Evaluation Research in the 21st Century: a more and more relevant tool for the educational and social development

La investigación evaluativa en el Siglo XXI: un instrumento para el desarrollo educativo y social cada vez más relevante

Escudero, Tomás
Universidad de Zaragoza

Resumen

Tras una amplia revisión de las recientes publicaciones alrededor del tema, se analiza y valora la actual situación de la investigación evaluativa, como instrumento estratégico para la toma de decisiones de desarrollo y mejora de la sociedad y de la calidad de vida de los ciudadanos, en ámbitos diversos como la educación, la sanidad, la economía, la cultura, la protección social, las políticas públicas, etc. Se describe y fundamenta la identidad científica de la investigación evaluativa actual, incidiendo en su carácter transdisciplinar, en el auge de la evaluación de organizaciones e instituciones, en su apoyo en metodologías diversas y en la importancia de las estrategias participativas. Se destaca también la utilidad y el uso apropiado de las evaluaciones como objetivo prioritario de este tipo de investigación, apoyándose siempre en principios y normas éticas y de calidad científica y los correspondientes estudios metaevaluativos.

Palabras clave:

Investigación evaluativa, Desarrollo social, Disciplina transversal, Metodologías diversas, Estrategias participativas, Utilidad y uso de la evaluación, Normas ético-científicas, Metaevaluación

Abstract

After a wide revision of the recent publications about the subject, the present situation of evaluation research is analysed as a strategic tool for the decision making of the development and improvement of society and citizens' quality of life in varied sectors such as education, health, economy, culture, social protection, public policies, etc. The scientific identity of the evaluation research is described and founded upon, stressing its transdisciplinary character, the rise of the evaluation of organizations and institutions, and the use of different methodologies and the importance of participative strategies. Also outlined is the utility and appropriate use of the evaluations as priority object of this kind of research, relying always on principles and ethical rules of scientific quality and on the corresponding meta-evaluative studies.

Keywords:

Evaluation research, Social development, Transdisciplinary discipline, Different methodologies, Participative strategies, Utility and use of the evaluation, Ethical-scientific rules, Meta-evaluation.

In the past few decades, evaluation research has become a primary focus of investigative activity in every scientific field, especially in fields related to social politics and programs in general and in education in particular. The number of researchers who work in this field has kept rising and they incorporate their principles, criteria and methodological strategies in their proposals.

Evaluation research, also known in the specialized literature as evaluation, is kept in

Autor de contacto / Corresponding author

Escudero, Tomás. Universidad de Zaragoza. Instituto de Ciencias de la Educación. C/Pedro Cerbuna, 12. 50009 - Zaragoza (España). tescuder@unizar.es

constant evolution, but it had its central period of growth in the second half of the twentieth century, primarily in response to the need to analize, assess and improve the politics, actions, plans, programs, institutions and systems supported by modern societies in order to grow and improve the quality of life for their citizens (Escudero, 2011).

During the aforementioned period, governments and other organizations experienced an increased drive for ambitious plans and programs for change and social reform related to education, health, culture, social security, their own scientific research, etc., which needed to be analyzed and assessed profoundly to better understand their function, efficacy, efficiency, and in order to make decisions concerning their continuity and transformation. Therefore, evaluation research is established as a strategic mechanism for just and equitable social change (Schawndt, 2002, Cook, 2015) and a necessary ally to responsible social politics to optimize their actions and decision. A clear example of this orientation towards social reform can be found in the number of Summer 2015 in the important journal ¨New Directions for Evaluation¨, dedicated solely to the evaluation of social justice programs and the resolution of problems of disadvantaged people.What therefore logically resulted was the approach of Thomas and Madison (2010) proposing and defending the integration of the discipline of social justice in the programs dedicated to evaluation, programs which, as indicated by La Velle and Donaldson (2010), have experienced an incredible growth in the last few years. In this context, it is logical, like it is done today (McClintock, 2003;Calderon, 2004), to define the evaluator as an investigative agent of change.

