

EVALUATION CONNECTIONS

The EES Newsletter

- ✓ Articles
- ✓ News
- ✓ Events



www.ees2010prague.org

The 9th European Evaluation Society International Conference
6th-8th October 2010, Prague, Czech Republic

number 3

CONTENTS

Editorial	1
A brief history of evaluation	2
The Policy for Monitoring and Evaluation of	
Primary Health Care in Brazil from 2003 to 2006	3
Sixth European conference on Evaluation of Cohesion Policy	4
Member Profile	5
The Limitations of Econometric Methods for Impact Evaluations	6

Editorial

Anne Mc Kee, Editor of EC

Welcome to the fourth edition of 'Connections', the newsletter for the European Evaluation Society (EES). In this edition we have contributions from the United Kingdom, Greece, Brussels and Brazil, reflecting the breadth of our membership and their engagement in evaluation activities.

Robert Picciotto offers an historical over-view of evaluation and the roles it has played and continues to play in society. A health research team from Brazil report briefly upon monitoring and evaluation policy in primary care. Veronica Gaffey from the 'Directorate General for Regional Policy' reports on the Sixth European Conference on Evaluation Cohesion Policy. A new member, Charalabos Bakopoulos

describes briefly why he joined the EES and what he hopes to get from the forthcoming conference in Prague. Scott Bayle, an evaluation specialist, provides quotations about 'Econometric Methods' from a diverse range of methodologists who identify limitations to this methodological approach.

We encourage you to use the newsletter to express your views on evaluation, share issues and findings arising from your research, report on conferences that you have attended or organized and alert us to forthcoming conferences.

In this edition, we introduce a new feature, short interviews with individual members. We hope that these

electronic interviews will help us 'put a human face' on the membership and develop both a sense of audience for the newsletter and community between face-to-face meetings. If you would like to take part in an electronic interview please contact either Karel.Jezek@czech-in.cz or Anne MC Kee Anne.mckee@anglia.ac.uk.

A brief history of evaluation

Robert Picciotto, EES Board member

As a practice, evaluation has an ancient pedigree. In prehistoric times, stone-cutters improved their work methods through experience. The Socratic dialogues of ancient Greece were evaluative. The civil service of ancient Egypt systematically evaluated the rich harvests of the Nile delta. Civil service examinations were a regular feature of ancient China's governance.

As a discipline, the foundations of evaluation can be traced to the scientific method which emerged in the seventeenth century with Descartes' "Discourse on the Method" (1637) and Isaac Newton's "Principia Mathematica" (1687). By rejecting dogma and stressing rationality as the foundation of human knowledge these two philosophers changed the course of intellectual history.

In parallel, William Petty, a physician and a statistician, invented the art of "political arithmetic" by publishing a pamphlet about tax policy in 1662. A century later enlightenment ideas expanded the quest for knowledge to laws and institutions. This is when reason was decisively freed from religious dogma; when critical inquiry overcame superstition and when tradition-bound beliefs gave way to notions of contractual rights, freedom and democracy.

In sum, the scientific method and the systematic use of observation to understand and guide social change – core evaluation principles – can be traced back to the seventeenth and eighteenth centuries. The seminal writings of Condorcet (produced between 1774 and 1794) presaged evaluation. They embodied conceptions of human progress that still resonate today.

Emblematic of the Age of Reason, Condorcet probed the art of government, the mathematics of democracy and the management of social change. His work was imbued

with the belief that universal human values are compatible with freedom of thought and individual liberty. He rejected divine revelation and embraced empirical and rational inquiry as did the writings of other enlightenment philosophers.

Another forerunner of the evaluation discipline consists in the use of government appointed commissions to assess social programs. This practice started in Sweden in the seventeenth century and still persists today. The practice has spread throughout the rest of the world. Its original focus was students' achievement, the performance of schools and the review of literacy and public health programs. Today's independent commissions span all fields and constitute a major instrument of United Nations governance.

