

Informative Sources for the Evaluation of the University Education Effectiveness in Italy

Stefano Campostrini, Simone Gerzeli

Dept. of Applied Statistics and Economics "Libero Lenzi", University of Pavia, Italy

Summary. The evaluation of the effectiveness of a study programme refers to its outcomes, and may be measured with the level of satisfaction of the programme objectives. In an educational programme, we can recognize three macro objectives referred to students:

- i) the formation of specific competences;
- ii) general competences and individual cultural development;
- iii) capability in finding a suitable job.

The measurement of these components is based on the construction and analysis of several indicators from surveys on teaching assessment, or on placement, or on teachers and employers, and, too, on linkage between databases.

Keywords: Education effectiveness; Evaluation; University; Informative sources.

1. The evaluation of effectiveness

The evaluation of the effectiveness of education involves several elements of complexity, from the definition of effectiveness itself to the multiplicity of programmes that pursue educational objectives. We will not linger over the delicate question of the definition of effectiveness, accepting, as an operational definition "*the degree to which the educational objectives are achieved*".

Hence, the evaluation of effectiveness implies a clear identification of the objectives and of the expected results, as well as of their measurement. The complexity of an educational programme makes the definition of the objectives even more important: only a clear, specific, precise declination of the objectives allows us to define the concept of effectiveness to which the programme refers. We could say that the definition of the objectives also contains the definition of effectiveness.

Thus, if an initial definition of the objectives leads to a precise determination of the possible results, a subsequent effort must be undertaken to define effectiveness, by indicating *what it possible to measure is*. Often at the initial stage just proxy measures are used, measures close enough to the concepts expressed by the objectives.

In defining objectives, particularly in the social context, we often use concepts (for example, the “quality of teaching”) which are not directly measurable. However, each of these can be made operational by identifying the components of the concept itself, which make it easier to identify possible measures, which nevertheless are proxy ones.

In the evaluation process, inadequate conclusions may be drawn if the outcomes are undefined. This is why it is risky to pay attention to only one outcome indicator and to evaluate the effectiveness with even small variations of this indicator.

The multi-dimensional nature of the social world implies a multiplicity of indicators. The basic idea behind evidenced-based approaches refers, on the one hand, to the impossibility of precisely measuring the complex situation the programme is on about and, on the other, the importance of choosing evidence indicators in order to make sufficiently well informed decisions.

A possible guideline for an operative definition of effectiveness may help also in identifying the information needed for its evaluation: how to make explicit the central feature of the definition of the objectives specific to the university programme and how it may be measured. The process is typically circular: from an initial definition of objectives, we determine which of these can be made operational and measured, and this often leads to the redefinition of the objectives themselves.

Thus, going back to the definition of effectiveness as the “degree to which the objectives are achieved” it may be added¹ “*provided the objectives are clearly defined*”.

The theory thus indicates that it is difficult to adhere to a single definition of effectiveness whatever this may be, that is valid for every project and every context; but this is due to the very same definition of the objectives that the educational programme intends to set forth.

An important corollary of this issue is that no evaluation of effectiveness is possible if the programme itself has not stated what concept of effectiveness it refers to. Hence, each evaluative project should explicitly mention which concept of effectiveness is referred to, if it is a partial or overall measure of it, and to what extent the “proxy” indicators are valid, reliable and sensitive.

¹ This should also solve the problem for defining effectiveness as the “satisfaction of needs”. It is the definition of the objectives that should account for how much and in what way the programme needs are to be satisfied; this should not be dealt with through the search for their (ex post) effective satisfaction outside of the objectives set forth by the project.

2. The components of university education effectiveness

In order to manage the complexity of programme evaluation, we attempt to determine (macro) objectives typical to the evaluation process without discussing the peculiar aspects of each of these. This approach aims to identify the macro-categories that are measurable.