In line with this reality we encounter with the declaration of 2015 as the international year of evaluation in the Third International Conference of National Politics of Evaluation, which took place in Sao Paulo between September 29th and October 2nd of 2013. The proposal came from ¨EvalPartners¨, an international global movement of evaluators to fortify the national capabilities in this terrain, with the declared objective of utilizing evaluation to improve the lives of the people, influencing in the improvement of public policies.

This great event of social programs in the development of evaluation research has sparked talk of the program evaluation as the name of this type of research (Smith, 2010). Nevertheless, the fact that the objective of the evaluations has futher amplified the programs, for example, politics in general, institutions, systems, organizations, resources, individuals, etc. has served to coin and consolidate the more global and suitable name of evaluation research.

This research acquires a great relevance and development in fields such as education, health, social intervention and welfare for the great importance they have for the functionality and development for modern societies, but we must specify that the scope of this type of research is transverse and influences equally the other disciplines and fields of social relevance such as economy, culture, science, research, sports, information, urban development, etc. That is, for example, taking advantage of the beginning of the journal "Revista de Evaluación de Programas y Políticas Públicas", Muñoz & col.(2013), accompanying a thematic revision in the economic sphere, defend the growing necessity for this type of research in the subject of public politics in Spain and the European Union as a whole. For these reasons, the well-known proposal of Scriven (2003), of transdisciplinary character for evaluation research, takes on more and more relevance, as its field of analysis extends not only to diverse objects of evaluation, but also to many disciplines.

Identity and the scientific basis of present evaluation research

When we talk about evaluation research, we are dealing with a complex concept with distinct positions which has been and will

continue to be a systematic process of evolution. In the field of education, for example, it is very interesting to observe this process of development from when evaluation was identified with the measurement of personal characteristics up until actual evaluation research (Escudero, 2003).

For these reasons, attempting to define evaluation research with a simple phrase, setting its functions in a precise way other approaches or research fields proves very difficult because there exist many zones of contact which blur that distinction. Only by highlighting the most specific elements of evaluation research are we able to establish a valid, definitive and useful outline to frame scientific work in this field of growing development.

The first thing to be pointed out is a distinctive essential element: that evaluation research is framed in a context of change, and more concretely, in a context of social change. It is a reactive focus of research, inside a paradigmatic school of thought of a critical base, which seeks to offer solutions to concrete problems from a pragmatic and contextualized position, from which a situational notion of social development is defended, instead of some fixed standard ones as the base and support for social emancipatory development (Bredo, 2006; Escudero, 2009).

This paradigmatic focus oriented towards practical resolution of problems has, alongside hermeneutics, its support in social constructionism, but also uses quantitative methods such as the experimental, as well as indicators of function which support the design of efficient interventions. Following the principles of pragmatism and contextualism, programs and projects which <<function well>> inside a particular real context are looked for, not being the fundamental proposition, description and theoretical construction, but decision making for management and planification, the solution of problems and the practical construction of programs inside of the applied research tradition (Escudero, 2011). This reality makes

it so that in evaluation research, the complete investigation of processes and results and of their significances, needs the use of diverse and mixed approaches (Sondergeld & Koskey, 2011).

We see that evaluation research is always approached in a real context to the service of social politics, more specifically, to the service of change and social development and that it therefore is an intermediary instrument helping other disciplines and fields such as education, public health, culture, social wellbeing, etc., so it looks directly for the objective of offering the best options for action from all alternative possibilities. It deals with helping to resolve problems which appear in these fields. Evaluation research, in summation, is moved in the context of problem resolution. And in this sense, the evaluator is seen as obliged to use theoretical knowledge in the resolution of practical problems, playing the incredibly important role of facilitator in the integration of basic research and practice (Urban & Trochim, 2009).

Another important identificative element of evaluation research is that it tackles the analysis of all types of information, qualitative as well as quantitative, in many cases with great volumes of data from diverse perspectives supporting itself in multiple and flexible methodologies of reseach (Maxcy, 2003; Escudero, 2009).

