

Debate 4

9 Keywords

- Methodological criteria
- Experimental methods
- Quasi-experimental
- When to evaluate
- Baseline
- Complementarity
- Rigour
- New assessment paradigm
- Instruments

◆ How to assess the impact of training methodologies and types of policies and programmes

✚ In the application or use of any methodology used to evaluate the impact of vocational training (VT) a number of **premises** should be borne in mind which, in my opinion, are indispensable.

Among these premises the following should be included:

Control and account for material, human and financial resources used in VT, as well as the time it takes. Although this is difficult to measure because it is often not taken into account, or it cannot be defined, a way must be found to do so from the beginning of the VT programme to be assessed.

Determine who is to receive the VT programme and what tasks they carry out. On the basis of this, direct outcome indicators can be defined, or indirect indicators based on levels of management reached by those who have undergone the programme.

Define the strategy of the organization implementing the VT programme (vision, mission, targets and objectives).

The heads of the organization carrying out the VT programme must be aware of the need to measure impact on the basis of the belief that only by evaluating impact can its effects be determined and based on these, carry out a process of continuous improvement of this VT.

Allow both qualitative and quantitative evaluation.

Clearly determine:

- For whom is evaluation carried out?
- What is being assessed?
- Who is being evaluated?
- When is the assessment being carried out?

- How is the assessment being carried out?

The methodology proposed must become a practical tool which will make it possible to evaluate outcomes and impact at different moments of the vocational training programme, independently of their features and characteristics (AN ASPECT WHICH IN MY VIEW IS COMPLICATED BUT NOT IMPOSSIBLE).

As essential elements in the impact evaluation methodology to be produced, the following should be considered:

General and specific objectives of the methodology.

Responsibilities and scope of the parties involved in the vocational training programmes to be evaluated (actions to be carried out at different stages).

Stages of the programme at which the evaluations are to be carried out, what is being evaluated and how. This is where the indicators and tools to be used at the different stages should be defined precisely; some of the ones we wish to highlight are:

- Ø Questionnaires
- Ø Interviews
- Ø Satisfaction surveys
- Ø Analysis of outcomes and information
- Ø Evaluation of the behaviour of specific indicators
- Ø Experts' criteria (Mario Miranda)

These are some contributions from the book *Metodologia de Avaliação* (Isaura Belloni, Heitor de Magalhaes and Luiza Costa de Sousa): Institutional and public policy evaluation becomes more significant when it can provide information not only about impact, but also about the broader and more diffuse outcomes of the actions carried out. In this regard, evaluation should include the process of designing and implementing actions and their outcomes. It thus becomes an essential instrument for decision-making and for perfecting or reformulating resulting actions.

The area of evaluation aimed at institutions or public policies is characterized by its internal similarity, from the methodological point of view. The differences between them involve the object of evaluation: in institutional evaluation the aim is to deal globally and systematically with an institution. Institutional policies and the programmes and projects developed in order to implement institutional policies are analysed; the achievement of the objectives of each of the policies and projects is examined, as is the attainment of the objective and purposes of the institution; that is, its "institutional mission". In the evaluation of a public policy, the institutions responsible for its design or political implementation are not a part of the evaluation objective. In this case, only the policies are examined – understood as a government's group of trends and actions in order to achieve specific objectives – and their outcomes and consequences.

An institution or a public policy always has broader and more meaningful purposes and roles than those expressed in its explicit objectives, mainly with regard to its coordination with other social policies and institutions. In this respect, it is necessary to develop conceptual efforts and

experiences with a view to producing methodologies which will enable the evaluation of effectiveness, efficiency and the social effect of public policies or institutions. (Andrea Barboza)

✚ "The operationalization of the methodology of evaluation in vocational education"

Evaluation research on institutional evaluation is similar to the type of research developed in the area of human science, inasmuch as it covers two of its basic principles:

(1) To contribute to the advance of knowledge regarding the object of study / evaluation, and

(2) To submit to the conceptual and methodological rigour which is characteristic of the scientific method, bearing in mind the criteria of validity and reliability.

The differences derive from the objectives pursued and the characteristics of the object of analysis itself.

Evaluation research generally focuses on the analysis of the adaptation and relevance of policies, programmes or projects with regard to explicit objectives, with the purpose of identifying the factors which promote or hinder their development.

