

EVALUATIONS
CONNECTIONSGuest Editors: **Mita Marra and Kim Forss**. Editorial Assistant: **Eva Petrová**

EES

EUROPEAN
EVALUATION
SOCIETY

SPEAKING JUSTICE TO POWER

SEPTEMBER 2015

ARTICLES – NEWS – EVENTS

CONTENT

Presidential Message	1
Speaking Justice to Power: Will Evaluation Help Rebuild Europe's Social Protection System? An Editorial	2
Renewing Democratic Evaluation in the 21 st Century	3
Speaking Justice to Power: A Plea for Evaluating Gender Equity	5
Getting Good Data to Evaluate Equity	6
Equity Monitoring Systems: Assess before you use	7
Using Theories of Reach to Enhance a Focus on Equity in Evaluations	8
Evaluating Equity in a Development Context	10
The Authors	12

PRESIDENTIAL MESSAGE

Dear Members and Colleagues

The heated and agitated summer in Europe – exceptional high temperatures linked to global warming and turmoil in the Eurozone – brings once more the need for effective and influential evaluation to the fore. This Special Edition **Speaking Justice to Power** explores the social forces underlying these events.

As noted by Kim Forss and Mita Marra in their editorial the issue of equity is still neglected in evaluation work; yet this edition comprises articles highlighting efforts for better integrating the equity and gender perspectives in evaluations such as the design of methodological tools – the Theory of Reach elegantly presented by John Mayne – as well as how to carefully select measures and indicators that are based on these values along with technical judgements for monitoring, evaluations and studies.

Evaluation needs to be democratic too, i.e., the conduct of evaluation needs to be democratic, but evaluation needs also to evolve within democratic institutions. Robert Picciotto's article nicely shows how the concept of democratic evaluation has evolved and he comes up with a promising new model: *independent and equity-focused democratic evaluation*.

But is it possible to discuss evaluation, equity and democracy in Europe without addressing the Greek situation? This on-going crisis

raises a whole range of questions. Why did politics and policymaking fail in securing the Eurozone? Where were the rating agencies, the economists, the auditors and the statisticians? What are the implications for evaluation governance? How should evaluation be shaped to strengthen democratic processes in Europe? These are difficult and complex questions, but the Greek crisis is a tough awakening and calls for in-depth thinking on the role of evaluation in the international space.

This is the overarching objective of the 2015 International Evaluation Year. Your Society has planned several activities. On September 17th in Helsinki, we will co-host a seminar: *"The Future of Evaluation: A Nordic Perspective and Beyond"* which will address the future of evaluation, and develop a Nordic contribution to the Global Evaluation Agenda. On September 19th, the first virtual conference of Emerging Evaluators led by our own Thematic Working Group will probe a related theme *"The Future of Evaluation for the Future"*. The conference is already booked with a waiting list, and the conference website (www.emergingevaluators2015.org) is up and running!

On September 30th, in collaboration with the OECD, UNESCO and the French Evaluation Society we have organised a joint seminar: *"Making effective use of evaluations in a complex world"* to be held in Paris. For more information on these events, please visit our Website

www.europeanevaluation.org. On October 28th the Institute for Managing Sustainability (Vienna University of Economics and Business), in collaboration with the Austrian Development Agency and the EES, will sponsor a symposium in Vienna about “*Evaluating the Sustainable Development Goals – New Challenges for Research, Policy and Business*”. For details please click: www.sustainability.eu/?k=SDG_symposium&s=home&u=&lang=en.

And last but not least, our pilot initiative, the Voluntary Evaluator Peer Review, which we launched jointly with the UK Evaluation Society will soon be in the process of

recruiting EES members who volunteer to be reviewed. Let me underline that this is not a test -pass or fail. Instead it aims to help enhance and develop evaluators’ skills. We will also need volunteers with extensive experience to carry out the evaluators’ review. More details about this process will be provided in a few weeks through emailing and/or you can visit our Website: www.europeanevaluation.org/search/node/VEPR

Finally EES will be present at the closing event of the Evaluation Year where we will be active in shaping the Global Evaluation Agenda. Let me also take this opportunity to remind you that

our 12th EES Biennial Conference: **Evaluation Futures in Europe and beyond: Connectivity, Innovation and Use** will take place in Maastricht, The Netherlands on 28, 29, 30 September 2016, with the pre-conference on 26 and 27 September.

And lastly I very much hope that you’ll enjoy reading this Special Edition of Connections and that you will see fit to use our Website to interact with evaluation colleagues and send us feedback about EES work www.europeanevaluation.org.

Claudine Voyadzis, President

SPEAKING JUSTICE TO POWER: WILL EVALUATION HELP REBUILD EUROPE’S SOCIAL PROTECTION SYSTEM? AN EDITORIAL

Kim Forss and Mita Marra

Europe has long understood the importance of addressing equitable issues by providing a system of social protection. Most European countries have recognized that good systems of social protection can lead to improved overall economic performance, as individuals are more willing to take the entrepreneurial risks that lead to higher economic growth.

Yet, in many parts of Europe, today, high unemployment (11% on average, 25% in the worst-affected countries), combined with austerity-induced cutbacks in social protection, has resulted in unprecedented increases in vulnerability and social inequalities. The implication is that the decrease in societal well-being may be far larger than that indicated by conventional GDP measures – numbers that already are bleak enough, with most countries showing that real (inflation-adjusted) per capita income is lower today than before the crisis – a lost half-decade.

Last October, the 11th EES Conference held in Dublin explicitly addressed the issue of social equity as the central theme for evaluators to articulate in many plenary and parallel sessions. This emphasis is warranted since it is still overlooked in evaluation work. Of the three Es pursued by governments around the world – efficiency, economy and equity, both

policy analysts and evaluators have devoted far more effort to measuring and evaluating the first two. With the intent of helping to redress the balance we edited for Transaction Publishers a volume of essays entitled: “Speaking Justice to Power. Ethical and Methodological Challenges for Evaluators.”

In theory, equity should be one of the main criteria in the assessment of worth and merit. Equity is concerned with grand societal themes such as the distribution of power and influence, the rights and duties of people, and the quality of life. These themes have attracted the attention of influential thinkers since ancient times: Plato and Aristotle, Thomas Aquinas, Machiavelli, Rousseau, Smith, Marx, and in recent times Rawls, Nozick, Nussbaum and Sen, etc. The same themes should attract – and challenge – evaluators too.

Most of us refer to equity when we talk about public policy; the tax structure, recruitment to schools, participation in elections, access to health care and the way we are treated in it – to name but a few subjects. We often refer to equity when we talk about the private sector as well, and equity is a central concern in many civil society organizations. Equity is everywhere. But does it play an important role in the evaluation discourse?