In this context it is also necessary to emphasize what it means for expansion of evaluation to a new perspective, the evaluation for development or empowerment evaluation, promoted by Fetterman as a focus, based in democracy, on the participation of those involved in the evaluated program to promote the autonomy of they themselves in the resolution of their own problems (Fetterman, 2001; Fetterman, Kaftarian & Wandersman, 2015).

At the moment we also need a sufficiently active current of evaluators who are assigned to the vision called evaluation based in theory and who defend the necessity of the evaluator to count previously, and analyze critically,

with the logical model (Renger & Hurley, 2006: Wasserman, 2010) or the theoretical one which serves as the base for the program, project, politics and interventions to evaluate (Christie, 2003; Donaldson and Gooler, 2003; Donaldson & Scriven, 2003; Mark, 2003; Weiss, 2004; Donaldson & Lipsey, 2006; Donaldson, 2007).

This vision gathers diverse approaches, but defends the contextualization and specific foundation of each research, in accordance with the particular characteristics of the problem, program or object of evaluation, of those involved in it and the more effective possibilities of action. The adaptability, flexibility, integration of foci and methodological eclecticism are definitive elements of this approximation of evaluation research, that part of the development of a theoretical basis of the program or intervention, from distinct sources and procedures which form the base for priority questions of evaluation to respond.

Another prevailing approach in the last few years is basing the evaluations on the theoretical perspective of the complexity of the objects and programs evaluated, so much as to define the objectives as the methodological approximations. Walton (2014) offers us a revision of the theoretical basis of this approach, of its principle characteristics and the consequences of its application in the last decade which converge in great measure with other approximations emphasized in this work, highlighting aspects such as the use of multiple and mixed methods and participative methodologies.

Urban, Hargraves and Trochim (2014) defend evaluation research as a process evolving in distinct phases in correspondence with the evolutionary process in the evaluated programs themselves, in a way in which each phase lays out diverse objectives and methodologies. More concretely, they talk about four phases: a first initial phase of the program in which the answer to the process is evaluated by the participants; a second phase of development of the program where the

changes produced in the individuals, context, etc. are evaluated; a third where the stability of the program is analyzed, controlling and evaluating its efficacy in comparison with other groups; and finally a fourth phase of dissemination of the program in which the capacity for generalization is evaluated.

A comprehensive evaluation of the analyzed proposals drives us to a possible definition of current evaluation research as a ¨type of applied research which has an effect on social objects, systems, plans, programs, participants, institutions, agents, resources, etc. which analyze and judge their static and dynamic quality in line with criteria and multiple rigorous scientific standards, both internal and external, with the obligation of suggesting alternative actions for they themselves for different propositions such as planning, improvement, certification, accreditation, fiscalization, diagnosis, reform, penalization, incentivation, etc.¨(Escudero, 2006, p. 271)..

Consolidation as a cross-discipline

In the transdisciplinary perspective, evaluation research, maintaining its idiosyncrasy as an autonomous discipline, provides basic tools and assistance to other very diverse disciplines (Scriven, 2003). We can therefore talk of two components: the conjunction of camps of application of evaluation and the content of the discipline itself. Something close to what occurs in disciplines such as statistics and measurement. Definitively in practice, evaluation research is a discipline which includes its own content and those of many other disciplines which are subject to study. And as Jacob (2008) shows us, the evaluations require the extensive collaboration between specialists in distinct fields and subjects of knowledge.

Strengthening this basis, relevant investigators such as Stufflebean (2001a) have defended and imposed programs in an academic formation in evaluation research of interdisciplinary character as an important medium to facilitate, strengthen and extend the evaluative practice.

