that research and research findings did not pick up matters of broad public interest. For example, large-scale studies were undertaken to explore the effects of different forms of schooling starting at the end of the 1960s and coming to an end early 1980s (Stockmann, 2004b: 29). However, seldom such studies were called evaluation or program evaluation or evaluation research. The peculiar term, literally translated, meaning “accompanying research”

(“wissenschaftliche Begleitung”, “Begleitforschung”) emerged instead and is widely used till today. To clarify the concept: this term is not meant as action research or any form of incorporating advocative elements into the research task (at least not per se). Using evaluation jargon: such “accompanying research” highlights conceptual use and knowledge gain over instrumental purposes. On the other hand, this idea of research does involve in the field, does connect to practice. So there is the notion of feeding research results back into the ongoing process and by that possibly improving the object of analysis—much like in a formative evaluation. In fact, both terms, “accompanying research” and evaluation, are used in Germany till today, the latter becoming more prominent, though. Sometimes they are used in a way exposing visible differences (e.g., an evaluation generating explicit value judgments), sometimes they are used interchangeably. Then, it can only be fathomed in which ways such an approach to applied research overlaps with evaluation and to what extent. A thorough debate on this is still pending.

So, even though the legal and executive interests in evaluation activities did exist in Germany just as in the U.S., an evaluation *profession* did not evolve until the mid 1990s, at least not under the term ‘evaluation’, unlike in the U.S. That does not mean government did not involve any measures of control or accountability. For example, the institution of the “Bundesrechnungshof” (self-portrayal: “The Bundesrechnungshof is a supreme federal authority [...] an independent body of government auditing”<sup>1</sup>), does conduct regular extensive checks on government spending, or ministries, e.g., the Federal Ministry for Economic Cooperation and Development, did set up a central evaluation department (Stockmann, 2004b, p. 31). Finally, a last sweep towards a broader institutionalization of evaluation came

---

<sup>1</sup> <http://www.bundesrechnungshof.de/en/1024.html>

with the introduction of New Public Management also in Germany (Stockmann, 2004b, p. 32).

Therefore, looking at the history of evaluation in Germany, by and large and despite certain ups and downs there has been continuity concerning such tasks as evaluation studies. In contrast to this, though, has been the lack of professionalization towards an evaluation discipline, not only as a field of application (cf. above Wollmann's analysis of consultancies conducting "success control" studies), but also as an academic discipline.

What makes it look like that there had been relatively little interest in evaluation is partly due to the difference in language. The term 'evaluation' (spelled the same in German and pronounced only slightly differently) has become used only, roughly speaking, throughout the last decade. Before that, neither in the academic nor the political world this term really caught on. In fact, you can count almost on one hand the books that carried the term 'evaluation' (or the German adaptation 'Evaluierung', meaning the same) in the title. If so, it occurred mainly in conjunction with close ties to U.S. developments and U.S. authors. For example, as early as in 1972 Wulf published a reader mainly presenting translations of articles by U.S. authors, among them, e.g., Scriven's "Methodology of Evaluation". In a similar vain Hellstern and Wollmann issued an extensive publication in 1984 with chapters by German, but also again a number of U.S. authors. Weiss's "Evaluation Research" from 1972 was published in German language in 1974, and in 1988 Hofmann translated and adapted Rossi/Freeman's "Evaluation: A systematic approach".

Which looks like an impressive list here, was not exciting from the German perspective: literally a handful of books specifically related to evaluation in more than 30 years!

Ties to the U.S. evaluation community have generally been strong. Also Koch/Wittmann (1990) and most recently Stockmann (2000, 2004a) incorporated chapters on U.S. developments and by U.S. authors in their handbooks, Beywl's work (1988) drew on U.S. methodological developments in evaluation. Maybe Wottawa and Thierau's book from 1990 (in the meantime second edition 1998 and third 2003) could be considered a turning point towards a more German-centered evaluation-related body of literature. For example, next to introductions of evaluation concepts stemming largely from U.S. writings, they tried to focus on specific German developments, including, e.g., related concepts of quality assurance like quality management systems. Though, it may be difficult to pinpoint to a certain book since a general development has taken place.

## **Current Status**

Starting from the early 1990s and certainly since the mid and late 1990s the relative scarcity of writings about evaluation in German language simply belongs to the past. Book publications and articles have popped up from all kinds of disciplines, within Germany and the German-speaking countries Austria and Switzerland (German-speaking part), respectively. Fortunately, there is not any “border mentality”—in the contrary: many intellectual and personal exchanges occur between German-language academics, journals and at conferences facilitated by the common language. Noteworthy in this regard is, e.g., that so far evaluators from Austria simply join the German Evaluation Society ([DeGEval](#)), while the current DeGEval president is Austrian, and the last annual conference in

November 2004 took place in Vienna. Currently there are considerations underway towards changing the society's name reflecting the German and Austrian membership (but that decision has not been made at the time of this article). Moreover, within the coming years a joint conference of the DeGEval with the Swiss Evaluation Society (SEVAL) is planned. Also, the DeGEval entertains ties with the European Evaluation Society (EES).