The historical intellectual and institutional antecedents described above notwithstanding, evaluation as the free standing discipline that we practice today only flourished in the mid-20th century. It arose out of the ashes of World War II. The post war period was characterized by optimism, "can do" attitudes and trust in government. These were halcyon days for the social sciences – and pioneering days for evaluation.

A new discipline had emerged. By the late 1950's evaluation was embedded in the public policy cycle of education, health, public housing, law enforcement and community based programs. In these and other social areas evaluation acted as a transmission belt between social science departments in universities and policy making units in government.

As research methods and availability of statistical data improved, evaluation was called upon to study social processes and to inform government policy. During the 1960's the market for program evaluation was boosted by

Lyndon Johnson's War on Poverty and the basic intellectual foundations of the discipline were laid.

The development of evaluation as a distinct academic specialty was buttressed by textbooks and research articles in the 1970's. Evaluation journals began to appear and the first evaluation associations were founded. Beyond its social program applications, evaluation branched into environmental protection and energy conservation.

Large social experiments to test new fiscal and housing allowance policies were undertaken. The same period witnessed the use of cost benefit analysis by Robert McNamara's Pentagon whiz kids aimed at delivering "more bang for the buck".

In the following three decades, the discipline has had its ups and down. Under the Reagan administrations of the 1980's social programs were cut and the demand for evaluation declined sharply at the federal level. It revived to a limited extent under the Clinton administration and it remained stationary in the Bush years. But throughout this period its reach has expanded remarkably at state and community levels as well as within the private and voluntary sectors.

By now the centre of gravity of the evaluation community has shifted towards Europe and the developing world. The demand for evaluation was boosted considerably by the 1988 statutory requirement that incorporated evaluation in the management of all programs funded by Europe's Structural Funds. In developing countries, the aid enterprise was and remains a consistent prime mover of evaluation capacity development.

Evaluation has gone global and in the process it has had to adapt to a wide variety of operating environments. A bewildering diversity of practices characterizes evaluation today. But

gradual progress is being made towards coherent evaluation standards across borders and sectors. Of course, the role of evaluators has been and is likely to remain context specific. Adaptability is the strength of our discipline.

Looking ahead evaluation will continue to be shaped by distinct organizational requirements and contrasting stakeholders' expectations. But there is

little doubt that we have much to offer to policy makers faced with the complex and interconnected challenges of a turbulent world. Contemporary evaluation methods reflect not only recent advances in economics, sociology, the political sciences and psychology but also the time tested traditions of the accounting and auditing professions. Evaluation is maturing and its future is bright.

JOIN TODAY !
www.europeanevaluation.org

The Policy for Monitoring and Evaluation of Primary Health Care in Brazil from 2003 to 2006: contextualizing its implementation and effects

Eronildo Felisberto(1), Eduardo Freese(2), Cinthia Kalyne de Almeida Alves(3), Luciana Caroline Albuquerque Bezerra(4), Isabella Samico(5)

1,3,4,5Research Group on Health Evaluation. Research Department. Instituto de Medicina Integral Prof. Fernando Figueira (GEAS/IMIP). Recife – Pernambuco, Brazil. 2Centro de Pesquisas Aggeu Magalhães/Fundação Oswaldo Cruz (CPqAM/Fiocruz). Recife – Pernambuco, Brazil.

The strengthening of policy for primary health care in Brazil developed during the last 10 years. At the municipal level, responsibilities for policy implementation established new roles for state and federal management. The Health State Secretariats (HSS) and the Ministry of Health (MH) took on more normative and regulatory responsibilities with emphasis upon the strategic function of evaluation of health systems. Then, there are various movements towards the institutionalization of evaluation on the field of primary health care. These actions stimulate changes in the practice of health professionals and organizations with the objective of qualifying the management process, health care and surveillance activities. All this contributes to the technical and scientific debate and it follows the subject of evaluation more recently.