But the best demonstration of the transdisciplinary focus of evaluation research having totally taken over today, are the foci and contents which are published in scientific journals promoted by the National and International Associations of Researchers and in the relevant journals on evaluation. By way of example we can cite the journals "African Journal of Evaluation¨, ¨Canadian Journal of Program Evaluation¨ and ¨New Directions for Evaluation¨ promoted by the African, Canadian, and American Associations of Evaluation, respectively. Also we can cite journals of great international relevance as "Evaluation and Program Planning" and "Evaluation", which show in an expressive manner their interdisciplinary character in their editorial line.

Logically, the scientific journals which form the basis for the transversality of evaluation as a discipline, which are many more than the aforementioned examples, coexist perfectly with many others which are centered on evaluation in concrete fields of great social relevance such as are education, public health, economy, social politics, etc., which reinforce the transdisciplinary character with which we currently understand evaluation research.

The great development of institutional research

Between the milestones of recent fortification of investigative research, we cannot forget to talk about the heights reached in the last few decades by institutional evaluation, in a very accused way in the field of education and its institutions in all levels of obligatory character and in the great majority of developed university systems in the international context (Escudero, 2007). Taking north American evaluations as an initial reference point, this process has been extending gradually in the majority of university systems, inside its specific characteristics, in a very accused way in Europe and Latin America (Aguilar, 2001).

A good example of the peak of evaluation research in this field can be found in the XIII General Assembly of ALAFEC (Latin American Association of Faculties and Schools of Accounts and Administration) celebrated in Buenos Aires in October of 2012, in which a presentation offered an analysis of investigative research and its perspective in the evaluation of institutes of higher education (Mira et al., 2012).

Institutional research is synthesized and centers on the great majority of objects of interests and in fields of work of evaluation research, politics, programs, organization and governance systems, internal and external relations, personal and material resources, infrastructures, etc., picking up and integrating also the principles and methods defended from the field of institutional quality and the processes of its control, assurance and improvement

This reality has imposed and fortified the use of certain methodologies in this field of research such as, for example, the development of auto-informs, the use of indicators of efficiency and the contrasts between internal and external visions (Escudero, 2002). Since the evaluation of institutions and organizations, not only educational ones but also those of other fields such as health (Yusa, Hynie & Mitchell, 2016), it has strengthened extraordinarily in evaluation research the use of internal and external evaluations such as necessary and complementary instruments to secure the access to all types of relevant information, to reach a complete diagnosis of the problems to resolve and elements to improve and reinforce the credibility and acceptance of evaluation and its proposals, to facing the decisions to improve. Pérez Juste (2002) also highlights the importance of both approximations, internal and external, to strengthen the quality of education from the evaluation of its programs.

A great example of this outlook is one offered by the recent publication of the European Commission (Eurydice 2015) in which they analyze how evaluate the quality of its schools thirty two European countries, comparing their distinct approaches and

structures and dealing in a special way the internal and external systems of evaluation, their procedures and types of evaluators and the coordination and use of their results.

Centering on the field of higher education, Nicoletti (2013) offers us an interesting analysis of the current situation of this type of evaluation research, of its objectives and principal methodological foundations. Likewise, in a recent doctoral thesis at the University of Zaragoza, Professor Javier Paricio (2015) offers a complete analysis comparing the conceptions of quality and quality assurance systems which exist in the whole of European universities, at the same time designing a methodological model of multidimensional analysis.

This impulse, that the cited references leave clear, has strengthened the approaches and organisms of external evaluation, creating commissions, agencies, intermediary offices and nets of regional, national and international character, of support for evaluations and processes of accreditation and certification of quality. Claverie, González and Pérez (2008), for example, describe and analyze critically the model of evaluation of the quality of higher education in Argentina, sustained by the coordination of the CONEAU (National Commission of University Evaluation and Accreditation). Reina Ferrández (2008) offers us an analysis comparing the programs of institutional university auditing in Spain and Great Britain, emphasizing the influence of organizations such as the Spanish National Agenda of Evaluation of Quality and Accreditation (ANECA) and the British QAA (Quality Assurance Agency for Higher Education).

