Comments on "An Association to Improve Evaluation of Development Aid"

Hellmut Eggers

1. General Observations

1.1. This is a remarkable text treating, with solid know-how and a vivid sense for the essentials, a highly important issue in the current development assistance debate. So, if the author would opt for not heeding any of my critical observations, that would be his/her right. The contribution could be accepted by JMDE as it stands, and my critical review would then pertain to the published version, say under "Comments" or, still better, under "Creative Controversy". I do believe, however, that this contribution still could (and should) be improved along the lines suggested below and might then (hopefully) mark the beginning of a series of successful measures designed to improve the effectiveness of development assistance. So, why is this text really worthwhile publishing, studying and to be acted upon?

1.2. Having painted a shockingly dismal picture of the state of evaluation in development cooperation (is it really THAT bad? I hope not but don't dare maintaining the contrary!!!), the author concentrates on two aspects to be tackled with a view to dramatically improving the present situation: S/he proposes:

(a) a new and practical way of coming to grips with the prevailing haphazard structures of evaluations: focusing on two key aspects that are of paramount significance in *any* evaluation: (i) cost-effectiveness and (ii) impact. If this is done, then these two elements would act like a piece of yeast permeating the entire evaluation dough; and

(b) the establishment of a new structure s/he calls "Evaluation International" (EI) to support the "peaceful revolution" s/he has in mind of radically improving the usefulness of evaluation in development cooperation.

I fully share his/her view that this is the two-pronged attack on the prevailing mediocrity that appears to be the most promising.

1.3. I do not believe, however, that the choice of the two key aspects as mentioned under 1.2 (a) above is optimal (in fact I think it's only "second-best"); and I think that the brand of EI s/he proposes under 1.2. (b) is a non-starter and has to be replaced by an organizational set-up that is radically different from what s/he has in mind. I will argue these two points in detail below.

1.4. Please note that these critical remarks in no way impinge on the decidedly positive judgment I have formed about the approach the author has taken. In fact, I believe that this approach is the only one that has a chance to succeed in practical, down-to-earth reality.

2. Specific Observations

2.1. It is true that the structure of development assistance management as described favours a "positive" bias concerning the conclusions evaluations will arrive at. My own experience as the head of the evaluation division of the EU Commission Directorates General for Development and for External Affairs shows, however, (at least I hope!), that such pressures can be resisted. I told evaluation contractors routinely that their chances of getting re-employed by us would diminish dramatically if they gave me the impression that they wrote what they thought my colleagues, my bosses, our "clients" and/or I wanted to hear. I made perfectly sure that they understood that message. We got quite some "unpleasant" evaluation reports as a result...

2.2. It is not correct to say that so far there has been no effort to provide "a means to achieve consistency between evaluations," and I would be really surprised if it were true to maintain that "the proportions of evaluations that estimate cost-effectiveness is probably declining". I hate to preach "pro domo", yet is it a verifiable fact that when I was responsible for the work of the Evaluation Division, we started, in September 1987, to develop what my boss, the Deputy Director General then called "The Integrated Approach", that I then re-baptized "The Integrated Approach to Project Cycle Management" of which, after a while, only the title "Project Cycle Management" (PCM) remained. As such, it has been taken up by development organizations across the globe and is still in official use today (even though not quite in a way I find entirely satisfactory; but having retired in 1993 from active service, I can no longer exercise any influence on current events). PCM offers, however, a complete mechanism not only to "achieve consistency between evaluations" but to achieve such consistency across the entire project/program/policy life cycle, including planning, implementation, monitoring AND evaluation. I include an up-dated version of an article I published on PCM in the professional journal *Evaluation*, Vol. 8(4), in 2002, as well as another one on "Planning and Evaluation," published in JMDE, June 2007. *The present "Comments" cannot be entirely understood without a thorough review of these two articles.*

2.3. Throughout his/her article, the author insists on the need to include two precise headings in any evaluation study, to wit "Cost-effectiveness" and "Impact". Of course, this is necessary (even though for "impact" it might prove difficult, cf. point 2.4. below). What is surprising, however, is the fact that the author insists on these two headings only. They are not at all sufficient to arrive at consistency between evaluations, let alone between evaluations, planning, implementation and monitoring (cf. also the "Second Main PCM Practice). The "Basic Format" of PCM (cf. annexes to my articles) contains ALL of the most important headings suggested by experience, and *this format is, of course,*

open to further improvement. It does, however, contain "inter alia" (and has done so from the beginning in 1987), the necessary headings to assess "cost-effectiveness" as well as "impact."

