Bull. Séanc. Acad. R. Sci. Outre-Mer
Meded. Zitt. K. Acad. Overzeese Wet.
55 (2009, Suppl.): 000-000

Evaluation of Development Research: Difficulties and Questions

by

Ivan BEGHIN*

KEYWORDS. — Development; Development research; Evaluation.

SUMMARY. — There is a general consensus that evaluating development research is not an easy task. In addition to the difficulties of evaluating scientific research in general, evaluation of development research suffers from specific limitations, such as the need to assess relevance to development, and the disagreements about the definition of development research. This paper attempts to offer an overview of the problems and their causes.

A broad array of persons need evaluation of development research: (i) the researchers themselves, (ii) persons who award degrees or prizes, or who appoint or promote researchers, (iii) funding agencies, (iv) development organisations, governmental or NGOs, etc.. Evaluation may be used for strengthening institutions, for accountability towards sponsors and/or public opinion, for disseminating knowledge, etc. There is thus a wide variety of situations, stakeholders, objectives to be met, etc. The main point here is that such a diversity implies flexibility in the design and choice of the most appropriate evaluation tools and criteria - in each situation and for each purpose.

Among the conditions for a "good" (meaningful) evaluation one should include clear objectives, comprehensiveness, due consideration of the specificity of each discipline, transparency of criteria and procedures, independence of the evaluator, etc.. Experience indeed shows that too many evaluations do not meet all these conditions.

To the main dimensions of evaluation of development research are scientific quality and relevance. Other dimensions include valorisation, performance, innovation, etc.

One comes to the conclusion that guidelines are required to assist persons in charge of evaluation of development research. Such guidelines should not be a handbook or a manual, and should leave way for adaptations to fit local needs and conditions. They should be short and concise. The Academy, hopefully in collaboration with other interested agencies, may endeavour to write such guidelines and widely disseminate them.

1. Introduction

— The purpose of this paper is to present an overview of the problems of evaluating development research and of the causes of such problem. A few think paths for finding answers to such problems will be followed. From the contributions of the guest speakers and from the closing Round Table solutions for these problems are the expected to emerge.
— The starting point of this paper is based on a basic assumption and on three observations, that most participants certainly share:
— The **basic assumption** is that our major interest is to help researchers in the South to do better research and receive due recognition for it.

* Honorary Professor, Institute of Tropical Medicine, Antwerp; Member, Royal Academy of Overseas Sciences; rue de Bordeau 53, 1060 Brussels (Belgium).

— The **three observations** are:
- our present tools for evaluating scientific research in general are unsatisfactory;
- evaluation of **development** research meets additional difficulties;
- researchers in the South, in that respect, have a serious disadvantage in comparison with researchers in the North.

1.1. Inadequacy of our Tools for Evaluating Scientific Research in General

— This is an important problem and the subject of much debate in the scientific press, in academic circles, and among funding agencies. The point is not discussed here since it is considered in the following papers.

1.2. When Development Research is concerned Evaluation becomes even More Difficult

— There are essentially two reasons. On the first place, in addition to evaluating scientific quality, one needs to evaluate also the relevance to development. On a second place, the disagreements about the definition of development research troubles the discussion. Even within the Academy there is no consensus on the best way to define the central theme of this meeting. The French have in Montpellier the IRD (*Institut de Recherche pour le Développement)* and they widely use the term «recherche au service du développement». The English have in Brighton the IDS (*Institute of Development Studies*). A term commonly used by our colleagues in the Flemish Community is "ontwikkelingsrelevant-onderzoek". The *Belgian Development Cooperation Price* uses a more elaborate, yet much better focused concept: that of "research that contributes strongly to knowledge that could benefit development in the South".
— Such definitions, it should be pointed out, do not necessarily include – but may include - scientific research conducted either in the South or in the North, but which does not have development as objective, at least on the short and medium.

1.3. Researchers in the South are at a Disadvantage

— Researchers in the South are at a disadvantage with respect to researchers in the North. I have covered this point in a presentation to the Academy on the occasion of its 75th anniversary (Beghin 2004).
— Their problems are basically:
- less access to scientific publications, congresses, networks, etc.;
- reduced independence in choosing research topics;
- greater difficulties in getting articles accepted in "mainstream journals";
- less recognition, and therefore less access to funding sources;
- isolation.