Diversity and flexibility of methodology

Just like in the methodological approximations, the approaches of evaluation research can be very diverse, including the evaluative approaches free of objectives, that, although, are not very habitual (Youker, Ingraham & Bayer, 2014), can be very

valuables in some process of external evaluation.

Expósito, Olmedo and Fernandez-Cano (2004) offer us an example of this diversity in their revision of employed methods in the evaluation of programs in Spain. Christie and Fleisher (2010) also provide proof of this varied panorama in their analysis of publications in specialized American journals.

Diversity and flexibility in the methodology employed in evaluation research is something forced by its own transdisciplinary nature and the paradigmatic basis of this type of research, analyzed in a previous section, and by the multiplicity of fields of work, goals, objects and contexts of analysis that the evaluator is obligated to work, to develop their various diagnoses and proposals for action for the different ones implicated, responsible for and affected by the research.

This pluri-reality makes the vast majority of methods and research procedures useful for evaluation research, but, logically, its applicability is conditioned and defined in each case by the type of problem, appearance or theme that is being investigated. The same applies for procedures and sources for collecting information, a key aspect in many evaluation studies that they must rely on variables, information, opinions, etc., very different in terms of its type, diversity of those involved and possibilities of collection. This complexity that the evaluator faces, makes the triangulation of methods, techniques, sources and procedures for collection of information and the combined use of different methodologies (mixing methods) to acquire special relevance in evaluation research, as a means of strengthening and validation of the diagnosis and results (Perassi, 2009; White, 2013; Betzner, Lawrenz & Thao, 2016; Makrakis & Kostoulas-Makrakis, 2016).

In the past few years, it has already been investigated about the gathering of data in surveys through internet, something that, naturally, has opened new and interesting perspectives for the evaluators. The number of fall 2007 of the journal New Directions for

Evaluation, for example, is dedicated solely to works on the use of surveys (online) in evaluation, influencing the strengths and limits, as wells as the conditions that are favourable for utilization.

This openness and methodological adaptability that we are talking about, is noted perfectly in the publications of research studies in specialized scientific journals, as it is very habitual that before the research of topics, programs, plans, etc., novel or complex, the research team designs, establishes, and describes previously a model and methodological procedures adapted to their problem or specific topic, that afterwards applies in the development of the research project.

Participative Strategies

For several decades, and for many and varied reasons, theoretical as well as practical, the experts have been worrying and defending the involvement of the stakeholders in the evaluation of their programs (Rodriguez-Campos, 2012). One of those reasons is to promote the use of the results of the evaluation in the various processes of decision-making that can affect a diverse group of those involved (Johnson et al., 2009; Patton, 2012). The basic idea is that if they participate and collaborate with the evaluation, it is much easier to assume its results and suggestions of action and, consequently, accept and promote the necessary changes, although they may be expensive and entail some sacrifices. With the participation of those affected, the evaluation, overall, increases their viability and potential usefulness. In addition, reinforces the principle of co-responsibility among those involved in the problems detected and their possible decisions and formulas of horizontal management in programs and institutions are strengthened.

One of the keys to produce this type of participation is that there is a culture that exists in the institutions, this is, a whole of basic assumptions, norms, values and cultural views shared by its members, which affect its functioning, integrating what we can call evaluative culture (Scheerens, 2004). In other words, the belief that the organization learns and improves beginning with critical reflection and systematic assessment carried out by its members is necessary. One of the most effective ways to overcome many of the inhibitions of, for example, teachers before the assessment, are the processes of critical reflection on their own work and institutional functioning.