2.4. Indeed, and as the "Basic Format" specifies, the classical chapters of "economy" (buy quality inputs to be acquired at lowest cost—tendering); "efficiency" (transform inputs into outputs at the lowest cost: do "things right"!); "effectiveness" (do the "right things" with a view to achieving the project/program purpose or specific objective, *always* to be expressed in terms of *"sustainable benefits for the target group"*); and "impact" (make sure the specific objective, once reached, will contribute to the wider objectives of the development policy pursued). These are the classical terms that have been known and applied for decades already in "Logical Framework Analysis" as taken over by PCM. The article is much less precise in that regard and one gets the impression that "impact", in the author's thinking, is hovering somewhere between the classical "project purpose" (or "specific objective") and the classical notion of impact as the project's contribution to the general development objectives. The entire article, by the way, is rather "lose" as far as terminological rigour is concerned.

2.5. The author, however, certainly has a point when s/he insists on the need for prioritizing certain key issues, as that appears necessary in order to focus the minds of the evaluation community. They have a general tendency to re-invent if not the wheel, at least the issues to be evaluated and the structure of evaluation reports, over and over again. If it is really true, as the author suggests, that the "proportion of evaluations that estimate cost-effectiveness is declining," then evaluation would be in a sorry state! If, indeed, this key aspect were being increasingly neglected, then such focusing of minds on certain key issues would be all the more absolutely indispensable. However, the focus on cost-effectiveness and impact, as suggested by the author, seems to me less than optimal. In fact, to me it appears to be inadequate. "Cost-effectiveness" has to be considered at the beginning of the project/program cycle while "impact" is to be considered beyond its end (although the author seems to mean "outcome," not "impact" as understood in generally accepted evaluation terminology). None of the two focuses on the central issue which is the project purpose or "specific objective." Since the project purpose is *always* geared towards the creation of sustainable benefits for the target group, it is entirely sufficient as the centre of gravity of any project/program. If, indeed, sustainable benefits for the target group are reached, then this means that the rest of the other key aspects, that is, "economy," "efficiency," and "effectiveness" (and thus "cost-effectiveness") have also been realized. Politicians, not the project/program management, will then be responsible for ascertaining (but that will usually be the case) that these sustainable benefits contribute to the realization of the wider project/program objective (or "goal"), that is to positive "impact" (cf. first main PCM Practice as presented in my articles). "Cost-effectiveness" can be reached, without any difficulties, for nonsense projects that serve no real need, and impact can be negative, so these two don't guarantee any positive project/program

outcome at all! Only the project/program purpose, if the said sustainable benefits are realized, does.

2.6. There is, however, one aspect that is neither covered by the "cost-effectiveness/impact" "twin brothers" nor by the focus on the creation of sustainable benefits for the target group, that is, by the project purpose, and that is "relevance": Solving the problems the target group faces or seizing the opportunities that present themselves to it must *really be important for them,* if not, there's no point in making the required effort. This is why PCM includes "relevance" among the key issues to be faced (cf. Basic Format). The author's "twin brothers" don't.

2.7. The author suggests the establishment of an association s/he calls "Evaluation International (EI)," that would control the entry to the evaluation profession and set and enforce appropriate standards, with a view to serving the need for learning and accountability in the international development community, that is, the need, through solid evaluations, for conceiving, executing and evaluating successful projects/programs. The author underlines that cost-effectiveness should orient the methodology to be promoted by EI (I have already explained why I don't share this view, while not at all wanting to diminish the importance of cost-effectiveness, cf. above). All of the more detailed responsibilities and tasks of EI are already covered without exception, by PCM, and would therefore not have to be re-invented by EI. To be sure, PCM should remain *always open to further improvements* (and Scriven's *Key Evaluation Checklist* may certainly play a positive role in that context, although it is not directly geared, in its present form, to immediate