To be sure, the concept of equity entails ambiguous value considerations, and it is culturally grounded. Thus it is challenging for evaluators of public program to operationalize, contextualize, measure and judge it. Even stakeholder engagement, an ubiquitous practice for evaluators, does not necessarily reduce, and may even increase, the challenge of measuring equity fairly and consistently (Newcomer 2014).

The report by the International Commission on the Measurement of Economic Performance and Social Progress (chaired by Nobel Prize Joe Stiglitz) emphasized that GDP is not a good measure of how well an economy is performing. UNDP reports remind us of the importance of this insight. They demonstrate that too much has been sacrificed on the altar of GDP fetishism. Regardless of how fast GDP grows, an economic system that fails to deliver gains for all of its citizens, and in which a rising share of the population faces increasing insecurity, is, in a fundamental sense, a failure. Equally policies, like austerity, that increase insecurity and lead to lower incomes and standards of living for large proportions of the population are deeply flawed.

What contribution can evaluation offer to address such equity-based issues, specifically

within the EU? What value basis can evaluators refer to when they analyze results and impact? How do evaluation teams arrive at conclusions on fairness, justice, equity and equality, and how can those processes be made transparent and gain credibility and legitimacy? What are the challenges evaluators face in data collection and analysis when pursuing the equity questions?

“Speaking justice to power” seeks to address these controversial issues organizing chapters around three themes. The first is ‘Equity in Theory: What Evaluation Implications’. It presents different social science strands that have underpinned such notions as social justice, equity and equality to the extent to which they are used in evaluation practice. The second “Equity in Evaluation Approaches: What Challenges for Evaluators?” highlights equity in evaluation approaches, evaluators’ mandate as well as stakeholders’ expectations around evaluation. The third “Equity in Program Evaluation: Lessons Learnt” presents real world evaluations, which have dealt with equity-related issues, such as climate change, fair trade, reproductive health, employment promotion, child rights, and the rights of ethnical groups and minorities, as well as women in international aid programs.

For all the authors of this book public policy is not merely a set of mechanistically

juxtaposed measures. It is instead conceived as endogenously generated measures implemented within complex contexts, where values, rules, and social norms shape social change. Authors’ efforts to incorporate equity into evaluation work range over vastly different geographical settings. They have also addressed equity in interventions from social and environmental perspectives. The articles in this special issue of Connections draw on their insights with a view to influence evaluation work in the European space and to contribute to European political, economic and institutional reform.

The articles in this special issue of Connections draw on authors’ insights with a view to influence evaluation work in the European space and to contribute to European political, economic and institutional reform. In particular, from a democratic perspective, Picciotto provocatively stigmatizes dominant public policy paradigms and related evaluation approaches. The latter are underpinned by those very premises, which do not adequately promote inclusion (working with underrepresented and powerless groups); dialogue (getting stakeholders to understand each other) and deliberation (reasoned discussion of issues, values and findings). Marra’s article is a plea to address gender equity through a multi-level approach to assess power relations and the distribution of work and caring between men

and women. By referring to the Canadian experience in evaluating the access of underrepresented groups in the federal public service, Barrados stresses the importance of defining the groups or individuals that are underrepresented, while getting good data to measure the size of the designated groups and the proportion expected in the public service. Wilkin uncovers the different perspectives underlying monitoring and reporting systems, which – even as framed in terms of poverty, opportunity, well-being, progress, disadvantage, quality of life, social inclusion, development and capabilities – may still miss issues for particular sub-populations, and conceal highly relevant trends. Mayne highlights the methodological relevance of theories of reach, which describe how and why the activities undertaken by a program are expected to (or have) lead to the realization of the desired results for a specific reach group. Finally Klove’s article wraps up this special issue’s collection, relaunching universalism as the welfare state approach that in Europe and, specifically across Nordic countries, over several decades has proved to assure more equitable outcome within society. This contribution, which is not drawn on the book chapter, may inspire a new social equity policy agenda both within Europe and throughout developing countries. ■

RENEWING DEMOCRATIC EVALUATION IN THE 21ST CENTURY¹

Robert Picciotto

Is evaluation up to the challenge of an unprecedented social crisis that is sparing no region and no country? Are current evaluation approaches fit for purpose? What is to be done? To address these questions this article surveys global trends, examines the trajectory of evaluation policy and considers the adequacy of dominant evaluation models.

Whereas in 1942 there were only twelve democracies 122 countries or 63 percent of the 195 United Nations members are now classified as electoral democracies². However progress has stalled. For the eighth

consecutive year Freedom House’s annual report has reported setbacks³. Modern authoritarianism is on the march. Its appeal lies in promises of rapid economic development.

Thirteen countries have managed to grow for twenty five years or more at an average annual rate of 7 percent or more using export led strategies⁴. The list includes China where the Communist Party has a stranglehold on power; Hong Kong and Singapore classified by Freedom House as ‘partly free; Oman where political parties are banned as well as Thailand where democracy is under threat.

Vested interests have captured the commanding heights of the global economy. They are increasingly undermining the workings of western democracy. Conflict of interest is now rife within the evaluation world. For example pharmaceutical companies’ controls over evaluation have introduced systematic biases in the conduct of drug control trials⁵. Evaluation practice has not been immune to these trends.

Dominant public policy paradigms have always shaped evaluation concepts and policies⁶. Don Campbell’s “*experimenting society*” was emblematic of the 1950’s. A *dialogue-oriented*,

	Traditional democratic evaluation	Deliberative democratic evaluation	Independent and equity focused democratic evaluation
Authorizing environment	Liberal democracy; “town meeting” or village pump democracy.	Partial or illiberal democracy; developmental state.	Authoritarian, patrimonial or captured by vested interests.
Evaluator role	Neutral brokering; facilitation, etc.	Activist; inclusive; empowers disadvantaged groups.	Autonomous. Owns evaluation product.
Fee dependence on program manager	Yes	Yes – mostly	No. Funded independently.
Recommendatory	No	Yes	Yes as part of civic engagement in decision making.
Advocacy	No	Yes – to level the procedural playing field.	Yes – as commitment to a more equitable and democratic society.

Table I: Democratic evaluation models.

participatory and pluralistic wave surged in the late 1960’s. The political winds shifted to the right in the eighties and a *neo-liberal* wave engulfed the evaluation discipline. We are now surfing an *evidence based* wave that restores experimentalism as the privileged approach.