Probably the most significant turning point towards the establishment and professionalization of evaluation in Germany meant the foundation of the DeGEval in 1997. Since then, the society has continued to grow and spurred the intensity of intellectual discourse on evaluation-related topics. To date the DeGEval has about 370 individual and more than 50 institutional members, these numbers increasing steadily since its foundation. Unlike, e.g., the American Evaluation Association (AEA) or the UK Evaluation Society (UKES), the DeGEval does not have any regional chapters or regional networks.

The DeGEval's internal structure compares to the AEA's formation in TIGs (Topical Interest Groups). The number of TIGs (in German called "Arbeitskreis"—"working circle" or "working group") within the DeGEval has also increased steadily and amounts to 14 as to date. Like in the AEA, the TIGs are mostly centered on a certain field of application, e.g., evaluation in schools, evaluation of developmental aid, environmental evaluation, evaluation in the field of human services, and so on. TIGs are created and may break up again once a task is done and depending on the actuality of a certain topic. For example, in the beginning of the DeGEval a working group was formed to develop German "Standards for Evaluation". Once accomplished, this group was terminated. As there is a first revision process of the standards in the making, again such a task force has been formed. A few TIGs deal with aspects of broader interest that

mainly concern matters of an evaluation profession, e.g., one TIG prepared the DeGEval's "Recommendations for education and training in evaluation" to specify evaluator competencies necessary for sound evaluation practice.

Currently, the 14 TIGs are dealing with:

- 1) Training and education in evaluation
- 2) Vocational education
- 3) Developmental aid
- 4) Research, technology and innovation
- 5) Health sector
- 6) Higher education
- 7) Media
- 8) Schools
- 9) Human services
- 10) Urban and regional development
- 11) Structural funds
- 12) Environment
- 13) Public administration
- 14) Corporate sector (in preparation)

Comparable to the AEA, each DeGEval TIG has a chair and vice-chair. The TIGs are largely autonomous in their activities. Some TIGs exist since the foundation of the DeGEval and not only sponsor meetings during the annual conference but also organize meetings (like small conferences or workshops regarding a certain topic)

throughout the year. At the two to three days long annual conferences the TIGs sponsor sessions. So far, the chair and vice-chair or TIG members look for appropriate and interesting presenters and invite them. Also, presenters can directly address the TIG and offer a topic they would like to present on. Since there is not a general call-for-papers (only by certain TIGs, if they opt for one), on the one hand, it opens the opportunity for people to be invited who normally would not answer to a call-for-papers by themselves (e.g., when they work at government agencies, foundations, corporations or other institutions with only loose ties to the academic world); on the other hand, it somewhat limits the range of presenters to the perspective of the TIG. The annual conferences sponsored by the DeGEval take place since its establishment in 1997, prior to the conference professional development workshops are offered. The conferences are held in fall (mostly October; this year it will take place at the University of Duisburg-Essen, located in the Ruhr valley, from Oct. 12-14) and in German language. However, visitors with, let's say, a working level of German language proficiency should feel welcome to attend, since there would be also ways to communicate in other languages, foremost English. Even though rarely, but there have been already occasional presentations in English.

Despite the DeGEval's attempts to encompass various disciplines (like the AEA is devoted to “evaluation in all its forms”) and the growing interest in the DeGEval's annual conferences, there are still “parallel universes” where evaluation is dealt with, mainly dominated by the traditional disciplines like education, psychology, and sociology. Their professional associations do pick up evaluation topics, e.g., in the form of own TIGs or working groups, but linkages to the DeGEval are still rather weak and sporadic. Mainly such connections exist in the way that people attend the conferences and engage in both the DeGEval and another association—

like U.S. evaluators may attend both the AERA (American Educational Research Association) as well as the AEA meeting, e.g.. The next years will show how much overlap and integration will be possible to overcome the “disciplinary segmentation” (Stockmann, 2004b, p. 35) in order to develop a common understanding on core elements of evaluation and synthesize scientific debates.

Apart from the foundation of the DeGEval, other developments foster the professionalization of evaluation in Germany. One of the society's founding members, Wolfgang Beywl, set up and administers the German-language mailing list called “forum-evaluation”, which engages several hundred enlisted members in discussions about concepts and ideas, exchanges of references, announcements of events, calls-for-papers, and the like. Similarly beneficial to the field is the first German-language “Journal for Evaluation” (*Zeitschrift für Evaluation—ZfEv*<sup>2</sup>). In its third volume (2005) it incorporates articles on theory, methods and practice of evaluation, book reviews, updates on activities of the DeGEval, and other pertinent information of interest to evaluators, sponsors, and anybody else concerned. International readers: the journal does include English-language abstracts!