Many initiatives were developed to implement strategic mechanisms with this objective and since 2003 the Ministry of Health of Brazil implemented a policy of evaluation of primary health care. The objective of this policy is the institutionalization of evaluation in the three management levels of the Unique Health System (SUS). It is developed on the basis of a process of elaboration and implementation of projects and activities to support the evaluation of policies and programs in primary health

care. This is characterized by a structured intervention through an organized system of action with its objectives, structures, actors, practices, process of action and context.

Taking all these considerations above, a study of the implementation of the Policy for Monitoring and Evaluation of Primary Health Care in Brazil during the period of 2003 to 2006 was carried. The aim of the study was to estimate the degree of implementation of the institutional mechanisms and the intervention, and evaluate their influence on the effects produced from their interaction with political and organizational contexts. This evaluative research (Implementation Analysis) was a case study. To estimate the degree of implementation, it was considered the institutional mechanisms produced that are representative of the components of the intervention logic model. In order to estimate the degree of implementation, a scale of points was used for each criterion: 1 = positive convergence of discuss; 0,5 = divergence of discuss; 0 = negative convergence of discuss. The degree of implementation of intervention was defined by the proportion between the sum of points achieved and the sum of maximum score of all mechanisms. It was considered four ranges of implementation: advanced: 75% - 100%; intermediate: 50% - 75%;

insufficient: 25% - 50%; incipient: 0 - 25%. 20 interviews and 29 documents were analyzed.

Findings showed that the implementation could be classified as advanced, with 78% of actions being completed. Institutional political context was characterized by a conjunction of political institutional decision, financing resources, means of technical and organizational strategies of human resources qualification for evaluation towards institutional capacity building. Findings show a significant amount of no-intentional effects as a consequence of various implementation mechanisms used. The Policy discloses a great effort towards actions integration on institutional level and to the decentralization of evaluation actions. In this way, two characteristics were relevant: the development of technical capacity by formative evaluative processes and the focus on the empowerment of the Health State Secretariats.

Key words: Implementation Analysis, Evaluation, Health Program Evaluation, Primary Health Care

Note: The full article of this research can be found at: Revista Brasileira de Saúde Materno Infantil; 9 (3):339-357, 2009.

Sixth European conference on Evaluation of Cohesion Policy, Warsaw, 30 November-1 December 2009

Veronica Gaffey, Head of Evaluation Directorate General for Regional Policy European Commission

This is a report on the successful 6th European Conference organised by the European Commission in Warsaw, entitled: "New Methods for Cohesion Policy Evaluation: Promoting Accountability and Learning".

Approximately 500 participants attended – policy makers, managing authorities, evaluators and academics – this sixth in a series of European Conferences, all with the overall objective of contributing to higher quality evaluations which can lead to more effective Cohesion Policy programmes. From former Commissioner Samecki's opening speech to the closing remarks by Danuta Hübner, MEP and chair of the European Parliament's regional policy committee, the consistent message was the need for evaluation to provide credible evidence on the performance of Cohesion Policy. Evaluation methods can make an important contribution to public policy making in general. For Cohesion Policy, evaluation needs to demonstrate how and to what extent programmes work and explore how they can work better –while letting the public know how taxpayers' money has been spent.

Other key messages included: the need for the European Commission as well as Member States and regions to deliver evaluation evidence; the need for political debate on the performance of the Cohesion Policy; the need to be clear on the objectives of the Policy and to generate evidence through evaluation on the achievement of those objectives.

A series of 8 workshops reflected on the experience gained in Cohesion Policy evaluation and had a critical look at the opportunities and challenges ahead, while also exploring more rigorous methods to evaluate impact.

Some highlights from the workshops:

Counterfactual impact evaluation (CIE) Practitioners presented and discussed the potentials, challenges and limitations of this technique. All agreed that it is an essential tool in the

evaluator's armoury, helping to answer the key question of what works and what does not. However, in the field of regional development, experience is still developing. The conclusion is to make CIE not only scientifically rigorous but also policy relevant – selecting target indicators that clearly reflect policy questions, rather than academic questions. For maximum policy relevance, CIE should be accompanied by theory-based evaluation to understand what determines the impact.

