Title: *Evaluation research in education: the importance of a paradigmatic debate of fundamentals and practices*

**Abstract**

Evaluation studies, being a form of applied research that intends to study how existing knowledge is applied to the development of new educational practices instead of the elaboration of new theories, have no specific methodological design. In fact, evaluation research studies combine the most diversified methods and techniques regarding social sciences research either of quantitative or qualitative formats: from descriptive to correlative, from experimental or quasi-experimental to pure qualitative designs. All formats are considered to be applicable though there will always be some methodological ambiguity over evaluation studies.

In this paper the author begins by presenting and discussing the conceptual frameworks – at ontological and epistemological level - of the main paradigms that sustain research practices in the social sciences: positivism, hermeneutic and critical theories will be analysed and debated. Subsequently, the most important positions regarding the paradigmatic debate in evaluation research will also be discussed and under consideration: incompatibility between paradigms, methodological complementarities and, more recently, the posture of those who simply argue for the need of integration and epistemological unity.

The author finishes by reinforcing that the paradigmatic debate must be understood by educational practitioners not as a mere philosophical exercise, but instead, as a way to stimulate evaluators to develop more creativity in the choice and use of techniques and instruments for data collection in real world contexts; such an attitude can only be achieved by combining quantitative and qualitative methods, and thus, enabling the evaluation studies to become more efficient, realistic and useful to the progress of educational knowledge.
1. Evaluation research

The main characteristic that identifies and distinguishes the evaluation of other formats of educational research studies is, as mentioned by Shaw (1999), more in relation to its purpose, objective or criterion, rather than the research methods or methodologies themselves; in fact, the evaluation is always oriented towards action, this means:

- it is a format of applied research, as opposed to fundamental or basic research which is more concerned with the strengthening of theories or the support of existing knowledge;
- it compares the achieved results with objectives/aim/criteria previously defined;
- it intends to support a decision making process regarding what is being evaluated: to approve, to reject, to change, to implement;

Considering the methodological “procedure” itself, the evaluation design may take very different formats, from purely descriptive to rather experimental and in correlative (Knapper, 1980), or even more qualitative (Mason & Bramble, 1997), and therefore, some kind of methodological ambiguity may arise. According to Weiss (1975),

The evaluation applies all methodologies of social research. The principles and methods that are considered valid for other types of research are also valid. All that is known regarding the procedure, measure and analysis comes into consideration in order to elaborate and implement an evaluation plan. What distinguishes the evaluative research is not the method or the subject of study, but instead, it is the objective or purpose: decision-making. (Weiss, 1975, pp. 18)

Regardless of the model, timing or format, the evaluation is always a process of data collection and processing that may be obtained through various methods and techniques, namely enquiries, interviews, tests, or even other methods of direct or indirect observation. Considering a frame of evaluation, all methods are considered to be reasonable as long as they are able to provide valid and feasible data in order to support decision-making. However, it is known that different methods may lead to contradictory results if applied to the same problem/question/program under investigation, i.e. evaluation. Inappropriate methods may achieve misleading results, and hence, incorrect decisions maybe implemented whether in favour or against a specific program, method or educational measure under evaluation.
2. The paradigmatic question

According to Patton (1998), it is fundamental for the evaluator to make supported decisions. This can only be possible if the evaluator is well aware of the human and cultural conditions that restrict his “choices”, and therefore, give the results of the evaluation an inevitable “time” and “place” limitation, which means considering a theoretical framework known in literature as a “paradigm”\(^1\).

According to Hussén (1998), the paradigm adopted by the researcher will determine the way in which the research problem is formulated and approached in a methodological point of view. This very same idea is defended by Koetting (1996) when he says:

The methodologies bear with themselves interests that restrict the expected and obtained results, and for this reason, the researcher should look for the human interests that always lie beyond the different research methods. (Koetting, 1996, pp. 1141).