These trends have affected evaluation thinking and practice everywhere. Private foundations have promoted the export of experimental evaluation to the developing world: the internationalization tide and the evidence based wave have merged. Evaluation is now “international in the sense of being at the same time more indigenous, more global and more trans-national”⁷. By the end of 2012 the International Organization for Cooperation in Evaluation had identified 114 evaluation groupings. Internationalization has also propelled evaluation towards patrimonial governance frameworks and illiberal regimes⁸.

Counter-currents have begun to materialize. Disenchantment with neo-liberal policies is spreading. Inequality is high and getting worse⁹. The hollowing of the middle class is generating popular pressure for change. By now, forty percent of the world’s wealth is owned by the richest one percent of the population while the poorest half own only one percent. Over a third of the wealth is owned by the richest 0.5% of the world population while more than two thirds of the world’s population (68%) share only 4%.

From a democratic perspective the three currently dominant approaches to evaluation fall short. The first emphasizes *accountability and compliance*. It examines how public resources are used to reach goals that are almost invariably set by the prevailing power structure. The second focuses on the pursuit

of *knowledge*: akin to social science research it emphasizes attribution oriented evaluations that evoke the *value free* scientific approach. Similarly the *utilization- focused evaluation* model¹⁰ is akin to management consulting and it has contributed to the social timidity of evaluation agendas.

What about the Barry MacDonald democratic evaluation model? It was full of promise when it arose in the mid-seventies in the United Kingdom. It is a non-recommendatory information service to the community. It tasks evaluators to act as brokers between differing groups. It offers confidentiality to informants and gives them control over the information. It works well in authorizing environments where communicative rationality prevails and ethical discourse influences decision making. But given its neutral brokering stance it is ill adapted to the very contexts where it is most needed¹¹.

To promote the interests of the least fortunate Ernest House refined the MacDonald model. His model stresses three principles: *inclusion* (working with underrepresented and powerless groups); *dialogue* (getting stakeholders to understand each other) and *deliberation* (reasoned discussion of issues, values and findings). In this new, value based incarnation the “evaluator is not a passive bystander, an innocent facilitator, or a philosopher kind who makes decisions for others, but rather a conscientious professional who adheres to carefully considered principles”¹². Undoubtedly this activist stance is better adapted to authorizing environments that are partially democratic.

However both of these democratic evaluation approaches are hobbled in contexts that

do not tolerate dissent and/or for assignments that are closely controlled by evaluation sponsors. In such situations progress towards liberal democratic ideals calls for an alternative model: independent democratic evaluation. It would embrace the vision of evaluation as a morally engaged occupation that amplifies the voice of citizens, resists capture by vested interests and promotes equity.

Independent democratic evaluation would break the chains of fee dependence; assert its autonomy in shaping evaluation agendas and methods and emphasize professional independence as well as fulsome engagement with citizens. It would draw on the procedural neutrality, process innovations and ethical canons of deliberative democratic evaluation but it would not shy away from advocating measures or recommending actions at the service of democratic ideals.

Evaluators operating according to this model would assume ownership of evaluation products and reject assignments that report to decision makers in charge of the intervention being evaluated, reporting instead to a supreme authority (board of directors, parliaments, etc.) or an entity that stands at arm’s length from the intervention (e.g. a non-governmental organization in the case of a government sponsored programme)

Undoubtedly all existing evaluation approaches have useful features and fulfil valuable services. But renewing democratic evaluation so that it is fit for purpose in contexts where democracy is absent, social inequities are rampant and/or governance has been captured by vested interests calls for a new approach: independent democratic evaluation (table I).

1. A longer version is available in the special issue of the Evaluation Journal dated April 2015 (Volume 21, Number 2).
2. <http://www.freedomhouse.org/sites/default/files/Electoral%20Democracy%20Numbers%2C%20FIW%201989-2014.pdf>.
3. <http://www.freedomhouse.org/report/freedom-world/freedom-world-2014#.U76nmZRdUuc>.
4. At this rate incomes double every decade.
5. House, E.R., *Blowback, Consequences of Evaluation for Evaluation*, American Journal of Evaluation, December 2008 vol. 29 no. 4 416–426. Sage Publications.
6. Evert Vedung, *Four Waves of Evaluation Diffusion*, Evaluation, July 2010, 16:263–277.
7. Chelimsky E. and Shadish W., (1997) *Evaluation for the 21st Century*, Sage Publications, London and Thousand Oaks, California.
8. Feinstein O. and Porter S. (2014), *Reframing Evaluation Capacity Development in Africa*; Evaluation Connections, European Evaluation Society, July (http://europeanevaluation.org/sites/default/files/ees_newsletter/ees-newsletter-2014-06-july-05-web.pdf).
9. Pedro Olinto and Jaime Saavedra, *An Overview of Global Inequality Trends*, World Bank, *Inequality in Focus*, Volume I, Number I, Washington DC, April 2012.
10. Patton, M.A. (2008), *Utilization-Focused Evaluation*, 4th Edition, Thousand Oaks, California. Sage Publications.
11. Simons, H., (2010). *Democratic Evaluation: Theory and Practice*, Paper prepared for the Virtual Evaluation Conference, University of the Witwatersrand, Johannesburg, South Africa, May.
12. House E.R. and Howe K. R. (1999) *Values in evaluation and social research*, London and Thousand Oaks, California: Sage Publications.

SPEAKING JUSTICE TO POWER: A PLEA FOR EVALUATING GENDER EQUITY

Mita Marra

By the mid-1990s, the EU gender mainstreaming approach had promoted activation and social inclusion programs to balance work and family, as well as subsidies for productive activities across regions and localities. The current challenge is to evaluate these policies through the lenses of gender equity in a systemic way, considering both equality of opportunities and equality of outcomes for men and women.

Evaluation approaches assessing gender-sensitive policies have thus far adopted both feminist and gender frames of reference, which present many overlaps but also major differences concerning methods of data collection and analysis let alone the role of the evaluator in using evaluative knowledge within the political arena.

Gender evaluation adopts mixed methods with an emphasis on quantitative methods to identify and analyze sex-disaggregated indicators, within countries or regions. Its major methodological challenge concerns estimation of non-market variables with respect to family care and unpaid work. This is essential in order to quantify the full contribution that women make to society even as they participate to a lesser extent than men in the

labor market. Feminist evaluation tends to be more qualitative and based on micro units of analysis, with a strong orientation towards action-research. It focuses on the uniqueness of women's lives and tends to favor qualitative inquiry based on dense and detailed descriptions of contexts and life stories that defy generalization.