Not only have evaluators and others interested in evaluation-related issues found their forums. In addition and appreciably so, as of now there are two German-language postgraduate, one of them master-level, degree programs for evaluation set in place: one in Berne/Switzerland<sup>3</sup>, up and running since 2001, the second one in Saarbrücken<sup>4</sup>, which had its first cohort fall 2004. In addition, during the last

---

<sup>2</sup> <http://www.zfev.de/>

<sup>3</sup> <http://www.evaluationsstudium.ch/>

<sup>4</sup> <http://www.htw-saarland.de/evaluation/>

years, professor positions within departments of social sciences and education have been set in place with an emphasis on evaluation, so more and more students will be trained more formally in techniques, methods and context factors concerning evaluation.

The first major product of the DeGEval has been to adopt and as professional association responsible to pass the “Standards for Evaluation”, also called the DeGEval-Standards (Deutsche Gesellschaft für Evaluation, 2002). The proximity to the wording used in the “Joint Committee Standards for Educational Evaluation” is no surprise, since the former are closely related to the latter. Like the Swiss Evaluation Society SEVAL had issued its respective evaluation standards in 2001 after a review and revision process, so did the DeGEval finalize its review process in 2001/2002 and prepared a brochure listing and explaining the standards to their members and others being interested. A new review process of the existing standards, based on a survey among DeGEval's members, is currently underway. By and large there have been only slight differences between the Joint Committee, SEVAL and DeGEval Standards, so the latter are based to a large extent on the work and experience of the Joint Committee.

In addition to the Evaluation Standards, in 2004 the DeGEval's TIG “Education and Training in Evaluation” also issued “Recommendations for education and training in evaluation—Required competencies for evaluators” (Deutsche Gesellschaft für Evaluation, 2004). These influenced already the existing academic training programs mentioned above. The latest major recommendation passed by the DeGEval is the adoption of the Evaluation Standards to the special form of self-evaluation (Deutsche Gesellschaft für Evaluation, 2005).

## **Specifics in Germany**

In one way, for German evaluators the situation has probably been much like in the U.S. in the 1960s when the profession just started and evolved—had to find its way. But since the English language is only a relatively slight barrier to many, those interested in evaluation indulge in the English-language literature on evaluation and pick up the ideas. As a consequence, an interesting mixture of concepts rooted in American culture and approaches stemming from German traditions of social science, policy-analysis (and every other respective field) merge and emerge.

For example, in the fields of social work and human services, e.g., concerning child and youth services, the approach of “self-evaluation” has become prominent. In fact, it proved to be a “gate-opener” during the mid and late 1990s in introducing evaluation to the field, not only for approaches of self-evaluation but also other “traditional” forms of external evaluation. The DeGEval responded to this in adjusting the DeGEval Evaluation Standards to applications of self-evaluation. A respective paper explaining the specifics of self-evaluation and how the Evaluation Standards respond to them was adopted by the DeGEval members at the general assembly during last year's annual conference.

More as a side note: An interesting discussion, indebted to the terminological differentiation that's made possible by the German language, sparks from time to time, e.g., in the mailing list “forum-evaluation”. There is a dispute regarding the differences between various forms of evaluation that could be distinguished by the attributes of “internal” (German: “intern”) and “external” (“extern”) as well as “self-” (“selbst”) and—well, here there's the German term “fremd”, meaning literally “strange” or more metaphorically “outside”, which is hard to translate into

English. Thomas Widmer, Swiss evaluation researcher, suggested to translate it as “heteronomous” (evaluation), thus, an evaluation in which the evaluatees are not in charge of the evaluation, i.e., have a say in the conduct of the evaluation. In contrast, in a self-evaluation they are in charge of both the evaluand and the evaluation (so a prominent, yet not undisputed definition). Also other attempts have been made to provide a German-language glossary of evaluation terminology (even referring to the English corresponding term, if applicable)<sup>5</sup>, in its make-up very similar to The Evaluation Center's glossary project<sup>6</sup>.