2. Who needs to Evaluate Research? What for?

2.1. Who needs?

— People who have a stake, an interest in good development research and/or need to make decisions about research or about researchers in this field Such "stakeholders" can be:
  - the researchers themselves, and their superiors;
  - universities and institutions delivering degrees, and appointing or promoting researchers;
  - organisations and foundations funding research;
  - editors and reviewers of scientific journals;
  - members of juries of prizes or awards;
  - international courses dealing with development;
  - potential users of the results of such research: governments, NGO's, bilateral and international agencies, etc.;
  - public opinion, the society at large.

2.2. WHAT FOR? THE OBJECTIVES OF EVALUATION

— The "stakeholders" can thus be very various people with diverse motivations. The objectives of evaluation of development research can equally be diverse, and a clear and explicit choice has to be made for each situation. The most frequent of these objectives are:
  – checking the relevance for development of a given research;
  – assessing and contributing to maintain or to increase the scientific quality of research, often including its valorisation;
  – allocating resources: providing funding or recruiting researchers;
  – strengthening of institutions;
  – insuring accountability towards funding sources, authorities, but also to the public;
  – evaluating researchers (individuals or research teams);
  – disseminating knowledge about development and its problems.

3. The Conditions for a Good Evaluation of Development Research

— The conditions under which a meaningful and satisfactory evaluation can be conducted are the same as in any other evaluation of research – yet with different emphasis on certain aspects:
  – the objectives of the evaluation must be clear. Why and what for is this evaluation conducted? For whom, and who will be using the results? These objectives should be explicit from the beginning;
  – comprehensiveness: all relevant elements of the context need to be identified and when necessary taken into consideration: place, opportunities, resources, etc. Also, evaluation should be taken as a whole, and not just as a sum of criteria;
  – separate disciplines should be approached differently: for example social sciences vs. natural and medical sciences vs. technical sciences. Objectives and tools of evaluation may be very different;
  – evaluation should be competent: criteria, methods and procedures should be well established, clearly spelled out, and based on generally accepted guidelines;
  – transparency of criteria and procedures should be made explicit from the start;
  – independence.
— These remarks may sound redundant. Yet, too many evaluation exercises and methods do not respect, in practice, all these rather obvious and simple recommendations.

4. What do we evaluate?
The Various Dimensions of Evaluation of Development Research

— We must make a distinction between the **dimensions** of evaluation and the **criteria** used to express such dimensions. Criteria vary widely according to the circumstances and have to be selected on the basis of each situation and need. In addition a certain number of criteria can be used to reflect different dimensions.
— Four rather classical dimensions are considered here (but there can be more):
  – scientific quality;
  – relevance for development;
  – valorisation;
  – performance.

4.1. Scientific Quality

This will be obviously in the first place. Criteria and procedures are the same as in the evaluation of research in general: bibliometric indices, peer reviews, panels, etc. Yet how to use such tools in the case of development research is one of the central issues of this meeting.

4.2. Relevance of the Research

Assessing the relevance to development of a given research project is not easy. Basically a research is relevant to the extent it answers – or has the potential to answer – a development problem. Yet a broad number of criteria can be used. Their identification and selection will depend upon the concept of development one has in mind, and on quite a few other considerations. As this topic is discussed in other papers of this volume, it is not considered here. Yet we would like to allude briefly to a connected question: who assesses relevance ?

Isabelle Stengers, in a recent book (2009) points to the potential dangers of research sponsored by private companies – that might have a particular stake in it (such as for example the pharmaceutical industry or agribusiness), or sponsored by certain state agencies that might have ideological biases (such as, for example, anti-environmentalists). Who is in charge of evaluation may therefore be an issue, particularly in the case of natural and technological sciences.

4.3. Valorisation

— It is maybe a less fundamental dimension than scientific quality or relevance for development, but it needs to be considered seriously. Actually, the impact of development research is seldom measurable, at least on short or medium term. Efforts towards dissemination of results and towards their effective application are therefore to be taken into account in evaluation. Valorisation in our case thus becomes an important, although too often neglected, dimension.
— Aspects to be considered in assessing valorisation may include **outputs** such as:
  – publication of scientific articles in peer reviewed journals;
  – other publications, including chapters in books;

- presentations in congresses;
- production of guides, handbooks, manuals, educational material, etc.;
- more generally the dissemination of results to all stakeholders in development
- advocacy;
- also: in special cases proposing solutions – or providing suggestions pointing to possible solutions.