Yet, both types of evaluations while rigorous and socially relevant have failed to draw up comprehensive comparative analyses across different political and institutional contexts. Nor have they influenced decisions at the policy level or helped in the design of broad-based strategies capable of modifying traditional gendered power relations, roles and stereotypes. In part this is because the adoption of these approaches has not been systematic and widespread in the evaluation community. Mainstream evaluators do not normally have expertise in gender analysis except for those who make up the networks of the so-called “femocrats” – the professionals who specialize in gender studies mostly within international organizations. Thus, both feminist and gender-based evaluations risk being perceived as ideological, rhetorical or technocratic exercises with limited influence on policy and evaluation circles.

In a recent article (Marra, 2014) I have proposed a multilevel approach to assess gender equity designed to even power relations and re-balance the distribution of work and caring. I contend that the issue is not developing new methods, but rather thinking differently about how to assess gender equity. In Europe as well as in North America, gender inequalities feed into social inequities (like racism, social-class inequalities). The challenge is to uncover whether and how program theories of change address existing inequalities emerging out of layered interpersonal, institutional, socioeconomic, and political interdependencies. Analyzing genders' cooperation in different realms would help characterize the multidimensional notion of social equity in the evaluation of public policies, raising attention on ethical criteria and standards that can be shared within the evaluation community.

From this perspective, power relations, perceptions of gendered roles and the ways in which men and women cooperate in the choice of work and caring are micro-dimensions of social equity. Any evaluation approach can then explore in the micro domain if and how the space for freedom for men and women is guaranteed and broadened. The level of socioeconomic and human

development of a region or country are 'meso' and 'macro' dimensions of social equity that may hinder or accelerate the realization of individual aspirations for professional development and family well-being as well as democratic social, political and institutional processes over time.

These factors interact in complex ways and affect all actors, who bear perceptions, opinions, cultures and values. Crucial is, therefore, to investigate the cultural, motivational, organizational, institutional and political contexts, in search of those emotional, cognitive, and economic mechanisms that generate, sustain, and propagate social cooperation – and specifically cooperation between men and women – over time and through different spaces.

Exploring gender inequities in different realms raises the question of whether gender sensitive ethical principles need to be explicitly shared within the evaluation community over and above the principles of systematic inquiry, competence, honesty, integrity, respect for people, responsibilities for general and public welfare, etc. as articulated by AEA (1995).

Recently the United National Evaluation Group's (UNEG) pointed the way in a Handbook entitled "Integrating Human Rights and Gender Equality in Evaluation – Towards UNEG Guidance" (2011) that provides step by step guidance on how to integrate these dimensions throughout an

evaluation process, helping, in this way not only evaluators, but also evaluation managers, and program managers both within and outside the UN system.

Along with "inclusion" and "participation," the Handbook explicitly articulates the key principle of "fair power relations." Such gender sensitive guidelines will raise awareness and increase compliance over time especially when combined with other initiatives in education, lifelong learning, and ongoing debating within the evaluation community and society as a whole so as to really make a difference against ingrained preconceptions, social norms and stereotypes towards a more equitable society.

References

American Evaluation Association (1995) "American Evaluation Association Guiding Principles For Evaluators," Task Force of the American Evaluation Association, available on www.eval.org/p/cm/ld/fid=51.

Forss, K. and Mita Marra (2014) [Eds.], *Speaking Justice to Power. Ethical and Methodological Challenges for Evaluators*. Comparative Policy Evaluation, Volume 21, Transactions Publishers, New Brunswick, US.

Marra, M. (2012) "The Missing Links of the European Gender Mainstreaming Approach. Assessing Work-Family Reconciliation Policies in the Italian Mezzogiorno."

European Journal of Women's Studies 19, no. 3: 349–70.

Nussbaum, M. C. 2000. *Women and Human Development in the Capability Approach*. Cambridge: Cambridge University Press.

Podems, D. R. 2010. "Feminist Evaluation and Gender Approaches: There's a Difference?" *Journal of Multidisciplinary Evaluation* 6, no. 14: 1–17.

Robeyns, I. (2007) "When Will Society Be Gender Just?" In *The Future of Gender*, ed. J. Brown, 54–75. Cambridge: Cambridge University Press.

Seguino, S. 2013. "Toward Gender Justice: Confronting Stratification and Unequal Power." *G. neros – Multidisciplinary Journal of Gender Studies* 2, no. 1: 1–36.

Sen, A. 1999. *Development as Freedom*. New York: Albert A. Knopf.

Sielbeck-Bowen, K. A., S. Brisolara, D. Seigart, C. Tischler, and E. Whitmore. 2002. "Exploring Feminist Evaluation: The Ground from Which We Rise." *New Directions for Evaluation* 96: 3–8.

United Nations Evaluation Group (UNEG) (2011) *Integrating Human Rights and Gender Equality in Evaluation. Towards UNEG Guidance*, UNEG Human Rights and Gender Equality Task Force, New York. ■

GETTING GOOD DATA TO EVALUATE EQUITY

Maria Barrados

Equity of access to and enjoyment of societal benefits are often expressed social goals in National Charter documents such as Constitutions and Bills of Right. These more general statements of principle are often further set out in laws, policies and programs. Many different approaches have been taken by governments to improve societal equity in different parts of the society.

Initiatives to improve the equality of access and participation of citizens in the benefits of employment have included affirmative action

programs where quotas have to be filled, access to overrepresented groups or individuals have to be restricted or limited, or goals and objectives that should be worked towards have to be formulated. For all these policy options the expected outcome is improved participation of underrepresented individuals or groups.

As is often the case the implementation of whichever option is chosen faces a number of challenges, many of which in turn become challenges for the evaluator in the assessment of effectiveness. If program designers

and managers have not resolved these issues, the challenge for the evaluator becomes even greater.

Even though decision makers, policy makers and program managers turn to evaluation to assess whether their programs and policies are having the desired results. A number of societal groups are often equally interested in the evaluative results.

The interest in the results of the evaluation of an equity issue is often high since with a limited

number of positions in the case of employment equity, for example, increasing access and representation of one group goes with reduction of the members of the overrepresented group. Evaluation can be seen an important arbitrator in often charged societal debate.

The experience in Canada in evaluating the effectiveness of legislation and policies to improve the access of underrepresented groups in the federal public service highlights some important methodological and measurement lessons.

The definition of the key elements – the groups or individuals that are underrepresented and what is to be improved – are the building blocks for programing and evaluation. In Canada rather than “affirmative action” the policy and legislative approach taken was equity of access and participation through continuous improvement. The legislation and regulations set out the expectations for programs and initiatives to provide greater equity and hence greater equality for four designated (considered disadvantaged) groups in Canadian society – aboriginal people, persons with disabilities, visible minorities, and women. It was generally accepted that

the definitional approach in the legislation of self-identification was a reasonable one.