Another topic that has been prevalent and mixed with evaluation debates in some sectors concerns approaches of “quality management” (“Qualitätsmanagement”), e.g., according to the approaches of Total Quality Management (TQM), the European Foundation for Quality Management's model EFQM, or the International Standard Organization's (ISO) norms (being transferred into German language and context by the German Institute for Norming—“Deutsches Institut für Normung”, DIN). A debate that has largely been absent from the evaluation community in the U.S., as far as my observation goes. In Germany, however, in some sectors there is a prevalence of quality management terminology whereas in others one of evaluation. And since terminology carries concepts, it has not been easy to pull the two strands apart. Several authors have worked out differences and similarities between these two approaches (Wottawa/Thierau, 1998, pp. 43-45; Beywl, 2001; Stockmann, 2002), but as of now it more looks like another “parallel universe”, with a conceptual conciliation still to be worked out.

---

<sup>5</sup> <http://www.univation.org/glossar/index.php>

<sup>6</sup> <http://ec.wmich.edu/Glossary/glossaryList.htm>

These are but two examples from my work context. Others surely could add more, e.g., concerning the fields of developmental aid, European structural funds, evaluation of sustainable development and so on—a list too long to be presented here.

Such disputes over various forms of evaluation and the assisting terminology are not yet settled. It's been taken serious what Michael Scriven suggested in his editorial in JMDE Num. 1: that “one must treat the definition of key existing concepts as an extremely serious matter, not a matter of casual linguistic convenience [...]. Conceptual schemes, and the definitions that go with them, are powerful instruments of analysis and hence persuasive support for particular interpretations, not minor precursors to it [...].” (2004, pp. 15-16). Indeed, there is this seriousness of—constructive—debate in the German-language evaluation community.

## **About the Author**

After an academic degree in Germany the author received her M.A. in ‘Evaluation Studies’ from the University of Minnesota, USA, in 2002. Since then she works in Germany again, currently concerned with evaluation in the field of children and youth programs. Her general interests are evaluation of human services and the theoretical foundations of evaluation. Correspondence to: Gerlinde Struhkamp, German Youth Institute, Nockherstr. 2, D-81541 Munich, Germany. Tel.: 0049-89-62306-340, E-mail: [struhkamp@dji.de](mailto:struhkamp@dji.de) or [Gerlinde.Struhkamp@gmx.de](mailto:Gerlinde.Struhkamp@gmx.de). The author wishes to thank both Sandra Speer and Karin Haubrich for their thoughtful comments on an earlier version.

## References

- Beywl, W. (1988): Zur Weiterentwicklung der Evaluationsmethodologie. Grundlegung, Konzeption und Anwendung eines Modells der responsiven Evaluation. [Development in evaluation methodology. Basis, conception and application of the model of responsive evaluation] Frankfurt a.M./Bern/New York u.a.: Peter Lang.
- Beywl, W. (2001): Evaluation und Qualitätsmanagement. Systemische Verfahren zur Entwicklung von Qualität im Bildungswesen. [Evaluation and quality management. Systemic procedures to develop quality in education] In: Bundesministerium für Bildung, Wissenschaft und Kultur/Bundesinstitut für Erwachsenenbildung St. Wolfgang (Hg.): *Konzepte der Qualität in der Erwachsenenbildung* [Concepts of quality in adult education], Materialien zur Erwachsenenbildung Nr. 2/2001, Aufsätze und Protokoll im Rahmen der Werkstatt am Bundesinstitut für Erwachsenenbildung St. Wolfgang vom 2. bis 3. Oktober 2000. Verfügbar unter [http://wwwapp.bmbwk.gv.at/medien/6048\\_PDFzuPubID88.pdf](http://wwwapp.bmbwk.gv.at/medien/6048_PDFzuPubID88.pdf) [20.04.2005], pp. 7-17.
- Deutsche Gesellschaft für Evaluation (DeGEval) (2002): *Standards für Evaluation*. [Standards for evaluation] Köln: Deutsche Gesellschaft für Evaluation.
- Deutsche Gesellschaft für Evaluation (DeGEval) (2004): *Empfehlungen für die Aus- und Weiterbildung in der Evaluation. Anforderungsprofile an Evaluatorinnen und Evaluatoren*. [Recommendations for education and training in evaluation. Required competencies for evaluators] Alfter: Deutsche Gesellschaft für Evaluation.

Deutsche Gesellschaft für Evaluation (DeGEval) (2005): *Empfehlungen zur Anwendung der Evaluationsstandards der DeGEval im Handlungsfeld der Selbstevaluation*. [Recommendations for application of the DeGEval evaluation standards to self-evaluation] Verfügbar unter <http://www.degeval.de/calimero/tools/proxy.php?id=139> [20.04.2005].

Hellstern, G.-M./Wollmann, H. (Hg.) (1984): *Handbuch zur Evaluierungsforschung Bd. 1*. [Handbook of evaluation research] Opladen: Westdeutscher Verlag.

Koch, U./ Wittmann, W. W. (Hg.) (1990): *Evaluationsforschung. Bewertungsgrundlage von Sozial- und Gesundheitsprogrammen*. [Evaluation research. Basis for assessment of social and health programs] Berlin/Heidelberg/New York: Springer.