Like Koetting (1996), many other authors (Patton, 1990; Deshaies, 1992; Shaw, 1999) support the need for the researchers to be aware of distinct paradigms in order to be able to make decisions that are supported throughout the development of the projects of educational research. Looking for the theoretical fundamentals of research is somehow like taking a journey into the scientific thinking, or into the different conceptions of the nature of knowledge, which is a question that is not always considered, since we educational researchers are not used to ask “why” regarding the methodological choices we take.

2.1 Two divergent paradigms

Within social sciences, the debate that opposes the relative methods to the quantitative and qualitative methods is long now, and also, among evaluation literature, the question of the paradigmatic “great war” has long been discussed (Gage, 1989); and so, there are those who believe that this is a secondary question that should be approached only by experts in epistemological questions, and hence enabling the evaluators to focus mainly on practical questions (Miles & Huberman, 1988), and there are others, such as Patton (1988:119) for whom

\(^1\) The concept of “paradigm” was first presented by Thomas Kuhn in 1962 in *The Structure of Scientific Revolutions*; according to Kuhn, paradigms refer to an entire worldview of a scientific community and are the basis of all sciences, since they guide the research efforts of practitioners and adherents.
this question “…may help the evaluators in some way, taking that if they are aware of their methodological prejudices, they will be able to make decisions that are more flexible, sensible, and adapted to the context under consideration”.

Regarding the reality of the evaluation studies in the USA, the author believes that focusing the practical questions, and thus neglecting the paradigmatic debate, would have turned the process of evaluation into simple routines somewhat out of phase with reality; the paradigmatic debate should be seen not only as simple metaphysics exercise but instead as the engine for more creativity in the choice and use of the different data collection techniques within the real contexts in order to take clearer methodological choices that can be more adjusted to the reality under evaluation.

At the centre of the debate an ontological question regarding the nature of reality, two opposite positions regarding the relative merits of the divergent paradigms: the quantitative, also called traditional, positivist, rationalist, empiric-analytic (Bisquerra, 1989; Latorre, Del Rincón & Arnal, 1996; Usher, 1996; Mertens, 1998; Shaw, 1999;), and the qualitative, also called in literature as hermeneutic, interpretative, naturalist (Creswell, 1998; Crotty, 1998), or even constructivist (Guba & Lincoln, 1988; Guba, 1990).

2.1.1 The quantitative approach

For the supporters of the quantitative paradigm:

- The reality under evaluation is “objective” since it exist regardless of the subject; events happen in an organised fashion, and hence it is becomes possible to determine the actual rules that govern them in order to prevent and control them;
- There is a clear difference between the “subjective” researcher and the outside “objective” world;
- The credibility of research evidence depends on the manner in which the observation is done; different observers facing the same data should yield similar conclusions – replication guarantees objectivity;
The social world is similar to the physical world; the purpose of science is to discover reality, and so, both the natural and the social sciences should share the same logic based on a common rationality and methodology.

Assuming that the methodological processes have been correctly implemented, the obtained information is taken as valid.

The problem that we face is in fact, a research paradigm that emphasises the determinism (the truth exists and can be determined), the rationality (no contradictory explanations are allowed), impersonality (rather more objective than subjective), prediction (the aim of research is to generalise in order to control and prevent phenomena), and also, according to Usher (1996) some lack of reflexivity since the acceptance of the results directly depend on a proper application of the research methods, neglecting in fact the research process itself.

2.1.2 The qualitative approach

The qualitative paradigm supports, in an ontological point of view, a relativistic position – there are many realities based on mental creations and socially localised -, and is inspired by a subjective epistemology that enhances the role of the researcher/creator of the knowledge, and so, the use of a methodological frame that is not compatible with the propositions of positivism and post-positivism is justified. In a concise form, one can say that the qualitative paradigm intends to replace the concepts of explanation, prevention and control of the quantitative paradigm by the concepts of comprehension, meaning and action by which the personal world of the subject is penetrated, “…to understand in what way the different situations may be interpreted and what does it mean for them…” (Latorre et al, 1996:42), attempting to “…understand the complex world from the point of view of the one who lives” (Mertens, 1998:11).