Once the key concepts are defined (which in the Canadian case was done through legislation) the challenge of getting essential data to measure the size of the designated groups and the proportion expected in the public service (the targets) arises. Since the approach taken was setting goals and targets that program managers were expected to work towards, the measures that set the baseline and the measures that assessed improvement became essential tools for managers. Both good measures of the population of designated group members and of the population that would be expected to be reached if there were equity proved difficult to secure.

Addressing this challenge in Canada took time. It required work with the National Statistics Agency and the government department responsible for labor as well as special data collection. A time series of data supplemented with additional special periodic evaluations gradually evolved.

Initial assumptions about data quality turned out to be incorrect. The evaluators learned

the importance of being prepared to closely re-examine all data and methodologies even if it meant questioning existing time series. A closer reexamination of the data and underlying methodologies pointed to problems with data collection that had not been recognized including one of the data series that had been used for a number of years and had not been challenged. A move to automatically collect background information at the time of application resolved the issues of unknown level of participation. Further data difficulties were identified as result of an early incorrect interpretation of survey results. This again resulted in recalculation and refinement.

Flawed data that appear credible and are used repeatedly confirm existing beliefs and undermine the validity of evaluation. The experience with the evaluation of the success of initiatives to improve the representation of disadvantaged groups in the public service of Canada underscores the importance of ongoing vigilance and care with data and their interpretation. Close examination is needed to uncover underlying methodological problems and to overcome data weaknesses. ■

EQUITY MONITORING SYSTEMS: ASSESS BEFORE YOU USE

Peter Wilkins

Equity monitoring systems have been framed in terms of poverty, opportunity, well-being, progress, disadvantage, quality of life, social inclusion, development and capabilities. In many cases there is little explanation of why a particular approach has been used, and a potential user of the information needs to assess its construction and presentation before putting it to use.

Lessons in this regard have been identified from a review of ten monitoring systems that address equity in school education, including the varied approaches adopted, how the needs of equity groups are identified, how indicators and targets are selected and how embedded values in the choice of goals and measures can be addressed (Wilkins 2014).

Each approach to equity monitoring results in a different perspective both in terms of what information is reported and how it is presented. For instance, rankings in league tables are an important component of the OECD presentation of its Programme for International Student Assessment (PISA) data and ranking in the table is highlighted in public debates and plays a significant role in shaping public policy in many countries. Australia’s Overcoming Indigenous Disadvantage (OID) Key Indicators reports are based on comparisons between Indigenous and non-Indigenous data. They focus efforts on six high-level Closing The Gap targets, for instance the target for “reading, writing and numeracy” is “Halving the gap for Indigenous students in reading, writing and numeracy within a decade”. The focus on disadvantage is not without its critics, most

notably that it focuses on deficits without looking at strengths.

Reporting systems that give emphasis to average performance for particular populations have the potential to miss issues for particular sub-populations, and composite indices can similarly conceal particular trends. A United Nation’s Development Program (UNDP) research paper on criticisms of the Human Development Index (HDI) noted that “A large group of critiques is concerned with the fact that the current HDI presents averages and thus conceals wide disparities in distribution of human development in the overall population” (Kovacevic 2011). An example of how limitations of this kind can be addressed through analysis and presentation of the data is provided in the presentation of an overview

of headline well-being indicators in OECD countries as part of the Better Life Initiative, the OECD using what it terms “traffic-lights,” these being for each indicator for each country an indication of whether the country is in the top two deciles, the bottom two deciles, or in the six intermediate deciles (OECD 2011).

The OECD “Your Better Life Index” addresses the problem of value judgements in the weighting of particular dimensions of well-being in its online service by enabling the user to give their own weight to each of the eleven dimensions involved and then produce information on their own country and make cross-country comparisons incorporating these personalized weightings (OECD 2012).

Values and technical judgement are also integrally involved in the selection of measures and indicators. Often, however, there is little explanation of the basis of indicator selection and how targets are set. An exception is the setting out of specific selection criteria by the OECD for its Better Life Initiative indicators. The OECD presents the case for individual indicators, which helps users to assess the appropriateness of each indicator (OECD 2011).

An innovative approach is evident in the Human Opportunity Index (HOI) developed by the World Bank for Latin America and the Caribbean that focuses on equal opportunities early in life, educational achievement being analyzed in terms of a component resulting from a set of predetermined circumstances (gender, mothers and fathers education, fathers occupation, and school location) and a second component encompassing individual efforts, talent, and luck. The HOI focuses on the ex-ante question how likely it is that children will fare well, in contrast to the Human Development Index ex-post focus on how well adults have fared. For instance, the HOI analysis provides insights into whether a girl’s probability of completing the sixth grade at school (a predictor of higher education) is affected by her race, mother’s literacy, or her father’s salary (World Bank 2009).

Equity monitoring systems are an important means of making transparent the challenges and progress involved, and prompting questions that may warrant complementary evaluations. However, users of the monitoring information need to be vigilant about how the information is shaped and presented.

References

- Kovacevic, M. 2011. Review of HDI Critiques and Potential Improvements. Human Development Research Paper 2010/33. http://hdr.undp.org/en/reports/global/hdr2010/papers/HDRP_2010_33.pdf.
- OECD 2011. *How’s Life? Measuring Well-Being*. Paris: OECD.
- OECD 2012. *Measuring Well-Being and Progress*. <http://www.oecd.org/std/Measuring%20Well-Being%20and%20Progress%20Brochure.pdf>.
- Wilkins, P. 2014. *Monitoring Equity and the Evaluation of Equity Monitoring Systems. Speaking Justice to Power: Ethical and Methodological Challenges for Evaluators*. K. Forss and M. Marra, Eds, Transaction Publishers.
- World Bank. 2009. *Measuring Inequality of Opportunities in Latin America and the Caribbean*. http://siteresources.worldbank.org/LACEXT/Resources/258553-1222276310889/Book_HOI.pdf.

USING THEORIES OF REACH TO ENHANCE A FOCUS ON EQUITY IN EVALUATIONS

John Mayne

Interventions often involve a number of different target groups whose behaviour and/or well-being they hope to change, as well as other groups who may be affected by such targeting. Keeping track of the impacts and effects on these various groups may not be straightforward, especially as an evaluation progresses. It is easy in these circumstances to lose sight of (or pay insufficient attention to) key beneficiary groups. This note discusses an evaluation tool derived from theory of change concepts that can help evaluations and evaluators focus on and assess equity aspects of an intervention, and in particular the distribution of impacts and effects from interventions. The tool also provides a way to sensibly disaggregate more complex interventions.