practical utilization in the area of international development cooperation and would have to be adapted accordingly). The "Guidebook" proposed by the author seems a good idea and should take over PCM entirely (no re-inventing of the wheel, please!), while also addressing those issues the author mentions that are not covered by PCM. Note that PCM covers, that is, the approach to sector-wide learning. Apart from sectors, PCM covers three other key aspects the article doesn't mention: development instruments, cross-cutting issues and country/regional analyses, and these should also be taken up in the guidebook. Furthermore, the dangers of the dilemmas mentioned in the article are certainly overestimated if PCM with its principles, instruments and practices is applied, as it is based on solid experience by countless experts over decades. May I repeat, by the way, that the strong focus on impacts and cost effectiveness throughout the article is, in my opinion, not only misleading but downright dangerous as it places the stress on aspects secondary to the creation of sustainable benefits for target groups, that is, on something different from the very reason of the existence of development cooperation. As to the proposed archive, the data base according to PCM affords, I think, a much better structured tool than the one proposed in the article, making experience, gathered around the world, directly available to sector, instrument, cross-cutting issues and country/region specialists. But this judgment may be (partly) mistaken, and so a common debate might ultimately lead to best solution results in the end. At any rate, I should like to advise strongly against starting to write the "Guidebook"

concentrating on a few sectors only. This suggests that experience does not allow tackling all of the above mentioned four key aspects at once, including sectors, *but it does*! Moreover, the approach suggested by the author to start with a few sectors and then proceed further on the basis of the experience thus becoming available, piece by piece, would lead to a long drawn-out process that might squash all enthusiasm of participants and might end in general boredom and hopelessness. The knowledge base *already available today* is largely sufficient to proceed much faster and to make sure the initial momentum is not lost!

2.8. As to the nature of EI, it strikes me as way too bureaucratic. Mobilizing donor contributions, providing for standards, selecting evaluators worthy of an EI stamp, defining what an evaluator must and mustn't know, providing for and enforcing stringent entrance conditions for would-be evaluators trying to make their way into the VIP lounge that would be EI for the chosen few, expel those that, alas, gave the appearance of being white but in the end turned out to be black sheep, authorizing (encouraging?) evaluators to identify the honest folk among their clients with a view to exposing the fraudsters, etc., etc. would make for endless squabbles, delays, rivalries, frictions and frustrations...No, scientific authority in this field cannot be "allotted" to a newly established entity! EI would have to make a name for itself over many, many years, and only then demand for its services may rise, and that would be another reason for creating endless delays that would frustrate the entire undertaking in the end and that could so easily be avoided! Furthermore, "evaluation of a defined part of donors'

operations to be carried out by EI members" would be legally inadmissible, at least within the European Union, as tendering for evaluation services is compulsory. And if, when all is said and done, this new set-up would lead, "after a substantial initial investment", to a modest rise in routine evaluation costs of "perhaps 50%," then you may safely bet your bottom dollar that the whole idea of establishing the EI will turn out to be a dead duck from the word "go"!

2.9. This does not mean that the functions the author proposes to entrust to EI would not be worthy of consideration, many of them to be found valid and well worth concerted action. But there should not be another bureaucracy! All of the functions considered are already well represented in existing organizations, and these should be encouraged to participate in a common effort. Why, for instance, should there be a new "Standards Committee"? Existing evaluation societies, like the European Evaluation Society, the German "Degeval," the US Evaluation Association, the Paris based OECD and many others have worked them out long ago! They should be compared, maybe merged into a common text acceptable to all, but certainly not "re-invented" by EI's "Standard Committee"! There is no need, either, to work out a guide-book "ex-nihilo." There are basic texts available covering competently and completely the areas under consideration. It will be sufficient to compare all of these different documents and merge them into common texts to be acceptable to and applied by all.

2.10. The author, in spite of all these considerations, insists on the idea of the need for establishing what I would call an "IE King-size": requiring (i) start-up

and running cost-coverage by private foundations, the UN, etc.; (ii) clients to use (compulsorily!) its services, superior to any other services, stands to reason, and therefore, already noted, a puny 50% more expensive only (and all that before IE would have had a chance to convince clients that its services are better than those of any competitor!!); empowered to (iii) "control entry to the (evaluation) profession; and to (iv) "set and enforce the standards" to be respected and applied by evaluators. Heaven forbid, dear author! We don't need that kind of Big Brother (BB) IE making and breaking careers and knowing it all, and better, than anybody else! So, you are quite right not to ask for "details," something that would be absolutely essential if your proposal had any chance of getting implemented as proposed. Take it from an old hand, a really "seasoned roadster" on the bumpy road to success in the area of development cooperation: It doesn't, I'm afraid. But rest assured, I fully share your basic concerns and the conviction that something has to be done about the desolate state of development cooperation evaluation. So, what I would propose instead of the BB or King-size variety of IE, would be an "IE Light." Let's see what that might look like:

2.11. In order to avoid any misunderstanding: I fully share the basic concerns underlined by the author and his/her conviction that the improvements proposed do require organizational support, for instance by an "IE Light". I think, it would be possible, for instance, to establish a "Network Development Cooperation Evaluation" (NDCE) among interested university departments around the world specializing in Evaluation and/or Development Cooperation, located in developed as well as in developing countries. One could start with, say, one university in Africa, one in Asia, one in Europe, one in Latin America and one in the USA (please note the alphabetical sequence). These would first have a thorough look, world-wide, at existing structures, knowledge and organizations active/interested in the area of development cooperation and its evaluation. NDCE would make absolutely sure that there would not be any "doubling" of needed knowledge or organizational resources that are already available among participants or among development cooperation agencies, NGOs, the Paris based "Development Assistance Committee" (DAC), professional consultants' associations, etc., etc... Only then would NDCE start to work out a proposal for a common action plan, to be proposed to existing organizations willing to take part in a common, concerted effort among equals. NDCE would make sure to include, from the start, among their partners representatives of intended beneficiaries (target groups) of development cooperation as well. I fully agree with the author considering that the evaluator should strive, above all, to give them a "voice" to be heard among (and above) the chatter and clatter of the powerful and the "professionals".

2.12. And how about the details to get the ball rolling? You believe, dear author, they are "less important" (cf. your "Conclusions")??? Let me assure you, they are not less but ALL important! Without them, people may nod their heads in agreement, and then nothing will get done, guaranteed! (cf. also the "PCM Third Main Practice"!) So, what about those details? Let's see:

2.12.1. Identify a handful of people who are *really determined* to take up the challenge of launching the EI in the form of NDCE. They might belong to one single university only, for instance Michigan, plus a couple of committed individuals, perhaps, that have decided to live their dreams instead of dreaming their lives.

2.12.2. Have them work out a draft strategy paper containing the basic intentions pursued both as to the substance, that is, NDCE's mission, and on start-up procedures, i. e. on how to identify universities and other organizations to be approached, on modalities of contact, timing, etc., etc...This work should not require any extra external financing.

2.12.3. Send the draft strategy paper to those organizations and ask for reactions.

2.12.4. After having received feed-back, concentrate on promising partners only and review the strategy paper accordingly, having it agreed by all those willing to cooperate.

2.12.5. Then, work out a first action plan, distributing tasks and responsibilities among participants. Agree on coordinating needs and central secretariat and discuss finance *for the first time.*

2.12.6. Take matters from there, as there is no need and no point in going any further right now with what would be nothing but speculation for the moment. Remember, however, that the procedure thus launched would have to be further pursued, in due course, along lines to be traced according to the "Third Main Practice PCM". The "Project" in

question might be called: "Improvement, through NDCE, of the Use and Usefulness of Evaluating Projects, Programs and Policies in International Development Cooperation."

3. Please, dear author, consider my critical remarks (including the post scriptum...) to be meant in a wholly constructive way. What would they be worth if they were not honest?! But they are not harsh, rest assured. Let us get together over your text an mine. If ever something positive comes out of this initiative, it will be you to have launched the idea!

Post Scriptum on Priest(esse)s

Like the author, I am unreservedly in favour of opening access to the priesthood (evaluation) to women (female evaluators) as well. Implying, however, as does the author, by a subtle (and exclusive) use of the pronouns "she" and "her", that supplanting priests (male evaluators) by priestesses (female evaluators) should be exaggerated to the point of driving the former to total extinction, seems to me grossly overplaying the hand of political correctness in flagrant violation of the principle of equality among the sexes. I propose, therefore, that the feminine expression "she" should be systematically replaced by the hybrid expression "s/he," and that the feminine expression "her" should be systematically replaced by the hybrid expression "him/her." Please note that the hybrid "s/he" starts with the feminine "s," followed by the masculine "h," while the hybrid "him/her" starts with the masculine "him," followed by the feminine "her." My proposal thus corresponds rigorously to a perfect standard of equality among the sexes and thus to unbiased gender equilibrium.