Those who support a qualitative evaluation also reject that appropriately applying the research methods and techniques (methodological tools) one can guarantee objectivity in the search for knowledge; they believe that there are fundamental differences between the natural and social phenomena and also, the methods considered by the positivistic paradigm turn out to
be inappropriate for the study of the latter. According to Clark (1999) the traditional quantitative
approach, though it is technically solid, it enables an evaluation rather based on a “black box”
logic type in which inputs and outputs are related despite of many other important details being
totally neglected, and contexts of evaluation simply not being applied; he concludes, the
important however, is to get near the reality to study it from the “inside”, from the point of view of
the participants in the action.

2.2 A paradigm of choices…

On an ontological (the nature of reality) and epistemological (relation between the
researcher and the object or subject of study) level, such differences gave rise to two distinct
methodological approaches considered as the possible decisions that the researcher is able to
make regarding the different stages of his research. The quantitative approach follows the so-
called hypothetic-deductive approach, by which the casual explanation and prevention are
based on a deductive logic: the research is closely tied to a theory that supports and justifies the
attempts to explain the phenomena under consideration (hypotheses within research); the next
step is to collect data and to test the hypothesis that will be whether accepted or rejected2.

The other methodological approach defends an inductive logic within the process of
research; data is collected not based on a previously defined hypothesis that ought to be tested,
but instead, with the objective of finding regularities that will support generalisation evermore
broader. The deductive approach requires a previous decision from the evaluator regarding
what will be considered as a succeeded the program/measure and from which the results will be
subjected to evaluation (measuring). According to the inductive approach, nothing is defined a
priori: it is assumed that the deep knowledge regarding a program/measure and the following

2 The experimental plans constituted both the “classic” model of quantitative research in education and
evaluation (Rossi & Freeman, 1993). The logic behind the experimental designs works in the following
way: the evaluator forms two different groups of participants. To one of the groups “treatment” is applied
(in the form of a new curricular program or other forms of planned intervention), and to the other group,
either a different “treatment” or eventually, nothing is done (Stern & Kalof, 1996; Vogt, 1999). The first
group is called “experimental” and the other group, in which something different will happen, or eventually
nothing is called “control” group. Both groups are compared in relation to its dependent variable in order to
verify if the differences in results are expected and are caused by the “treatment”. The relation of cause
and effect is based on the assumption that the groups are necessarily similar in every way except
regarding the independent variable, or treatment, that they were subjected to. The fact that the
experimental group received the “treatment” unlike the control group, implies that the differences in DV are
caused by VI (Coutinho, 2005).
results can only be achieved through insights on the personal experiences of the partakers/participants. Patton (1986) underlines the implications of these two different strategies in the process of evaluation:

In evaluation the classical deductive approach is measuring relative attainment of predetermined clear goals in a randomized experiment that permits precise attribution of goal attainment to identifiable program treatments. In contrast the classic inductive approach is goal-free evaluation, in which the evaluator gathers qualitative data on actual program impact through direct observation of program activities and in-depth interviews with participants, all without regard to stated, predetermined goals. (Patton, 1986, pp. 194).

Similarly, Clark (1999) considers that the attempt to control the influence of marginal variables by forming equivalent groups in experimental plans is absurd as it does not enable the evaluator to be aware of the factors that, in fact, may be determinant to the success or failure of a specific program/measure; this means that such plans may, to a certain degree, show whether a program/measure achieved its objectives, but will rarely determine why the results occurred.

Both approaches require different methods in order to obtain data: the quantitative evaluator needs structured tools (such as questionnaires or structures interviews) with standardised categories that allow the individual answers to correspond to the latter. The qualitative evaluator records the individual opinions (interview is not structured, active or passive observation) careless of the categories of the answers; he assumes that to consider the individual characteristics of the participants is fundamental for the success of the program/measure, given that this will depend on their effort.

The first attempts to find alternative forms of evaluation were characterised by the change of the focus for evaluation of the objectives regarding observation and analysis of the developed activities within the process of evaluation, this means, the study was more oriented to describe and interpret rather than measure and quantify. Parlett and Hamilton (1976) denominated this as “illuminative” evaluation, in which the evaluator describes the perspectives and experiences of the participants of the educational program or measure, based on qualitative methodologies, such as direct observation and extensive interviews. According to Denzin (1989), the point of view of those for whom the program/measure is conceived is crucial for those who, in fact conceive the program/measure; given the complexity of the educational programs and existence of different views of the same reality, the view of the of the evaluator is
considered to be just one out of many possibilities, and therefore, in order to make an adequate
decision, the view of whoever is “on the other side” takes up a great importance.