Theories of change are standard tools for evaluators. The left side of Figure 1 illustrates

a generic theory of change that has proven useful in a number of settings, showing a results chain from activities to reach to capacity development to behavioural change to direct benefits to livelihood changes, along with the associated causal link assumptions. (Mayne 2011; Mayne 2012).

For interventions involving activities aimed at influencing several different target groups (beneficiaries, institutions, organizations and partners), one approach would be to try and develop a theory of change that captures all these activities as they relate to various target groups and the resulting outcomes and impacts, thus identifying the links among the various pathways. However, developing and setting out such a model in other than as a general overview while helpful would be quite challenging, and the resulting complex

theory of change model could become cumbersome and hard to work with, either for explaining the intervention or for helping design the evaluation.

Instead, and much more useful, would be to develop a results chain for each target or reach group with an accompanying theory of reach, recognizing that these theories of reach may interact with each other in bringing about the desired results (Mayne 2014).

A *reach group* is a group of common or similar individuals, partners, organizations and/or institutions whose behaviours the program is trying to influence. These include those whose livelihoods the intervention is trying to improve and others whose support is needed for this to be realized. A *reach chain* is a model showing the links between the

activities undertaken by the intervention and the sequence of outcomes and impacts on a specific reach group.

Similarly, a *theory of reach* is a theory of change illustrating and describing how and why the activities undertaken are expected to (or have) lead to the realization of the desired results for a specific reach group. As indicated in Figure 1, theories of reach would be sub or nested theories of change within the larger intervention theory of change, focusing on a particular reach group. Developing and monitoring such theories of reach would focus attention on the reach groups of interest. Mayne (2014) discusses an example of an intervention aimed at improving education outcomes for girls through improves teacher training, where there is a special interest in the group of worst-off girls. Theories of reach could be developed for the teachers, for girls, for boys (to monitor unintended impacts) and for worst-off girls.

Complex sufficient interventions (Mayne 2014) aim to engage and influence relevant NGOs, civil societies, government actors and the private sector to ensure that an enabling environment for success is forthcoming, as well as the more traditional efforts to build the capacity of their reach groups. Thus, for example, providing training to women in cottage handicrafts will not lead to their betterment without availability of financial credit and working markets for the products, as well as supporting policies on small businesses and trade. Complex sufficient interventions are designed to address the various causal link assumptions in the intervention's theory of change so that the *collective* actions of the intervention actors and their partners are sufficient to bring about the desired benefits and impacts. Figure 1 illustrates a generic theory of change for such interventions.

When complex sufficient interventions are being evaluated and equity is an issue Figure 1 makes clear that there are quite a few nested theories of change (theories of reach for the different beneficiaries as well as for intermediaries one tries to influence) that could be usefully developed. As discussed, those for the worst-off reach groups would be particularly important to develop so as to monitor the extent to which the theory of reach is being realized. In this way, the distribution impacts and effects of the intervention would

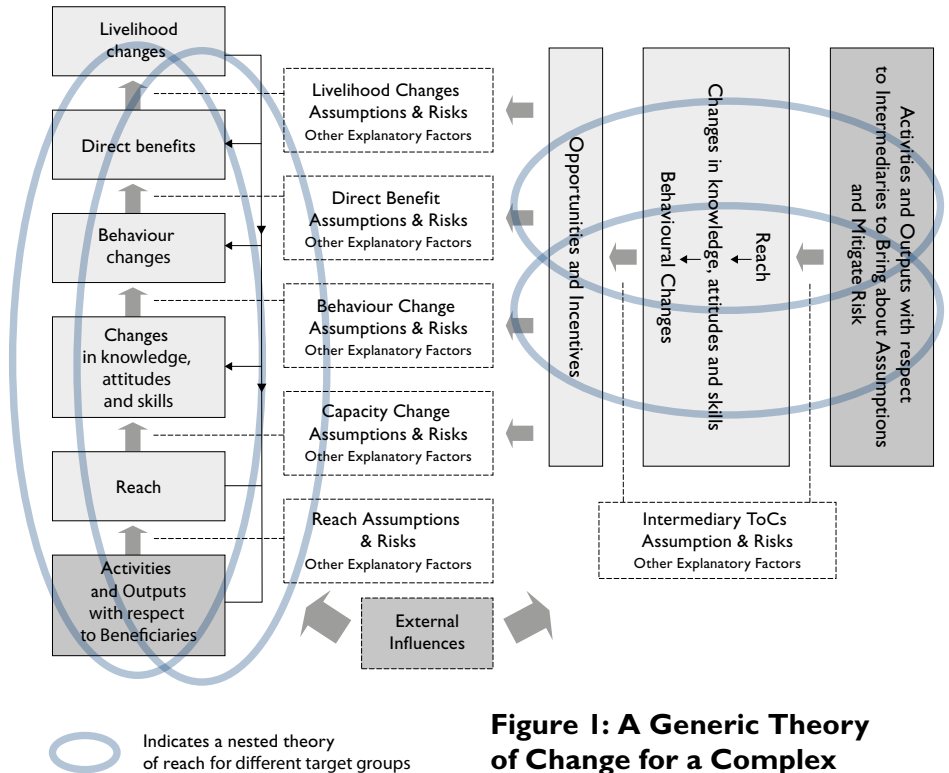


Figure 1: A Generic Theory of Change for a Complex Sufficient Intervention.

be assessed. Such nested theories of change also provide a sensible way of disaggregating an otherwise quite complex intervention.

Reach chains and accompanying context and theories of reach can all be usefully encompassed in a reach story. A *reach story* describes what results an intervention has had on a specific reach or target group. It does not try to present a complete picture of the impacts of the intervention, nor how all results were brought about. It is, however, a very useful way to capture, track and discuss the effects of an intervention on, for example, worst-off groups. A reach story can also be an effective way to communicate the value judgements that underlie the evaluation, discuss the extent of participation by the reach group, and to raise the political issues around the distribution of impacts and effects

Issues of equity are rightly gaining increasing attention in interventions and hence in evaluation. Given this interest and focus, this article has suggested a way using theories of reach that focuses on reach for interventions and their evaluations to make sure that equity issues for key reach groups are “on the table”, assessed and reacted to as interventions are implemented. Espinosa (2013: 180) argues

that in development evaluations, “little or no attention is paid to how the development action impacts on women and on men and how it changes unequal gender relations.” Developing and using theories of reach would go some way to addressing this gap.