Theory-driven evaluation

The workshop explored whether this evaluation approach can be helpful in identifying the effects of an intervention. The discussion confirmed that programmes are socially constructed, with certain actors developing the programmes, to be implemented by other actors with different interpretations of programme logic. Theory-driven evaluation pays due attention to all actors' perspectives, establishing stronger cause-effect links. Limitations include that performance does not exclusively depend on the soundness of a programme theory; unintended effects have to be considered as well as the policy context.

Major projects

The aim of the workshop was to discuss challenges linked to correct cost and time estimates, as well as incentives to improve the quality of Cost-Benefit Analysis (CBA). Participants agreed on the importance of ex ante CBA as a tool for decision making and accountability. A discussion on how to improve the potential contribution of CBA to selecting good projects was initiated. Particular attention should be devoted to simplification of CBA, benchmarking and the use of ex post CBA.

Innovation

As innovation is a complex process involving multiple actors and requiring time to mature, identifying the results and impacts of activities supporting it is particularly challenging. This workshop looked into specific methods for evaluating innovation. Participants

agreed that the same methods as for other policy areas can be used, taking into account the specificities of innovation. They recommended triangulation of methods and tools for evaluation, including interviews, social network analysis and statistical and impact analysis.

Macroeconomic models

Three approaches were presented: the HERMIN and QUEST macroeconomic models as well as an example of econometric techniques used at national level. There was general agreement that the different approaches are complementary, attempting to explain the complex topic of Cohesion Policy impacts. Good macroeconomic models depend on sound theoretical foundations but also the quality of the underlying data made available by regional and national authorities. The discussion also highlighted that all the approaches have been upgrading their techniques, thus improving the state of the art.

Case studies

A common evaluation tool, case studies provide an explanatory picture of how policies work. In this workshop, practitioners discussed advantages and limitations of the tool. The experience shows that carrying out a case study is more complex and time-consuming than perceived, but can be particularly useful for delivering insights into the performance of policies.

Indicators

The type of indicators chosen has to be strategic and pragmatic: this was the key message of this workshop, through presentations addressing common issues and challenges from different perspectives. Participants agreed that caution should be applied when measuring impacts, as they materialise in a longer period of time and may be influenced by other factors. No perfect set of indicators exists, but it is possible to design indicator systems producing valuable information for both monitoring and evaluation.

Concepts for Cohesion Policy: GDP, Sustainable development, Intellectual capital

This workshop discussed which of these concepts really matters for Cohesion Policy, as well as whether measurement challenges for complex concepts can be overcome. Conclusions highlighted the need to improve data and indicators to complement the currently used GDP per head indicator; Cohesion policies of the future will also be judged on their contribution to improving quality of life and wellbeing – existing measurement methods should be further explored and appropriate indicators developed; the "intellectual capital framework" can be considered in identifying effects of Cohesion Policy.

Full coverage of the Conference and detailed presentations can be found under:

http://ec.europa.eu/regional_policy/conferences/evaluation2009/index_en.htm



TO FIND OUT MORE ABOUT
THE EES CONFERENCE IN
PRAGUE'
www.ees2010prague.org

Member Profile

We have introduced a new section in 'Connections' called, 'Member Profile'. Our aim is to put a human face on our membership, listen to their expectations and encourage contributions to the newsletter. We begin with

What is your name and occupation?

Charalabos Bakopoulos, Director of Langhouse
LangHouse is an educational consultancy, management, research and training center specializing in e-learning and learning difficulties. Langhouse is a team of educational researchers, learning difficulties experts, speech therapists and psychologists working across the whole education and children's services spectrum. We are enjoying working with young people and their parents and our aim is to help students reach their goals.

How long have you been an EES member?

Almost a year.

What do you hope to get from being an EES member?