At some point during the 60’s, a third paradigmatic approach emerged as an alternative:
the social-critical method based, on a conceptual level, on the Marxist philosophy, on the
critiques of Adorno and Habermas to the liberal economy, ranging from Marcuse up to the
unrestrained consumption present in capitalist societies, and on a pedagogic level, based on
the ideas of Paulo Freire, Michael Apple e Henry Giroux.

The basic assumption that supports this model is the following: if education is not
considered to be neutral, then the evaluation should not be either, this means, the possibility of
obtaining impartial results is rejected. The two paradigms existing at the time were considered
excessively conservative in the way they were exclusively concerned with the explanation
(positivism) and comprehension (interpretative) of the educational reality, however, made no
attempt to improve it. In methodological terms, the use of models in which the participants can
take part as researchers and hence, everybody can contribute to the social measure. Tójar
(2001) believes that the indissoluble relationship between research and implementation is the
key element that distinguishes this paradigm and accounts for its epistemological and
methodological status.

Compared to the empirical-analytical perspective centred on the explanation of the
phenomena, or to the humanistic-interpretative perspective that intends to understand the
meaning of actions, such model aims to implement educational practices that will ease the
autonomous, liberating and critical development of active citizens who take part in social
changes; within this context, the processes for evaluation turn out to have function of
emancipation that will in fact, enable the student to develop capability for self-reflection and self-
evaluation (Alves, 2003).

2.3 The epistemological debate

From the point of view of the theoretical principles, paradigms seem to be purely not
compatible: in fact, this the “purist” (Rossman & Wilson, 1985) or “monotheist” (Lecompte, 1990)
position that has been shared by many authors including Guba & Lincoln (1988), for whom the
differences that exist on an epistemological and ontological level would eventually any possibility of “mixture” on a paradigmatic level:

Just like water and olive oil, paradigms cannot be mixed; moreover, mixing them is distorting both. Just like magnetic poles, they will repel each other: uniting them requires strength, and when such strength declines, the methodologies that support them will come apart (Guba & Lincoln, 1988, pp. 111).

In fact, a considerable part of the last three decades of the past century was characterised by the epistemological debate between the different paradigms, more precisely, between what is usually called the quantitative and qualitative focusing. In the perspective Lukas and Santiago (2004), the argument that came up was erroneous from start due to the following reasons:

a) A defined method with a specific paradigm is determined;

b) The paradigmatic debate is based on meta-theoretical assumptions;

c) The possibility of joint utilisation of quantitative and qualitative methods is neglected;

Rather than debating the core questions, the clash between the anti and pro positioned participants was limited to the discussion of a number of mutual accusations and remarks connected with a strong ideological and emotional sense that scarcely contributed to the quality of the educational research in general, and to its evaluation in particular. However, since the 80’s, the fierce confrontation that supported the implementation of the only existing model, gave way to a quieter situation in which the possibility of mutual influences began to find acceptance. According to Walker & Evers (1997), one can speak about the three moments of the epistemological debate between paradigms:

a) Incompatibility: it is considered that the different paradigms (two or three) are epistemologically different because of their immeasurability and inappropriateness – this is the “purist” (Rossman & Wilson, 1985) or “monotheist” (Lecompte, 1990) position that has been shared by many authors such as Guba & Yvonna Lincoln (1988);

b) Complementary: similarly to the approach above, it is believed that paradigms are different on an ontological and epistemological level, though it is accepted that the researcher is not obliged to take part in one of them, namely the best one; it is held
that the complementary of the quantitative and qualitative methods as a function of what appears to be the best solution for the problem under consideration;

c) **Epistemological unity**: rejects confrontation between paradigms arguing alternatives that may implement or improve the “old” confrontation.