University education typically unfolds in ways that are often quite diverse depending on the framework of application, and involves profoundly different objectives. Nevertheless, it is possible to identify some common features, often because it is not the basic objectives that change but their relative importance. Thus, for example, the objective of finding a job cannot be the primary of university education, but at the same time, it cannot be ignored. Likewise, in a high-level university education (for example, a Master programme) the cultural development of the person, while not a priority, is an important objective anyway.

In Table 1, we outline the macro objectives that an educational programme could typically seek to achieve. We will try to bring out typical features defined by both the theory and the evaluative practice.

Table 1. A framework for defining effectiveness in university education programmes.

<i>Objectives</i>	<i>What we can measure</i>	<i>How</i>
Learning / development of specific knowledge	Satisfaction of the students during and after their studies Self-perception of educational advancement Assessment of knowledge/skill acquisition Perception of employers	Customer satisfaction surveys on teaching and placement Partial and final verifications <i>Ad hoc</i> surveys of employers Impact studies of different programmes Surveys of teachers Surveys of experts
Development of cross-occupation competences/ "formation of a person"	Satisfaction of students during and after their studies Self-perception of (<i>ex post</i>) educational achievement/ usefulness of cross-occupation competences Perception of employers Enrolment in more advanced studies	Customer satisfaction surveys on teaching and placement <i>Ad hoc</i> surveys on employers Surveys on teachers Surveys on experts
Employment capabilities	Employability, such as time required to enter the labour market Quality of work (contract / position / professional nature) Coherence of the job with the studies	Placement surveys Comparison of archives Surveys on employers Surveys on teachers Surveys on experts

Each of the items presented in the table would deserve a full and deep discussion. This exceeds the aims of this paper, which seeks to take advantage of this synthetic presentation to discuss the possible sources for evaluating the effectiveness of University education, their pros and relative limitations. Moreover, the list is not meant to be exhaustive.

The scheme obviously does not consider specific cases and the classification cannot capture all the complexities of particular educational programmes. On the other hand, such an approach perhaps cannot reveal other important issues faced by university education². If we consider only three macro objectives, we accept the risk of leaving out many important aspects. However, given the complexity of the problem, we have sought some solid footing to start from, with the possibility of subsequently widening the scope and detail of these considerations.

3. The sources for evaluating effectiveness

We will now examine the main problems and possible methodological solutions that can be encountered when using the typical instruments for evaluating the effectiveness of educational programmes. We begin with an outline in order to discuss, though briefly, each element.

Table 2. Typical approaches to gather information on the effectiveness of education programmes, and related problems

<i>Sources and instruments for evaluating effectiveness</i>	<i>Possible problems</i>
Evaluation surveys of teaching methods	Subjective evaluation
Placement studies	Subjectivity Timing of data collection Selection bias
Linkage of archives	Data availability
Surveys on experts	Subjectivity / lack of quantitative elements
Surveys on professors	Subjectivity / partiality
Surveys on employers	Subjectivity Coverage and rates of response

² For instance, we lose the trade-off between quality of education and the increase in the number of educated individuals, a problem that is dealt with and not yet resolved by the reform of the educational system in Italy.

3.1 Evaluation surveys of teaching methods

The evaluation of teaching is quite a well-known approach around Europe. In Italy its application is recent (late 1990s), but it is routinely used in a most educational programmes, and is compulsory in public universities.

Even though the data-gathering instruments and the subsequent construction of indicators have become more refined, one would be wrong in thinking that the validity and sensitivity of these elements allow for indications of absolute effectiveness of the educational programmes being studied.

In fact, the specific role of this evaluation approach, especially when we limit ourselves to gathering information from assessments by students, can only represent a general stimulus for the individual teachers and a sort of alarm signal with regard to certain particularly critical situations.

It would be deceiving (and in many regards counterproductive) to attribute to these types of evaluation the role of an absolute judgement on the quality of teaching and, even more, on the single teacher. In fact, the reliability and sensitivity of the instruments can certainly allow us to make valid considerations at the macro level as well as regarding the “tails” of the distributions (the left one in particular, the relevant negative cases), while the validity of comparison in the central part of the distribution can be questionable. I.e., a teacher who receives an average score of 3 out of 10 is certainly a problematic case, while it is not necessarily true that the performance of a teacher with a score of 7.5 is below that of one whose score is 8.