The objective of evaluation research is generally an action or intervention carried out by one or more institutions, often public, together with a focalized social sector. Therefore, its conceptual and empirical object is clearly delimited by the policy, programme or project to be evaluated. The purpose of this type of research is to advance knowledge in a very specific way, since it focuses on the formulation of subsidies aimed directly at decision-making and action, particularly for the redesign of policies. Its purpose is also to contribute to conceptual progress in the area or subject of the policy which is the object of the evaluation. Its results not only enable arriving at conclusions and making recommendations, but also reflect a certain degree of valuation. Conceptual and methodological rigour is the instrument which supports the formulation of consistent and legitimate value judgments, always with relation to explicit definitions or conceptual assumptions. When the object is the evaluation of an institution, the focus is no longer on a project or programme, but on the institution as a whole; that is, the multiplicity of actions, projects and programmes which compose its institutional mission. The object of the evaluation is in this case broad, and contains several lines, including a process which should be carried out independently, but linked to the institution as a whole. In the case of the evaluation research into a specific policy, generally the institution designing the policy is not a part of the evaluation object, except with regard to the policy itself. (Andrea Barboza)

✚ Impact evaluation should be carried out at different points of the training programme – I have already talked about this idea – and the references should be established before implementation, during implementation and after implementation. Although impact on persons and on organizations can be evaluated after implementation, before and during implementation **partial outcomes and impacts** are obtained,

in our view (see definition of impact in DRAE – the Spanish dictionary – a trace or impression which remains), which should be assessed. (Mario Miranda)

Experimental, quasi-experimental, non-experimental methods

- ✚ When global programmes are involved, it must be borne in mind that that totality is constructed by the addition of several or multiple learning processes and what we should not do, in our opinion, is substitute one impact for another. In short, we need each person to learn and obtain a result, and on that basis, the whole must be constructed.

Although received doctrine says the opposite, control groups are often not necessary. When a worker proposes an impact to be obtained (for example: to resolve a problem at work, or improve his or her results, time, sales, etc.), we only need a point of contrast which will provide reliability. This contrast can be provided by the supervisor or a colleague.

(Carlos Gómez Plaza)

- ✚ According to definitions of quasi-experimental models proposed by Ernesto Cohen and Rolando Franco in their book *Evaluación de Proyectos Sociales* ("The Evaluation of Social Projects"), the application of this methodology is useful in the evaluation of active policies being analysed. Probably what the authors call 'periodic series'; that is, a 'before' and 'after', constitute alternatives to experimental models and the selection of a control group is substituted by comparison groups.

At the Buenos Aires meeting we shared with you the salient characteristics of the evaluation we designed in order to learn to what extent actions carried out had managed to install and develop a *Sistema Nacional de Formación Continua*, SNFC (National System for Lifelong Training). Progress has been made in a design which considers the System as a whole to be the object of evaluation – the network of relations and devices which generate the conditions for the development of lifelong training actions, as well as the impact of training on the employability of workers. In this regard, the idea was to establish a starting point from which vocational training actions were carried out by the Labour Ministry without the intervention of the *Sistema Nacional de Formación Continua*. The challenge of evaluating the impact of this policy was that the starting point was the absence of the System and during the process of its creation, the impact caused by its actions had to be evaluated.

It is important to acknowledge this, because it is the policy itself, deployed in this case by Argentina's Labour Ministry, contained within the social dialogue of actors representing the different sectors of activity, which is the object of evaluation. How have the actions developed in the consolidation of the *Consejos Sectoriales de Formación y Certificación de Competencias Laborales* (Sectoral Councils for Training and the Certification of Labour Competence) – tripartite settings for social dialogue – influenced design and implementation, and the commitment of actors with regard to training policies? Equally, have the actions implemented produced objective and subjective changes in the living conditions of workers?

That is, the specific nature of the object of the evaluation, in this and in every case, determines the objective of the evaluation and the methodological strategy, which includes the feasibility of the study designed; it is no use committing to an impact model which, for ethical or technical reasons, cannot be implemented. In this respect, **it is not advisable to consider a universal impact evaluation model**; alternatives to the dominant models should be assessed, which will make it possible to arrive at valid and reliable outcomes without ethical, political and technical questions arising.

Impact evaluation should make progress in determining to what extent, and as precisely as the object allows, have the situations described in the baseline been modified as a result of active intervention policies. In our case, whether progress in institutionalization and the improvement of employability and quality of employment are due, at least in part, to the creation of the SNFC. We should recall that we are considering a public policy, which is therefore wide-ranging, not the development of programmes which by definition are more circumscribed with regard to coverage and time.

The first decision we made was to define the data comparison model and determine to what extent changes in situations which gave rise to the actions would be considered meaningful. The difference between the initial diagnosis and the successive evaluation points will make it possible to learn how much and in what way the situation has improved amongst the population and what progress has been made in the development of the SNFC, regarding its basic institutionality and the actors who participate in it.

The comparison strategy, which is the core of the evaluation model, implies the design of longitudinal studies whose characteristics are determined by: the variables or aspects to be considered, the type of follow-up required and the possibility of having reliable data available prior to the development of the actions. A serial design will be used whenever it is possible to work with periodic measurements prior to the implementation of actions, with the purpose of constituting a preintervention trend. Serial designs are the most rigorous of the non-experimental methods, since although they do not use control groups, they enable a calculation of the trend which the phenomenon to be analysed would have followed if the actions implemented had not existed.

