

International Summer School on Public Policy Evaluation Research, Lille, France
29th August to September 2nd 2010

Some personal impressions and thoughts

By

Marlène Läubli Loud, DPhil
Email: marlene.laeubli@bluewin.ch

This year, an International Summer School was organised for the first time by the French Evaluation Society's Network of Researchers on Policy and Programme in partnership with the French Ministry of High Education and Scientific Research and the Lille Institute of Political Sciences. Some 100 participants came together to debate developments in evaluation research as well as to present and/or learn about ongoing evaluative research projects. Of these, 51 of the participants were research fellows, 27 were currently PhD students and the remaining 22 were in activities "other than research". 90% of those attending came from Europe, the majority coming from France. But there were also participants, other than guest speakers, from Africa and North America. The main intention was to develop an international network of researchers in and about evaluation, produce a research agenda for the coming years and ultimately build the foundations for legitimising evaluation as an academic research field in its own right.

The topics covered during the five days were: the development over time of evaluative research and research on/about evaluation, methods and issues for researching the theory and practice of evaluation, evaluation research by sector, by territory, evaluation vis-à-vis academic disciplines, and academic programmes for evaluation research. Full details of the programme and presentations can be found at

http://www.sfeasso.fr/blog.php?menu_id=264&menu=593&mode=page

My intention here therefore, is to not to provide a synopsis of the week's debate, but to share my thoughts and impressions about the topics that I found to be of most interest – namely the historical development of evaluation research, more about what is and/or could be evaluation research as a way of improving its academic credibility, the interdisciplinary nature of evaluation, as well as a possible research agenda.

In reflecting on the historical development and epistemological roots of evaluation research, Evert Vedungⁱ uses the metaphor of "waves" rather than "generations"ⁱⁱ or "trees and branches"ⁱⁱⁱ to describe the origins and developments of evaluation theories and practices over time. We now, therefore, have essentially three distinct metaphors to illustrate how theories come and go. Each accepts that there is an integration of "old" ideas with the "new", but that with each development, there is a shift in focus, or even sometimes, a combination of new and old resulting in a "substantially new way of thinking". For example, an analysis of contextual influences as an integral part of evaluating effectiveness has been given higher or lower priority over time. But happily more recently with "realistic evaluation", for example, Pawson and Tilley^{iv} have resurrected and further developed the importance of "context" in making judgements about what works.

But the novelty, to some degree, of Vedung's analysis is the clarity with which he links the emergence of each "wave" with a particular political doctrine, and how this in turn has acted as a driving force to bring about change. His first wave, the "science-driven wave", is coupled with a search for "radical rationalism" and is focused on the verifiable, objective measurement of goal achievement. It is expected to provide scientific evidence on the means being developed to achieve the end goals, and is therefore focussed on "experimentation" before launching a full-scaled intervention. The second is the "dialogue driven wave", post 1968; it is pluralistic and participatory accepting that there are multiple realities, values and perceptions. It is based on the principle of "communicative rationality" not, as before, on a "means-end rationality". With the introduction of "new public management" in the 1980s, a third "wave" emerges in response, the "neo-liberal wave", one that is particularly based on the concepts and language of the market economy. It is focused on results, value for money, and customer satisfaction. Determining the effectiveness of "results-based management" is a central preoccupation for evaluation, especially taking into consideration the "customer's point of view. The latest and fourth wave he describes as the "evidence-based wave", which is a return to experimentation and very reminiscent of the first, "science-driven wave". In fact, at first glance, there appears to be very little difference between the first and fourth wave and the supposed differences are not sufficiently developed. But you can read more on this in Vedung's article which recently appeared in "Evaluation, The International Journal of Theory, Research and Practice", vol. 16 No. 3, July 2010 provides a detailed discussion of these ideas. For me, the analysis is particularly relevant for teaching purposes and for understanding why for many of us who have been working in evaluation for some time, there is a certain "déjà vu" when it comes to learning about supposedly "new trends and fashions" in evaluation theory.