The dissemination of the research results to all stakeholders, in a given situation, is a whole issue in itself. On the one hand researchers in the South tend to publish comparatively little, and on the other hand there is a lack of good indicators for measuring the exchanges of scientific knowledge.

Valorisation can also be expressed in terms of **outcomes,** such as:

— effective use of research results in development programmes or in policy formulation: both their actual application and the efforts towards application. In the particular case of technical sciences: patents, start-ups, etc.;
— strengthening of research capacity in the South;
— increased autonomy of researchers and research teams;
— degree of success in attracting resources, financial and human.

4.4. PERFORMANCE

— performance is essentially expressed as a function of productivity and quality;
— it is a dimension of evaluation that comes well behind scientific quality and relevance;
— it is used less in evaluation of research "per se" than in the evaluation of researchers (individuals; research teams);
— it can also be used in the ranking of departments within a given discipline.

Criteria used to express valorisation and to assess performance may widely overlap. Usage will determine the interpretation of such criteria.

This leads to opening a parenthesis concerning the differences between evaluating individual researchers and research teams: here also criteria will be partly different. When a team is being evaluated, one would consider:

— the same criteria as for evaluating an individual researcher;
— additional criteria applicable to the team or the institution, such:
— size;
— degree of integration of the team;
— leadership;
— management;
— efforts towards the strengthening of research capacity;
— concern for valorisation;
— outside image of the research team.

5. Final Considerations: a Call for Guidelines for Evaluation of Development Research

5.1. The Need for Guidelines

As we have seen substantial differences exist of both criteria and methods between disciplines or groups of disciplines, between users of evaluation, or according to the objectives assigned to any specific evaluation. There is no universal instrument for all situations – actually there cannot be such an instrument. Persons in charge of evaluation will have to establish their own criteria and procedures, and elaborate themselves their own evaluation tools: grid, questionnaire, list of instructions, etc.

Yet it is possible to enounce general and broadly acceptable rules and principles, which then would possibly materialise as a set of general guidelines. The Academy has accepted her responsibility in this matter, reflecting its independent and multidisciplinary way of operating.

5.2. Nature and Characteristics of such Guidelines

The guidelines should be a short and well-structured document made widely available by the Academy. They would be "guidelines" sensu stricto: precise yet not imperative. They would provide their users with general principles to be respected — or at least to be taken into consideration. They should not be a handbook for evaluation. The text should be concise and short, allow for broad adaptations, and not be restrictive. Shortness will in addition favour their dissemination.

5.3. Guidelines Content

At this stage it is too early to propose a final list of content. It seems reasonable to assume that useful and practical guidelines should contain various aspects, possibly covered in different sections such as:

(1) A general section, in which the guidelines would be presented : justification, nature and characteristics, potential use.
(2) A section on specific technical points. For example:
   — the objectives of evaluation
   — a consideration of the specificity of disciplines or groups of disciplines. The three broad areas of interest of the Academy could be a starting point : the human and social sciences, the natural and medical sciences, and the engineering and technical sciences
   — the pros and cons of a variety of criteria and procedures
   — the main conditions for a meaningful evaluation.
(3) A final section on how to build an evaluation tool specific for a given place and use. In this section the guidelines should emphasise the needs and the expectations of researchers in the South.

5.4. Writing (?) and Dissemination of the Guidelines

While they are presented here as a prime responsibility of the Academy, it would be highly desirable that other interested institutions – particularly those that already use "ad hoc" evaluation tools for application to development research would share their experience and ideas with the Academy.

It may also be suggested that, after a prudent period – of for example four years – the effective use and the usefulness of the guidelines be assessed jointly.

ACKNOWLEDGEMENTS

This paper owes much to the work of the Academy's Task Force on Evaluation of Development Research (January 2006 to the present). Special thanks are extended to Morgan De Dapper, Jean-Jacques Droesbeke, and Georges Stoops.

REFERENCES

BEGHIN, I. 2004. *In*: New challenges for the Academies in a changing world. 75th Anniversary, Royal Academy of Overseas Sciences, Brussels, 9 & 10 October 2003, pp. 69-74

STENGERS, I. 2009. Au temps des catastrophes: résister à la barbarie qui vient. Editions La Découverte, Paris, 205 pages.