References

- Espinosa, J. (2013). “Moving towards gender-sensitive evaluation? Practices and challenges in international-development evaluation.” *Evaluation* 19(2): 171–182.
- Mayne, J. (2011). *Contribution Analysis: Addressing Cause and Effect. Evaluating the Complex*. R. Schwartz, K. Forss and M. Marra, Eds, Transaction Publishers: 53–96.
- Mayne, J. (2012). “Contribution Analysis: Coming of Age?” *Evaluation* 18(3): 270–280.
- Mayne, J. (2014). *Using Theories of Reach to Enhance Equity Considerations in Evaluation. Speaking Justice to Power: Ethical and Methodological Challenges for Evaluators*. K. Forss and M. Marra, Eds, Transaction Publishers.

EVALUATING EQUITY IN A DEVELOPMENT CONTEXT

Eva Kløve

Inequities and inequalities¹ are central to development problems. Yet these concepts are sometimes poorly understood, and reduced to a narrow focus on poor or marginalised groups. Inequities and inequalities exist through unjust power relations and discriminatory practices, and can be found in all segments of society. This has some important implications. First, equity should not be approached in a narrow manner. Second, equity is as relevant to economic development as it is to social development.

I start with the latter point. Neoclassical economics has claimed a trade-off between equity and efficiency, which is due to the alleged dulling of incentives created by a more equal distribution of output. Thus, equality and economic growth cannot be simultaneously achieved.

Over the years, distinguished economists, including Nobel laureates Amartya Sen and Joseph Stiglitz, have challenged this view. The Nordic countries have exemplified what is counter-intuitive to neoclassical thinking; following what would be seen as a recipe for serious economic trouble (Moene 2006). They have too small wage differences, too high and too progressive taxes, too large public sectors, too generous welfare states, and too strong unions. Yet these countries have achieved high growth, low unemployment, low inequality, and a fairly efficient allocation of resources. How come?

One of the key messages from recent research on Nordic experiences is that there is not always a trade-off between equality and efficiency: Certain policies, institutions and behaviours fit together and strengthen each other. “Under some institutional arrangements, equality and prosperity can work together and be mutually beneficial” (Baland et al 2009).

To mention two examples² where equity and economic development go hand in hand: First, in the 1960s Norway experienced increased female labour force participation. This raised the demand for public care for children and the elderly, which again made it possible for

**«No society can surely
be flourishing and happy,
of which the far greater part
of the members are poor
and miserable».**

Adam Smith

more women to work. Second, wage coordination in the Norwegian labour market has brought wage compression, which again has led to political support for higher welfare spending. With higher welfare spending, effectively working as a minimum wage, comes higher wages for weak groups. This has reduced wage differences even further.

In these cases, the predictions of neoclassical economics have failed. Because much of the premise of development is based the same predictions, so has development.

So why is this relevant? Answering this question brings me to the first point raised above. In tackling inequity, the development community can sometimes be too preoccupied with how to best target the poor or the marginalised. However, such a narrow focus can lead to development programmes that are flawed and short-lived. In the Nordic experience, universalism has been an important principle. Universal social provisioning means that the entire population is the beneficiary as a basic right, in contrast to targeted provisioning which involves selecting a few “deserving” (Mkandawire 2005). In a human rights framework, citizens are entitled to universal human rights.

The arguments are equally relevant in a development context. In fact, universal systems can be the best – and most sustainable – way of reaching the poorest. Targeting is difficult, and costly. How do you define and identify the poorest? How do you measure their incomes or assets? How do you take into account that

the most vulnerable today, might not be the most vulnerable tomorrow? Negative economic shocks, accentuated by climate change, are frequent amongst the poor. Often they send families just above the poverty line into permanent poverty. Narrow targeting could easily exclude such families. What perverse incentives does targeting create? Is it desirable to keep beneficiaries below a given income-threshold, and how do you monitor that they do? Do you kick them out if they do not, knowing that this may well leave them worse off than before the programme started?

Universalism ensures objective targeting, if any, and is therefore likely to be less prone to corruption and perverse incentives. As found in a number of studies, universalism can also be key to ensuring long-term political support; a benefitting middle-class is more likely to be supportive³. This will be increasingly important as countries introduce private income taxes. Social insurance mechanisms such as social pensions or disability grants should of course be set at a level adapted to the national context. Affordability, however, is often a question of political will.

This understanding of equity as part of holistic social policies with implications for both social and economic development should provide the basis for how we think about evaluating equity.

Equity is not a bivariate factor that you either achieve or fail to achieve, and it can sometimes be hard to measure. This can make it complex to evaluate. It adds to the challenge that it can take time, sometimes years, to see the returns on social investments, so that costs can appear unreasonably high at the start of an intervention. Further, equity can be a politically sensitive issue, and the incentives may not be there or they may go in the opposite direction. There can also be deliberate discrimination of ethnic or minority groups, for instance by underfunding certain regions.

Equity is often seen more as a social issue than an economic one, with the consequence that evaluations of economic development do not

assess equity, and evaluations of equity lack assessments of economic aspects. Evaluators should aim to understand the broader context of the evaluation object, which is often quite narrow in scope. Recently, some efforts have been made to integrate in particular equality into evaluation frameworks⁴. However, although these efforts include important points such as assessing the evaluability of equity in a programme and using sufficiently disaggregated data, they fail to acknowledge the need for holistic thinking.

Seeing development programmes in their larger contexts involves identifying synergies or trade-offs between policies, institutions and behaviours. Too narrow approaches risk neglecting social spillovers. A current interest in including unintended consequences of development programmes is a valuable step

in such a direction. Another is the inclusion of systems-thinking in evaluations. This can be thought of as going beyond merely looking at a particular development programme, to study structures, patterns, norms and policies – and then identifying what can be controlled within the programme, and what cannot⁵.

To conclude; economic and social development cannot be seen as separate. Addressing and evaluating equity cannot be the sole province of either one or the other.

References

Baland, R., K. Moene and J. Robinson (2009). Governance and Development. Chapter 69 in Rodrik, D. and M. Rozensweig, eds., Handbook of Development Economics, Volume 5. Elsevier BV.

Kushner, S. (2014). Democratic Evaluation and Equity Focused Evaluation: Cousins but not siblings. Evaluation Connections, European Evaluation Society, November 2014.

Mkandawire, T. (2005). Targeting and Universalism in Poverty Reduction, Social Policy and Development Programme Paper Number 23, United National Research Institute for Social Development, December 2005.

Moene, K (2006). "Equality, Social Organization, and Performance (ESOP), Confronting theory with Nordic lessons". Downloaded from <http://www.sv.uio.no/esop/english/about/reports/plans-reports/The%20CoE%20Application>.