Thus, the methodological solution to the problems linked to the subjectivity and partiality of the evaluations does not involve adopting some technical tricks but limiting ourselves to the role that evaluations of this type can play and understanding the importance of interpreting the results in light of other possible information, following an integrated approach (Nevo, 2001; Bernardi *et al.*, 2005).

3.2 Placement studies

Placement studies are also being used in different educational contexts. Typically, at a distance of six months, and one to two years from graduation, students are interviewed regarding their satisfaction as well as to obtain information on their employment situation. It must be emphasized that opinions given spontaneously and after some work experiences (even unsuccessful ones) can provide an interesting point of view.

Obviously, all *ex post* assessments provide us with a lot of information to define overall judgements, and are not in competition with the approach presented in the previous section, which focuses more on formative evaluation.

Timing is certainly a critical element: studies too close to the end of the educational experience do not sufficiently reflect the students’ rethinking of

their educational lives. However, if the surveys are too far from this experience, the rethinking can itself be strongly influenced by exogenous factors.

Once again, the specific evaluative objective indicates the appropriate timing with respect to this trade-off. In any event, experience seems to show a minimum distance of six months and a maximum of two years.

As far as the employment situation is concerned, we need again to pay attention to the value and the role that this type of informative source can play. The case of vocational training is emblematic. In many regional systems, evaluations on the employment capacity of formative programmes have been considered as points in favour for those seeking financing for their educational programmes. This use of the evaluation of results in selection processes is certainly positive and should be encouraged.

There is a mixture of aspects concerning both the definition of objectives and methodology. It is clear that, with regard to different study programmes available in a territory, a selection problem could hide the effectiveness – in terms of employment capabilities – of such study programmes. Students who choose subjects that are viewed as important by the labour market will certainly have more chances of success, for example in terms of employment rate six months after the end of the educational programme.

This typical problem of impact evaluation can be overlooked if we consider that the ability to select worthier and more appropriate students is itself part of the educational programme. Thus, we again have a problem of the definition of objectives: this would surely be a tenable approach.

Should we reward a school that offers a university education to less gifted students, or one that emphasizes quality, even with a rigid entry selection? If we do not consider selection problems, which can hide a result (positive or negative), we run the risk of wrongly judging the outcomes of a placement survey.

3.3 Linkage of databases

Linkage of databases could certainly be an interesting approach to overcome most of the above-mentioned problems. At a small cost we could compare the performance of persons that received a university degree with those without it (or who have had a different degree) by simply observing the social security agency status (INPS in Italy) after a certain period. When this has been applied to similar contexts, there have been encouraging results (see, among the others, Battistin *et al.*, 2001).

Apart from the methodological problem of linkage, the issue here is often a “political” one, involving the willingness of different institutions. The data exists, but what is often missing is the willingness to link the data for comparative purposes. A non-trivial problem is data privacy, although it is frequently used as an excuse.

Basic information for the evaluation, in terms of internal effectiveness or output data, could also be found in the university administrative archives. These databases allow us to construct indicators to measure the output of the university education, for instance the number of students and degrees, the number of courses, etc.

Of course, these archives are useless for outcome evaluation, although some information on effectiveness can be obtained.

In Italy, the universities' capability of carrying the students to final degree is taken by the Ministry for University as an indicator of performance of the universities.

3.4 Studies on experts

Often the data on the effectiveness of educational programmes is limited to those directly involved (e.g. students and subsequently to graduates), who undoubtedly are aware of the progress in their educational experience, but can provide only a conditioned point of view. A useful complement to this data can be "qualitative" studies on experts.

Though not useful to quantify effectiveness, this information can nevertheless be precious not only to complete the data that has been gathered but also to provide a key to interpreting the success or the failure of an educational programme.