Which brings me to the next issue, that of academia's (non) recognition of evaluation. Given that one of the major objectives of this Summer School was to debate, once again, how to improve the recognition of evaluation as an academic field of study, it is not surprising that a day was devoted to this subject. One of the major problems raised, of course, is about trying to secure research funding for a "discipline" that is considered to be more of a profession than an academic field of study. There was also discussion about where to place scientific articles about evaluation when evaluation journals per se were considered relatively low in the ranking order of scientific publications. Better to have the article placed in reputed journals of one's discipline such as sociology, social psychology etc, than in evaluation journals. The domination of a particular discipline in any one country and, as such, the influence on evaluation practice was also raised, but not further developed. For example, in France, economists seem to have been more active than others in the field of policy evaluation, at least during the first couple of decades, whereas in the United States, educationalists were the first to enter the field and have been ever present, and in large numbers ever since. More about the relationship between the disciplinary shaping of evaluation's development in a particular country, and a comparative analysis with developments in several others would be an interesting research project in itself. But the most important point for me that was raised during this discussion was the need to distinguish between conducting and presenting an evaluative research as different from conducting research on and/or about evaluation. Steve Jacob's^v research on the institutionalisation of policy evaluation in

several countries is a good example of “research about “a certain aspect of evaluation. For Patricia Rogers, there are many aspects needing research for example, she says “too many evaluation approaches are advocated without evidence of their effectiveness or appropriateness”. And as for managing or commissioning evaluations, an area very neglected in research, she suggests such questions as “which approaches to commissioning, designing, conducting, reporting, managing evaluation are effective, and in which kind of situations?” So the answer seems to be that whilst one shouldn’t give up on research *in evaluation* (evaluative research), there are clearly advantages in **conducting research on and about evaluation** within the currently acknowledged academic disciplines, especially as a means of promoting more academic interest in the subject.

Trying to place evaluation within a particular academic discipline, however, can run the risk, of course, of undervaluing the interdisciplinary nature of evaluation. Nicoletta Stamme gave a very clear and concise input to the session by illustrating the influence and contribution of sociology to evaluation. She especially reminded us of the need to return to the literature on sociology for theoretical frameworks when analysing data. Alain Trannoy highlighted many of the problems facing economists in judging the effectiveness of policies when using only economic evaluation models. Rogers stressed the need to look at the epistemological roots and influences and particularly focused on the dangers of poor evaluations. In the accepted hierarchy of evidence, for example, credible evidence can be excluded leading to faulty conclusions about what works. In this respect, she reminded us of the American Evaluation Association’s reply to the US government’s proposed criteria for effectiveness evidence – based on randomised control trials - with regard to home visiting program models:

« I. Randomized designs are not necessarily superior to Quasi-Experimental Designs in all circumstances.

II. Many other factors, especially real world conditions and fidelity of implementation, affect the quality of scientific evaluation evidence

III. Other evaluation designs can support causal attribution, especially when they can rule out other potential causal factors.

IV. Knowledge of program impact is enhanced by considering multiple studies and using mixed methods»^{vi}

These more theoretical presentations were then very well supported through the practical case studies presented by the researchers. They provided good examples of the need for mixed methods, but particularly for an interdisciplinary approach when tackling their research projects.

Finally, what next? Most importantly for me in thinking about research in relation to evaluation was again the distinction made between the different roles that research could play (a) research **ON/ABOUT** evaluation - when studying evaluation **as an object of research** e.g. social sciences funded research on the object ‘evaluation’, (b) research **IN** evaluation - when used to draw out concepts from evaluative research that could then be applied to social science research and finally (c) research **FOR** evaluation – research aimed at identifying how we can improve evaluation theory and practice. As for a research agenda, as there was little focussed attention on the matter, it was not surprising that this particular objective was not well achieved. More importantly perhaps, however, was agreement that the Summer

School had played an important role in bringing researchers together and that this should be repeated, possibly every two years. Meantime, it was felt that the now more international research network, as established through the Summer School, should remain active albeit with no concrete ideas about how at this point in time.

If you wish to join the network or find out more about it, please do get in contact via

<http://www.sfe-asso.fr/>

References:

-
- ⁱ Vedung, Evert (2000), *Public Policy and Program Evaluation* Transaction publications, USA
 - ⁱⁱ Guba E.G. & Lincoln, Y.S. (1989) *Fourth Generation Evaluation*, Sage Publications USA
 - ⁱⁱⁱ Alkin, M. (ed) (2004) *Evaluation Roots, Tracing theorists' roots and influences*, Sage Publications USA
 - ^{iv} Pawson, R. & Tilley, N. (1997) *Realistic Evaluation*, Sage Publications UK
 - ^v Jacobs, S. (2005) *Institutionnaliser l'évaluation des politiques publiques. Étude comparée des dispositifs en Belgique, en France, en Suisse et aux Pays-Bas*, P.I.E. Peter Lang SA, Brussels
 - ^{vi} American Evaluation Association (2010) on *Studies and Using Mixed Methods* see www.eval.org/EPTF/aea10.eptf.visitation.comments.box.pdf