Rossi, P. H./Freeman, H. E./Hofmann, G. (1988): *Programm-Evaluation. Einführung in die Methoden angewandter Sozialforschung*. [Program evaluation. Introduction to the methods of applied social science research] Stuttgart: Enke.

Scriven, M. (2004): Editorial: The Fiefdom Problem. In: *Journal of Multidisciplinary Evaluation*, No. 1 (Oct. 2004), pp. 11-18.

Stockmann, R. (Hg.) (2000): *Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder*. [Evaluation research. Foundations and selected fields] Opladen: Leske + Budrich.

Stockmann, R. (2002): Qualitätsmanagement und Evaluation – Konkurrierende oder sich ergänzende Konzepte?. [Quality management and evaluation –

- competing or complementary concepts?] In: *Zeitschrift für Evaluation*, 2/2002, pp. 209-243.
- Stockmann, R. (Hg.) (2004a): *Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder*. 2. Auflage. [2nd edition] Opladen: Leske + Budrich.
- Stockmann, R. (Hg.) (2004b): Evaluation in Deutschland. In: *Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder*. 2. Auflage. [2nd edition] Opladen: Leske + Budrich, pp. 13-43.
- Weiss, C. H. (1974): *Evaluierungsforschung. Methoden zur Einschätzung von sozialen Reformprogrammen*. [Evaluation research. Methods to assess social reform programs] Opladen: Westdeutscher Verlag.
- Wollmann, H. (1997): Evaluation in Germany. In: *European Evaluation Society, Newsletter (3)*, pp. 4-5.
- Wottawa, H./Thierau, H. (1990): *Lehrbuch Evaluation*. [Textbook evaluation] Bern/Göttingen/Toronto: Huber.
- Wottawa, H./Thierau, H. (1998): *Lehrbuch Evaluation*. 2. vollst. überarb. Auflage. [2nd fully revised edition] Bern/Göttingen/Toronto: Huber.
- Wottawa, H./Thierau, H. (2003): *Lehrbuch Evaluation*. 3. korr. Auflage. [3rd corrected edition] Bern/Göttingen/Toronto: Huber.
- Wulf, C. (Hg.) (1972): *Evaluation. Beschreibung und Bewertung von Unterricht, Curricula und Schulversuchen*. [Evaluation. Description and Assessment of instruction, curricula and schooling trials] München: Piper.

## **Evaluation—Making it Real in Aotearoa New Zealand: Leading by Example, Leading by Association**

Pam Oliver, Kate McKegg, Geoff Stone, and Maggie Jakob-Hoff

The second Aotearoa New Zealand Evaluation Conference, sponsored by the Auckland Evaluation Group, will be held 18-20 July, 2005 at the Tauhara Centre, Acacia Bay, Taupo.

This year's conference follows on from the very successful 2004 Auckland Evaluation Group Conference, at which evaluators and others from various parts of the country came together and shared their challenges around evaluation practice in New Zealand. In many respects, the theme for this conference builds on that work.

The theme "Evaluation—Making it Real in Aotearoa New Zealand" is about evaluator roles and what we actually do as evaluators. We will explore what is unique about those roles in the New Zealand context, particularly how we work in partnership with Maori and how we work with other cultures like Pacific peoples, Asians and refugees.

"Leading by example" means that as practitioners we are prepared to subject our practice to reflection, and to the scrutiny of others. It is about openly striving to understand, respond, learn and evolve in our work.

"Leading by Association" means that we take active steps together to grow professionally, and to develop our profession. We organise to gather about us

critical friends and supportive colleagues, to create a space for sharing skills, mistakes, insights, motivations and possibilities, and we create structures and systems to promote safe professional practice in evaluation.

Key aspects of this theme are evaluation as a profession, accountability, rigour, consciousness and relevance to New Zealand.

Charmaine Pountney and Dr. Te Kani Kingi will be the keynote speakers. Dr. Te Kani Kingi's talk will be entitled "Evaluation and the measurement of cultural outcomes." He will examine the process and practice of evaluation as well as the requirement to measure activities and outcomes that are culturally derived. Charmaine Pountney's address will be entitled "Doing evaluation: From magic marks to vital values." She will provide provocations and challenges on two key themes of the conference—what are the essential features of evaluation work across a range of settings? and what are the necessary attributes of a professional association which will promote effective and ethical evaluation while avoiding the risks of becoming a professional clique?