Whenever relevant, cohort serial designs will be used in order to obtain a more focalized analysis of change processes, particularly of indicators related to the labour insertion of workers, the conditions of the job obtained, and the conditions of employability. Serial designs are particularly efficient when the objective of the actions is to reach the whole of the population without specific limitations, in this case, the workers and the VTIs.

If quality data are not available for the construction of the trend prior to the intervention, a pretest-posttest design will be employed, in which indicators before and after the intervention are traced, evaluating change on the basis of the statistical significance of the variables. Although, on the basis of these designs, the effects of factors foreign to the anticipated actions cannot be isolated, it is possible to state that the

actions have contributed to the outcome, but together with other non-isolated factors.

We are convinced of the complementary nature of qualitative and quantitative methodologies. In this area, we propose interviewing actors who participate in our policies in depth and through a focus group.

In short, we believe it is important to consider evaluation models which contemplate local realities, the characteristics of the processes involved, the information available and their feasibility. If we can produce different techniques for the construction of data, which include the qualitative and quantitative dimensions of the policies under analysis, we could consider alternative models which would approach the aspects to be evaluated rigorously. (Susana Barasatian, Nora Morales)

◆ How to determine which methodology is the most suitable. Aspects to be borne in mind when making decisions

Methodological criteria

- ✚ General criteria which, for some researchers, should be included in the evaluation of training.

One of these criteria is based on three main elements:

1) The OECD's IMHE, which points to the need to contemplate five broad perspectives:

- The reference context
- The potential of the system
- The input variables
- The process
- The output

2) The evaluation of outcomes (vocational, social and personal) in which it is necessary to show and make public the evidence which validates the certification of the training profile involved.

3) The proposal of the Association of Dutch Polytechnics and Colleges which proposes five approaches: based on the discipline, economic factors, the consumer, the labour market and organizational innovation.

For other scholars, the criteria to be considered in training evaluation should be based on a new evaluation paradigm which is reflected in the following principles:

Ø **The primacy of purpose**; that is, the need to achieve the established objectives. This principle takes us from a rationale which focuses on input to a rationale focused on outcomes and the processes which produce them.

Ø **The principle of imputability**; that is, the assignment of responsibility. Institutions and the system of which they are a part, as a whole, are required to be accountable to society for their effectiveness and efficacy. To this end, it should be ensured that internal evaluation is associated with external evaluation, at all levels of the system.

Ø **The principle of subsidiarity**. It should be ensured that decisions are made at the same level at which they are to be applied, or, at least, that these levels participate in them. This is an essential principle in the systemic model.

Ø **The principles of self-organization and self-development;** that is, the autonomous operation of institutions and organizations. Every institution and every organization should be a fundamental core for reflection and innovation. Reflection through evaluation should be at the service of quality. (Humberto Aguirre)

Evaluation in general is conducted to determine the worth of a programme with the view to providing basis for decision making. Impact evaluation is used to determine whether a change has taken place as a result of an intervention. Hence, both quantitative and qualitative data can be collected and analysed to provide justification for decision to be taken.

To determine the impact of a programme it is vitally important to have a baseline data on the dependent variables, in this case, quality of vocational training and in-plant training. The question is whether we are seeking information on the impact of VT and In-plant training in the work place i.e., on productivity, or are we looking for variables that will help us improve the quality of VT and In-plant training? These are important questions to answer in order to narrow the focus of the evaluation and data collection.

In seeking to determine the impact of an intervention, experimental designs are useful in addressing evaluation questions. The design may be pre-experimental or quasi-experimental to provide for control and experimental group including randomization of the sample group. This gives validity to the result because one is certain that the outcomes are due to the intervention or treatment.

The selection of methodology will depend on the objective of the study and evaluation questions. Other extraneous factors that may impact the independent variables also need to be taken into consideration.
(Hassan Ndahi)

What, when, how, who evaluates

- ✚ It should be pointed out that this debate is addressing impact evaluation. In other words, it is not about the evaluation of satisfaction, or the evaluation of learning; nor even about the evaluation of applicability.

When we refer to measuring impact, we mean measuring, in the case of an employed worker, the improvement of outcomes which arise as a result of his or her learning process (productivity, working conditions, etc.) and if we refer to an unemployed worker, we mean whether or not he or she obtains a job as a result of the learning process.

At some point of the debate, we refer to "outcomes" and to "evaluating impact at different points of the training programmes". In this regard, it is necessary to distinguish between: the monitoring measures which are set in motion during the development of the programme in order to ensure that the training programme is maintained in compliance with the anticipated objectives, from the measurement of impact which can only

be produced when what has been learnt is applied to reality. In view of this, in most cases it should be considered that impact can only be measured after the learning process and not before or during.