1. Equity and inequality are understood differently in different context, but it is common to interpret equity as "equality of opportunities", and "equality" as "equality of outcomes". UNICEF states that "Equity-focused policies do not aim to eliminate all differences so that everyone has the same level of income, health or education. They aim at eliminating unfair and avoidable circumstances that deprive individuals of their rights".
2. From Moene (2006).
3. For references, see Mkandawire (2005).
4. It makes sense to see equity through a human rights lens, as the key principles of a human rights based approach provide a useful entry-point to assessing equity. These principles are participation, accountability, non-discrimination, human dignity, transparency, empowerment and rule of law.
5. See for example <http://www.insites.org/evaluations-and-resources/systems-thinking-in-evaluation>.

THE AUTHORS

Maria Barrados

Maria Barrados recently retired from the Public Service of Canada. She was President of the Public Service Commission of Canada for 8 years, and prior to that served as Assistant Auditor General at the Office of the Auditor General of Canada. In her retirement from the federal public service she continues to advise governments at home and abroad on public service reform, governance, performance measurement and evaluation, financial and human resource management. Over the past two years she has been actively involved in a project supporting Civil Service Reform in Mongolia. Maria serves on a number of boards and advisory committees. She has maintained her research interests in audit, performance measurement, evaluation and statistical analysis and continues to write on these topics. In February 2014, she accepted a position as Executive in Residence at the Sprott School of Business at Carleton University, Ottawa, Canada. She has a Ph.D. in sociology from Carleton University, M.A. from McGill University and an Hons. B.A. from the University of Saskatchewan.



Kim Forss

Kim Forss holds a Ph.D. from the Stockholm School of Economics. He is the founder/owner and manager of Andante – tools for thinking AB, and in the company works with research, teaching, and consultancy assignments in evaluation. Over the past 25 years, he has evaluated policy and strategy in the field of development cooperation, public administration, the national system of Parliamentary inquiries, as well as policies for research and development. He has worked extensively with organisational analysis, for example of the Swedish Government Offices, the national museum structure, music exports, Swedish regional administration and of several UN agencies (governance systems in UNEP, reform processes at WHO and UNDP, organisational learning in UNICEF, and a comprehensive evaluation of UNESCO). His recent research publications include a recent volume on evaluating complex policy, and he contributes to a forthcoming volume on evaluation cultures, as well as evaluation and equity. He is a member of the European Evaluation Society since 1992 and a founding member of Swedish Evaluation Society, where he has served on the Board for six years, in 2010 and 2011 as the President.



Eva Kløve

is senior adviser in the evaluation department of the Norwegian Agency for Development Cooperation (Norad). She



holds a Ph.D. in economics from the Centre of the Study of Equality, Social Organisation and Performance at the Department of Economics, University of Oslo. Her education also includes a post-graduate diploma in African Studies from University of Cape Town. Eva has more than 14 years professional experience in development; both from Norad, the Norwegian MFA and the NGO sector. Her expertise is within the areas of social and economic policies and how they relate.

Mita Marra

Mita Marra teaches Public Policy and Public Administration at the University of Salerno, Italy. She also teaches in The World Bank's International Program in Development Evaluation Training (IPDET) at Carlton University (Canada). Her current research interests revolve around the evaluation of complex policies for human and social development. Dr. Marra studied Economics and the University of Naples in Italy and received a MA in International Relations from the School of Advanced International Studies (SAIS) of the Johns Hopkins University (1998) and the PhD in Public Policy from the Trachtenberg School of Public Policy and Public Administration of the George Washington University (2003). Since 1998, she has consulted with the World Bank, the United Nation Economic Commission for Europe, and a number of regional governments in Italy, conducting evaluation research on institutional development programs as well as gender equality and social policies. Dr. Marra has authored a number of articles and books. For Transactions Publishers, NJ, in 2011, she co-edited with Kim Forss and Robert Schwartz "Evaluating the Complex. Beyond Attribution and Contribution." Mita Marra is Associate Editor of the Journal Evaluation and Program Planning (July 2015) and the president elect of the Italian Evaluation Association (April 2013).



John Mayne

Dr. Mayne is an independent advisor on public sector performance. Over the last 11 years he has focused largely on international development evaluation and results based management work. He has been working with a number of government, NGOs and international organizations. He has authored numerous articles and reports, including several on contribution analysis, and co-edited eight books on program evaluation and performance monitoring. In 1989 and in 1995, he was awarded the Canadian Evaluation Society Award for Contribution to Evaluation in Canada. In 2006, he was made a Canadian Evaluation Society Fellow. Dr. Mayne's current research interests are on approaches for strengthening impact evaluation, useful theories of change and dealing with attribution.



Robert Picciotto

Robert ('Bob') Picciotto, (UK) Professor, Kings College (London) was Director General of the World Bank's Independent Evaluation Group from 1992 to 2002. He previously served as Vice President, Corporate Planning and Budgeting and Director, Projects in three of the World Bank's Regions. He currently sits on the United Kingdom Evaluation Society Council and the European Evaluation Society's board. He serves as senior evaluation adviser to the International Fund for Agricultural Development and the Global Environment Fund. He is also a member of the International Advisory Committee on Development Impact which reports to the Secretary of State for International Development of the United Kingdom.



Claudine Voyadzis

has spent some 20 years working in the evaluation field. Her work consisted mainly of appraisal and monitoring of projects and programmes, impact studies and ex post evaluations in West and East Europe and in Africa. She earned her Ph.D. in the United States in Demography and Sociology and has a dual experience of Anglo-Saxon and European approaches to evaluation. After working in the Independent Evaluation Group of the World Bank in Washington, she joined the Council of Europe Development Bank in Paris as technical adviser for project appraisal and monitoring. Subsequently she was appointed to set up the Ex Post Evaluation Department and remained for six years as Director of the Department. Claudine Voyadzis works now as an independent consultant, and leads evaluations for multilateral organisations and consulting firms.



Peter Wilkins

Peter is an Adjunct Professor at The John Curtin Institute of Public Policy at Curtin University and an Honorary Research Fellow in the Sir Walter Murdoch School of Public Policy and International Affairs at Murdoch University in Western Australia (WA). He holds a PhD from Southampton University England and conducts research into evaluation; performance measurement, reporting and improvement; accountability; collaboration and governance. He has served as Western Australia's Deputy Ombudsman and prior to this had been Western Australia's Assistant Auditor General Performance Review. He is a member of the International Evaluation Research Group, the Australasian Evaluation Society, the Australasian Study of Parliament Group, and is a National and WA Fellow of the Institute of Public Administration Australia and Deputy Chair of the Board of Australian Volunteers International.