I want to exchange ideas and be kept being informed about current evaluation issues.

What areas of evaluation most interest you and why?

I am most interested in educational evaluations of distance e-assessments, e-education and e-training because LangHouse shift of business is towards that.

Have you considered coming to the

EES conference in Prague this October?

Yes. I see this as an opportunity to exchange ideas and thoughts and be informed on different aspects and approaches on evaluation

What could we do to improve the experience of being an EES member for you?

I would value more active and interactive presence on the website.

If you would like to contribute a conversation to Member Profile, please contact....

The Limitations of Econometric Methods for Impact Evaluations

Scott Bayle, an evaluation specialist, provides quotations about 'Econometric Methods' from a diverse range of methodologists who identify limitations to this methodological approach.

Quotation:

Regression analyses reveal relationships among variables but do not imply the relationships are causal. Demonstrating causality is a logical and experimental, rather than statistical, problem.

From: Tabachnick & Fidell, 1989, Using Multivariate Statistics, Haper & Row.

Quotation:

Without an experiment, a natural experiment, a discontinuity, or some other strong design, no amount of econometric or statistical modeling can make the move from correlation to causation persuasive.

From: Jasjeet S. Sekhon. 2009, Opiates for the Matches: Matching Methods for Causal Inference; <http://sekhon.berkeley.edu/papers/opiates.pdf>

Quotation:

Causal inferences can be drawn from non-experimental data. However, no mechanical rules can be laid down for the activity. Since Hume, that is almost a truism. Instead, causal inference seems to require an enormous

investment of skill, intelligence, and hard work. Many convergent lines of evidence must be developed. Natural variation needs to be identified and exploited. Data must be collected. Confounders need to be considered. Alternative explanations have to be exhaustively tested. Before anything else, the right question needs to be framed. Naturally, there is a desire to substitute intellectual capital for labor. That is why investigators try to base causal inference on statistical models. The technology is relatively easy to use, and promises to open a wide variety of questions to the research effort. However, the appearance of methodological rigor can be deceptive. The models themselves demand critical scrutiny. Mathematical equations are used to adjust for confounding and other sources of bias. These equations may appear formidably precise, but they typically derive from many somewhat arbitrary choices. Which variables to enter in the regression? What functional form to use? What assumptions to make about parameters and error terms? These choices are seldom dictated either by data or prior scientific knowledge. That is why judgment is so critical, the opportunity for error so large, and the number of successful applications so limited.

From: Prof David Freedman, Professor of Statistics at University California Berkeley, 2008, lecture.

Quotation:

The 16 completed design replication studies offered a range of conclusions about the value of the nonexperimental

methods they examined and whether those methods produced findings that were similar to those of randomized experiments. Five studies concluded that nonexperimental methods performed well. Three studies found evidence that some nonexperimental methods performed well while others did not. The remaining eight studies found that nonexperimental methods did not perform well (or found insufficient evidence that they did perform well). Question: Do 'nonexperimental' methods replicate the findings from experimental impact evaluations? Answer: Occasionally, but not in a way that can be easily predicted.

From: Steven Glazerman Dan M. Levy, David Myers, 2002, Nonexperimental Replications of Social Experiments: A Systematic Review, Mathematica Policy Research Inc.

Quotation:

Econometric methods are a useful technique when undertaking exploratory research, generating hypothesis, developing theory, and for making predictions. However, these methods are problematic when undertaking impact evaluations. The results obtained from econometric analysis are heavily dependent upon the specific type of model chosen, and we usually have several different plausible models to choose from. Results are also very sensitive to violations of the chosen model's statistical assumptions, and such violations are the norm. Finally, the results obtained are strongly affected by the omission of relevant variables from

TO FIND OUT MORE ABOUT THE EES CONFERENCE IN PRAGUE'

www.ees2010prague.org



www.ees2010prague.org

The 9th European Evaluation Society International Conference
6th-8th October 2010, Prague, Czech Republic