If we refer to the points at which impact can be measured, we shall say it can be done at the time when the person who learns applies what he or she has learnt to the job, and a reasonable perception can be constructed of the new outcome obtained. This means that the closer to that moment we measure, the easier it will be to distinguish the effect produced by learning on the outcome, and separate it from other factors external to training.

This does not imply that measurements cannot be taken much later (6 months, or a year), but to disregard the first moments when learning is applied (2-3 weeks) implies complicating the clear perception of the origin of the outcomes obtained.

The objective of impact from the first moment when learning is designed should be to measure it; for which purpose we require a method which will tell us whether training is worthwhile or not. To this end, from the very beginning it is necessary to have a person who needs to learn something, to determine what he or she needs this learning for, and then ascertain that **what has been learnt is applied and improves his or her work and, as a result, impact is obtained in the outcomes, which is translated into benefits**. If not, we shall not be measuring impact but evaluating or performing a follow-up of the training process (satisfaction, quality, level reached, etc.)

In this respect, the impact data to be obtained should be determined at the time when the learning objectives are established. That is, at the point when needs are identified, it is necessary to obtain the impact data to be achieved, both qualitative and quantitative.

With regard to who should evaluate, there are two matters to be considered. The first is that the person who learns will be the person who will experience the desired change; we should then be able to access this person and his or her perceptions. The second is that if we are seeking a contrast, we should attempt to get the person closest to the issue (supervisor or colleague) to participate.

Teams external to the learning process usually generate a very high cost which is related to the cost of the understanding process for someone who is removed from the issue in question (the learning process). In our opinion, such teams do not contribute greater reliability than that provided by the person who learns, as long as we can guarantee freedom of expression. (Carlos Gómez Plaza)

Instruments

- ✚ It is indispensable for evaluation to develop objective and valid methods which will ensure the **reliability** of the system and that it is recognized by the different sectors.

This is achieved by involving all of the sectors through the transfer of knowledge (competency standards) which they offer the education

system, so that, in turn, the system may align its curricula and education becomes more relevant.

The sectors and the education system should develop objective and valid methods with which to determine that programmes are relevant, but the most important aspect is that students or workers should have access to the knowledge, abilities, skills and behaviour they need in order to perform a function or an occupation at a high level of efficiency. According to the above, it is important to:

- Produce measurement instruments with which to control to the utmost the subjectivity of the evaluation.
- Design tests in order to assess the effectiveness of workers' or students' performance, clear evidence which will determine that they have acquired practical abilities, the technological knowledge required to carry out the various operations of their profession or occupation, knowledge of tools, equipment and material, and precision and speed in order to perform the tasks involved. (Magali Soria)

- ✚ Regarding the instruments or tools to be used, we believe that in the case of small learning groups (for example: one person, or a group of 15-20), no special support is necessary. However, in the case of habitually larger groups, it is necessary to seek suitable support, or computer or data transmission tools. In such cases, it is essential for the system to determine every aspect, from the identification of needs to the evaluation process and to enable the participation of the person who needs to learn and of those who need to corroborate data.

Impact should aim to produce the return of investment in quantitative terms. That is, we should be aware of how much the learning process has cost for each economic unit, what return has been obtained, in order to know the profitability of the investment.

In qualitative terms, it is also necessary to attempt to determine the return, although in this case outcomes may not be so conclusive.

Impact evaluation should always bear in mind the quantitative and qualitative aspects. The key elements of the return should always involve two indicators: length of time at work and/or income variation. Time and money are the two solid elements which can lead to the most significant estimates in order to be able to change decisions. (Carlos Gómez Plaza)

- ✚ Amongst the instruments to be used in evaluation, the literature I have consulted highlights the following:
 - Surveys, which can measure levels of satisfaction of people receiving the training in relation to their expectations and need of the training received, as well as of the beneficiaries of the training carried out.
 - Tests, presentation of team work, role playing, etc. which enable an evaluation of assimilation levels (learning) of training in those who have received it.
 - Evaluation of performance enabling an assessment of the application or implementation of knowledge and abilities acquired during the training received, based on the outcome of work compared to previous stages.
 - Analysis of behaviour or tendency of productive, services or management indicators related to the tasks carried out by those who received vocational training, enabling an evaluation of outcomes or

impact of the application of what was learnt during the training received (cause-effect relation).

- Estimate of the rate of return of the investment made in the vocational training provided, related to being able to measure in monetary terms, how much was invested and what outcomes were obtained. Using this instrument means being able to convert to monetary values all of the resources used during training and the benefits and impacts related to it. There are theories and models associated with this tool (Phillips vocational educational and training model). (Mario Miranda)