Within the supporters of the last position - epistemological unity - we can find different positions. Hence, the extreme position that utterly supports the adoption of a new paradigm, to which Miguel Diaz (1988) and Nisbet (1988) call the “paradigm for change” and “research oriented to decision and change” respectively. Along this same line of paradigmatic integration but not as radical as above, is the position defended by Kenneth Howe in several papers published between 1985 and 1992; starting by denying the obligatory of the methodological choices, the author argues that the necessity (and opportunity) of the methodological combinations, because if the paradigm is positivist it will become unbearable and the interpretative incomplete. Therefore, an epistemological platform that is able to overcome the antagonism by integrating the aspects of both paradigms must be found (Howe, 1985, 1988, 1992).

Alternatively, one could neglect all theoretical and epistemological related aspects of the paradigmatic debate (central to the positions mentioned above) and focus the discussion on a more methodological and technical level. Such position was shared by Cook & Reichardt (1979, 1986), who consider the debate to be outdated not because it denies the importance of the paradigms and their support on a methodological level, but instead, because of its inflexibility regarding the “monotheist” position that in their view has only the creativity and innovation within the research under evaluation. This gave rise to the acceptance by the evaluators of more flexible paradigmatic choices that were also more adapted to the real problems under evaluation:

There is no need to choose a research method solely based on a paradigmatic frame, let alone choose between two opposing paradigms. Therefore, there is no need for a dichotomist choice on a methodological level, and there is all the reason (based on pure logic) to use both paradigms jointly in order to satisfy the requirements of the evaluation that we expect to be efficient (Cook & Reichardt, 1979, pp 27).
The methodological integration within the development of the evaluation that we wish to be valid and realistic is nowadays the feeling shared by most of the authors that were concerned with such questions. In the view of Salomon (1991:10) “…though the winds of approximation have only recently started to blow, not due to a mutual love between “positivists” and “qualitativists”, but because researchers have understood how much complementary could the paradigms be to on a practical level of research”. In reality, we have seen attempts for the integration of the quantitative and qualitative methodological perspectives that suggest a future of methodological complementary, rather than the traditional antagonism. Patton (1990) speaks of a “paradigm of choices”, while Dendaluce (1995) supports a “integrator pluralism” by proposing a number of formal basic and common aspects – critics – to all research methodologies; also, he points the need of the educational research to “transcend the qualitative debate” (the title of his article) and to be aware of its object of analysis: a complex reality (classroom, school, family, culture) where the implementation of inter-related variables converge (behaviour, perception, attitude, expectation) and whose analysis and study cannot be approached in an exact science manner in which variables and individual facts can be isolated. Since educational research investigates the “whys”, rather than a one only methodology, it is expected a number of diversified methodologies that are able to combine the best of each paradigm: to combine the “analytical” accuracy of the quantitative paradigm with the “authenticity” of the interpretative systematic approaches to educational research is – citing Salomon – “…a combination that is nothing but superfluous; is rather a necessity if we really wish to achieve valuable results” (Salomon, 1991:17).

In conclusion, what determines the methodological choice of the researcher should not be the use of a specific methodology, a specific paradigm, but in fact, the problem under investigation. Understand the essence of the education phenomenon should be the centre of the question around which all methodological choices should be organised by the researcher; Bachelard (1971) called it “the meaning of the problem”, which he considered as a sine qua non condition for the existence of a “real scientific spirit”.

3. Synthesis
In this paper it was intended to discuss some aspects that should be considered if to start a research project in which the main objective is the evaluation of a specific program, curriculum, methodology or even a new software or experimental prototype.

Given the different perspectives that have lead to the discussion regarding the specificity of evaluation research, the main conclusion is that in a investigation of this kind if solely practical questions are considered, and the paradigmatic debate is neglected, the processes of evaluation may turn into simple routines somewhat out of phase with reality; the paradigmatic debate should be seen not only as simple metaphysics exercise but instead as the engine for more creativity in the choice and use of the different data collection techniques within the real contexts in order to take clearer methodological choices that can be more adjusted to the reality under evaluation.

References