## **Further Information**

If you have any queries, please feel very welcome to contact any of the organizing committee:

Pam Oliver—09 372-7749 / [pamo@clear.net.nz](mailto:pamo@clear.net.nz)

Kate McKegg—07 870-1665 [kate.mckegg@xtra.co.nz](mailto:kate.mckegg@xtra.co.nz)

Geoff Stone—04 460-3052 [geoff.stone@corrections.govt.nz](mailto:geoff.stone@corrections.govt.nz)

Maggie Jakob-Hoff—09 360-0827 [maggie.jh@evaluate.co.nz](mailto:maggie.jh@evaluate.co.nz)

# **A Review of the Chinese National Center for Science and Technology Evaluation**

Laura Pan Luo

China is now in a transitioning stage from planned economy to market economy. There is a growing interest in China to have a strong evaluation process in place so that planning and decisions can be based on valid and credible information. Evaluation also provides a guide for resource allocation.

The Chinese Ministry of Science and Technology (MOST) is the highest administrative body responsible for formulating and implementing science and technology (S&T) policies and programs in China. To provide accountability for government funding and improve management practices, in 1997, MOST commissioned an independent entity, the National Center for Science and Technology Evaluation (NCSTE), to conduct evaluations of science and technology policies and programs in China. The goal of NCSTE is to provide an objective peer review of government-funded S&T research programs.

Ms. Deng Nan, former Chinese Vice Minister of the Ministry of Science and Technology, noted that the evaluation system is important in the following four aspects: (1) improving the decision-making process; (2) enhancing the macro-level management of technology; (3) promoting innovation in the science and technology management system; and (4) reinforcing the implementation of the national science plan (People's Daily, 1999).

Over the last several years, NCSTE has evaluated over 1,000 projects focused on technical, institutional, economic, and financial aspects, ranging from information technology, to health care, environment protection and sustainable development. As the leading evaluation organization, NCSTE aims at providing timely and accurate information for both government agencies and private organizations to assist in their decision-making process. Additionally, it strives to promote dialogues among central and local governments, private sector, and academia.

Mandated by MOST, NCSTE issued China's Science and Evaluation Standards (Standards) in 2001. MOST made the Standards an annex to the government regulation on evaluation management. Since its issuance, the Standards have been well observed in science and technology evaluation in China. It has been selected as the training material on science and technology evaluation. More than 600 people across China have participated in the training workshop on Standards.

NCSTE consists of employees who specialize in areas such as management, system engineering, public policy research and economics. NCSTE also hires consultants to work on various evaluation projects. For example, in 1997, NCSTE conducted evaluations on a number of National Engineering Technology Centers to assess the effectiveness of their management and operations. NSCTE also conducted a policy review and analysis of China's new and high tech industrial development zones, technology transfers in the Sino-Japan computer industry and the role and impact of foreign investment on the development of new technology oriented industries in China.

In recent years, NCSTE has also conducted evaluations of foreign aid. The aid evaluation project teams at NCSTE have studied the relevant OECD development aid policies and the policies on utilization of foreign government loans to China.

As a local partner, NCSTE has conducted joint evaluations with international institutions. For example, NCSTE has evaluated Norwegian Mixed Credits jointly with a Norwegian consulting agency, Institute of Applied Social Science of Norway (FAFO).

NCSTE has had collaborations with many countries in the world, including the US, France, Canada, Japan, New Zealand, UK, Netherlands, Thailand, Korea and India in addition to providing services to the World Bank, the United Nations Development Program and other NGOs. According to Chinese officials, NCSTE has improved management practices at MOST research programs, and the evaluation of science and technology has contributed remarkably to the development of Chinese society by making the policy and decision-making process more objective.

## **References**

Chelimsky, Eleanor & Shadish, Williams R. (1997), Editors. *Evaluation for the 21<sup>st</sup> Century*. Thousand Oakes, CA: Sage.

Chen, Zhaoying (time unknown). *Making S&T evaluation the tools for government decision-making practice in China*

National Center for Science and Technology Evaluation (2001). *The Uniform Standards for Science and Technology Evaluation*. Beijing, China: China Price Publisher.

People's Daily (November 1, 1999). *China to strengthen evaluation of scientific projects*.

Scriven, Michael (1991). *Evaluation Thesaurus*. Thousand Oaks, CA: Sage.

# Evaluation in Japan

Ryo Sasaki

## Overview

Two years have passed since the “Government Policy Evaluations Act” (GPEA) became effective in Japan on April 1, 2002. In that time, evaluation has been well accepted as an essential part of the policy management cycle at each ministry of the Japanese government. It is reported that evaluation results have been utilized for budget formulation by governmental ministries, and it is also observed that policies have been prioritized and, conversely, abolished based on the evaluation results. Other merits of introducing evaluation are that the so-called ‘policy diagram’ has been frequently developed at ministries, and policy goals have become more outcome-oriented with more quantitative measures. Now the Act is under discussion for amendment with some major points proposed for change.