3.5 Surveys of teachers

In many countries (and Italy is one of these) teachers are rarely surveyed for their opinions, almost as if researchers were afraid to inconvenience those on the other side of the fence. Actually, teachers can also provide interesting information, if not on the final effectiveness of programmes, at least on those elements of the *process* that favour or obstruct such programmes.

In fact, it is becoming increasingly necessary to include process elements in *ex post* evaluations: these allow us to understand not only if an outcome has been achieved but also how and why, and under what conditions. Those who manage this process can certainly provide this information. Thus, it is not simply a question of asking teachers to do a sort of self-evaluation of their work but to involve them in a more comprehensive evaluation programme.

In this regard, we must quote the disputes between those who believe that an evaluation should be internal to the concerned organisation (i.e. the organisation that must "learn") and those who feel it should be external, requiring third-party evaluators. It depends on the evaluative objectives, on who requires the evaluation, and the point of view the study will take (Chen, 1996; Campostrini, 2001).

A third-party evaluator is certainly necessary to guarantee evaluations that are independent of the context and free of peculiar interests, among the other targets, when it is used to formulate judgements for programme funding.

In the same way, we cannot deny the validity of internal evaluations, undertaken as a standard practice by the organization itself (Patton, 1997).

3.6 Survey of employers

As far as the educational objective of labour market entry and specific skills training are concerned, the point of view of employers is surely important, if not fundamental.

The problems in this context are technical and cultural. In fact, the basic limit to employer surveys is the difficulty of getting employers to collaborate. It is not only the problem of interviewing who are busy, often on the move and have little time available for activities that take them away from work. It is not only the problem of selecting the most appropriate person to interview in highly complex organizations. The problem is (also and above all) cultural.

This is the case in Italy, where employers have shown little interest in educational programmes. Recently there have been some notable changes, but these involve training on the job more than training for a job. Even companies that believe strongly in training programmes, as opposed to basic education, prefer to train their personnel themselves to provide specific skills than turn to some training agency.

Some changes are taking place, and it could be worthwhile to study the methodological problems for this type of survey and to launch some pilot projects. In the meantime, we feel that the qualitative approach involving experts presented above still represents the best proposal.

4. Final remarks

In this paper, we have tried to systematise the complexity of an evaluation project of university educational activities. Our purpose was to provide a possible framework, discussing the typical sources for evaluation. This has been done with reference to the Italian situation. Giving these limitations, our work can represent only a possible starting point for approaching the matter.

Further research is needed to better define the effectiveness of educational measures, in particular university ones. Although some approaches (such as customer satisfaction or placement surveys) are well developed both from a theoretical point of view as well as from the methodological one, some others (e.g. surveys on employers or data linkages) need to be further studied.

References

- BATTISTIN E., GAVOSTO A., RETTORE E. (2001) Why do subsidies firms survive longer? An evaluation of a program promoting youth entrepreneurship in Italy. In: LECHNER M., PFEIFFER F. (eds) *Econometric Evaluation of Labour Market Policies*, Physica-Verlag, Heidelberg: 153-182.
- BERNARDI L., CAMPOSTRINI S., PARISE N. (2005) L'impatto della valutazione della didattica sull'Università: il caso dell'Ateneo patavino. *Rassegna Italiana di Valutazione*, **31**: 103-123.
- CAMPOSTRINI S. (2001) Metodi e tecniche nella valutazione di politiche e programmi nella realtà del nuovo millennio, *SIS 2001 Convegno Intermedio Processi e Metodi Statistici di Valutazione*, CISU, Roma: 65-74.
- CHEN H.T. (1996) A comprehensive typology for program evaluation, *Evaluation Practice*, **17**(2): 121-130.
- NEVO D. (2001) School evaluation: internal or external?, *Studies in Educational Evaluation*, **27**: 95-106.
- PATTON M.Q. (1997) *Utilization Focused Evaluation. The New Century Text*. Thousand Sage, Oaks.