## Background to Introduction of the Act

‘The Basic Law for the Reorganization of Central Government Ministries and Agencies’ came into effect in June 1998, marking the start of serious reform of the Japanese public sector. Though the law comprises almost all the subjects of administrative reform, strengthening of *policy evaluation* is pointed out as one of the major tools for government-wide reform. One thing should be pointed out: the word *policy* is used with a very broad meaning in Japan, and this includes all three levels in the hierarchy of governmental activities, namely, *policy, program and projects*.

Reflecting the concept of this basic law, the Ministry of Internal Affairs and Communication (hereafter the MIC)<sup>7</sup>, prepared the ‘Standard Guidelines for Policy Evaluation’ in 2001, and the MIC encouraged each ministry to test them on their

---

<sup>7</sup> Ministry of Public Management, Home Affairs, Posts and Telecommunication was the formal name in 2002.

policies, programs and projects. Three approaches were suggested in the guidelines, namely ‘project evaluation,’ ‘performance evaluation’ and ‘comprehensive evaluation.’ These names do not match with the internationally accepted academic norms. Roughly saying, ‘project evaluation’ is a different expression for *ex-ante project-level evaluation*, or simply *appraisal*. ‘Performance evaluation’ is equivalent to *performance measurement*. ‘Comprehensive evaluation’ is almost the same as *program evaluation* as has been developed by evaluators for the long term. (see Box 1)

After a certain period of examination, the GPEA was prepared by the MIC and passed through the Diet in 2001. The law required all governmental ministries to evaluate their policies and report the results to the public. It also asked ministries to reflect evaluation results in policy and budget formulation, albeit not by mandate.

### **Box 1. Summary of Standard Guidelines for Policy Evaluation**

#### **Evaluation Method and Performance Ideas**

Based on the following three standard evaluation methods, each government office must select an appropriate evaluation method and carry out evaluation in accordance with the characteristics of its own policy and the need for policy evaluation in each area.

- (1)“**Project Evaluation**” to provide information useful for adoption, rejection, and selection of administrative activities by conducting evaluation beforehand, and carrying out verification during and after the implementation.
- (2)“**Performance Evaluation**” to provide information on the extent of policy achievements. This is accomplished by setting up the goals to be achieved beforehand in the wide-ranging areas of administration, measuring the performance, and evaluating the extent of goal achievements.
- (3)“**Comprehensive Evaluation**” to provide a variety of information useful for solving problems by setting up a specific theme, carrying out comprehensive evaluation by looking at the theme deeply and from various angles, and finding out policy effects

Source: Ministry of Internal Affairs and Communication, *Summary of Standard Guidelines For Policy Evaluation*, 15 January 2001

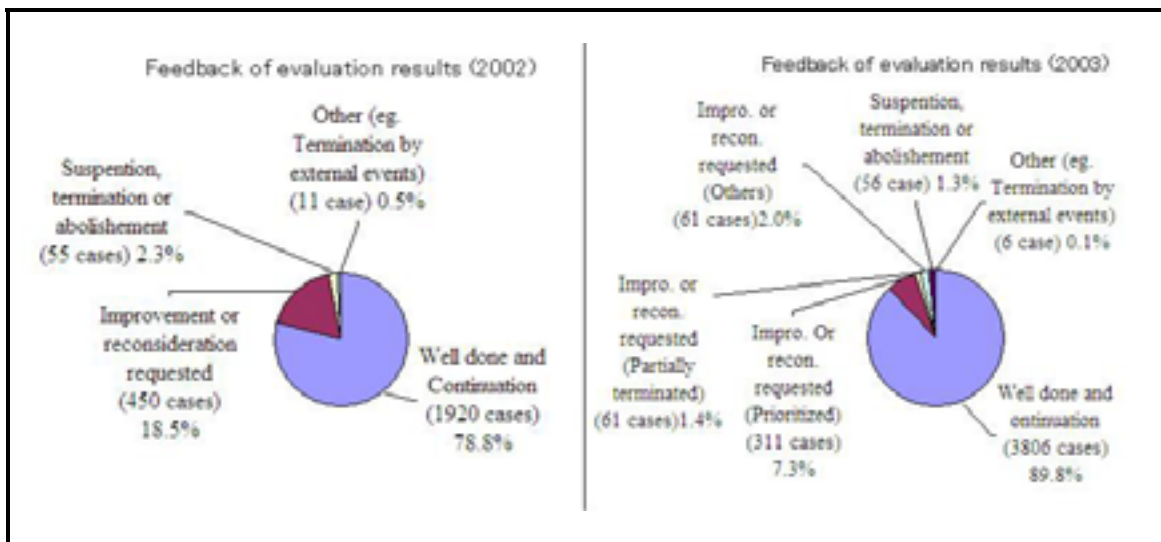
## **Utilization of Evaluation Results**

As has been already mentioned, two years have passed since the Act came into effect. The MIC has conducted survey for each ministry and published a report concerning the extent and degree to which evaluation results are utilized.

## Evaluation Results are Well Utilized for Policy and Budget Formulation

In 2002, a total of 2,436 ex-post evaluations were conducted using one of the approaches suggested above. Out of the total, 1,920 cases (78.8%) were evaluated as ‘well done and should be continued as is;’ 450 cases (18.5%) were evaluated as ‘should be improved or reconsidered;’ and 55 cases (2.3%) out were judged as ‘should be suspended, terminated or abolished,’ which has actually transpired. In 2003, a total of 5,923 ex-post evaluations were conducted and with the breakdown of results as shown in the following figure.

**Figure 1. Feedback of Evaluation Results (2002, 2003)**



Source: Ministry of Internal Affairs and Communications, *Implementation Situation of Policy Evaluation and Feedback for Policy Formulation*, 2004 and 2005

## Policy Diagrams Have Been Developed at Each Ministry

Along with the introduction of evaluation activities, the so-called policy diagram was developed in more than half of all ministries. A policy diagram is like a hierarchy of policies, programs and projects, or a hierarchy of mission, vision, strategic goals, programs, and associated activities. For instance, the Ministry of Agriculture, Forestry and Fisheries has developed a policy diagram consisting of 5 major goals, 12 intermediate goals, 59 policy areas with 142 numerical targets, and associated programs and interventions. This kind of framework was not considered in Japan until the introduction of evaluation activities. It has been unanimously

reported that policy diagrams are shared throughout whole organizations and are used as effective internal communication tools.

## **Policy Goals Have Become More Outcome-Oriented with More Quantitative Measures**

It is reported, for example, by the Ministry of Education and Technology, that certain words such as *outcomes* and *performance indicators* were broadly accepted and their concepts were shared by the entire organization. The MIC reported that the ratio of cases where performance targets are set in a quantitative manner has increased from about 30% in 2002 to more than 50% in 2003.

## **Discussion for Amendment of the GPEA**

The GPEA states that the Act shall be amended based on the lessons learned after three years. The professional committee of the MIC, the formal name of which is the *Committee for Policy Evaluation and Independent Administrative Institutions*, published a report entitled 'Major points for amendment of policy evaluation system' in December 2004. A summary of the report is given below (see Box 2). Based on these points, discussion will heat up through this year, and it is expected that amendment of the GPEA will be actually proposed to the Diet at the end of fiscal 2005. Professional associations, such as Japan Evaluation Society, are strongly requested to contribute to this discussion and take a lead on the appropriate use of professional terms on evaluation and the diffusion of various evaluation concepts.

## **Box 2. Major Points for Amendment of Policy Evaluation System**

### <Feedback of evaluation results for policy formulation>

- More feedback for budget formulation as well as policy formulation should be done
- 'Units' to which evaluation is applied should be set more clearly. For example, 'units' can be recognized by development of an appropriate policy diagram.
- Mindset of staffs should be changed. Concepts of management cycle and results-oriented management should be diffused.

### <Promotion of more objective and rigorous evaluation>

- Target setting should become more quantitative.
- Information of *cost* invested for policy implementation should be gathered and cost-effectiveness analysis should be conducted more frequently.
- Knowledge of academic and professional expertise should be utilized more.
- Possibility of re-examination and double check by outside expertise should be maintained.

### <Sophistication of evaluation activities>

- Various evaluation activities should be appropriately prioritized and conducted in a more cost-effective way.
- Ex-ante evaluation on introduction and amendment of public regulation should be more sophisticated and amplified.

### <Public report of evaluation results>

- Evaluation report should be prepared in a more reader-friendly manner.
- National discussion on evaluation should be stimulated.

### <Other issues>

- Cooperation with regional authorities, - Role of the MIC, - Tie-up with related fields

Source: Committee for Policy Evaluation and Independent Administrative Institutions, *Major points for amendment of policy evaluation system*, December 2004

## **References**

Ministry of Internal Affairs and Communication, *Summary of Standard Guidelines for Policy Evaluation*, 2001

[http://www.soumu.go.jp/english/kansatu/evaluation/evaluation\\_04.html](http://www.soumu.go.jp/english/kansatu/evaluation/evaluation_04.html)

Ministry of Internal Affairs and Communications, *Implementation Situation of Policy Evaluation and Feedback for Policy Formulation*, 2004 and 2005.

Committee for Policy Evaluation and Independent Administrative Institutions, *Major points for amendment of policy evaluation system*, 2004.