The collaboration and participation formulas in the evaluations are multiple and varied in their intensity and implication (Taut, 2008), although all of them are sustained in some way in organizational models known generically as democratic. Sometimes with a single phase process in which everyone participates, and others with several phases, with varying degrees of involvement of each other. Askew and collaborators (2012) defend the need to reconcile collaborative evaluation techniques with responders focuses to diverse cultures of those involved, to avoid discrimination against minority groups. Geist (2010), for example, presents and explains the use of the Delphi method as a procedure to promote the participation of the stakeholders in the evaluation process. Rodriguez-Campos (2012a) offers an interesting review of recent advances in collaborative evaluation and its foundation, with its strengths and contributions, additionally proposing a model focus utilizing the same. At the same time, it continues the empirical research about participating in evaluations and the relationship between involvement and use in the evaluation processes and formulas of more effective collaboration in different contexts and types of evaluation (Brandon & Fukunaga, 2014 ; Daigneault, 2014 Roseland et al, 2015;. & Chouinard & Milley, 2016).

The idea of democratic evaluation has been developed with diverse focuses and aproximations (Ryan, 2004), that reflect the organizacional aspects and methods, in the generation of knowledge, in the role of evaluators and the practices of communication

between those involved. An extraordinary extended formula in institucional evaluation, for example, is the one of combined phases of self-evaluation and external evaluation, seeking the integration of results and complementarity of the positions and interests of different audiences. Some authors go even further (Christie et al, 2004), arguing that both evaluations can be carried out by a single team, consisting of internal and external agents, working together every time in collaboration.

In the field of health, for example, faced with making decisions, Adelson and collaborators (2003), behind a wide revision of the topic, highlighted the importance and usefulness of the use of deliberative methods, with the participation of diverse groups involved. Also in this field, a group of Austrian investigators, pertaining to the Ludwig Boltzmann Institute of Research in the Promotion of Health in Vienna (Nitsch et al, 2013) offer an interesting review of nearly fifty evaluative scientific papers, of health promotion programs, analyzing the patterns and levels of participation the stakeholders, as well as the strengths and weaknesses of such processes.

The strengthening of the collaborative strategies in the last years, in the practice of evaluation research, is unquestionably, becoming at the present moment something of habitual use in projects of evaluation (Preskill & Boyle, 2008). The great mentor of empowerment evaluation (Fetterman, 2001a), comes to tell us that in the 21st century, evaluation has become a collaborative process between those involved in it.

Fitzpatrick (2012) offers a review of how it has spread the use of collaborative evaluation since the last century, as well as the different models and approximations of cooperation contained in the scientific literature, their strengths and limitations, and how his application is conditioned by the nature of the problem evaluated, the context in which it takes place and the types of those involved.

Utility and use of evaluation

The analysis of scientific literature confirms to us that, in the present moment, few things are as in accordance to experts in evaluation research as the fact that this is done with the intention of being useful and influence in the making of subsequent decisions. In fact, the evaluators themselves devote great efforts to get their suggestions to be taken into account and properly applied (Christie, 2007). However, the use of evaluation results, for many reasons, is always a matter under controversy (Escudero, 2013).

To strengthen appropriate use of evaluations, it is important to analyze the contexts of decision of potential users well and focus the study on aspects relevant to the task (Leviton, 2003; Ledermann, 2012, Neuman et al, 2013.). The development of evaluative culture in the organizations is another facilitator element of the appropriate use of the evaluation processes (Cousins et al. 2014).

Reflecting on the practice, Grasso (2003) tells us that, on beforehand, we will never have complete assurance that an evaluation be used, but this possibility is increased if the conditions are optimized for it. Specifically, an evaluation is more useful and more usable, if the evaluator has previously responded to the following three questions:

- Who will utilize the evaluation?
- What will they need for the evaluation?
- When will they need the information?

In short, it seems necessary to determine which are the audiences, how they can arrive at the information, what are the priorities of each one of them and with what speed do they arrive at them. Therefore, the contact and consultations with various audiences prior to designing and carrying out the assessment is fundamental, to promote a comprehensible communication and at the proper time of the results of the evaluation and suggestions of action. Today, based on this type of principles, the one develops and applies different systems of audit of the information produced by the

evaluations (Schwarth & Mayne, 2005), to ensure its quality and, at the same time, which is available when needed, that responds to the demands of the problem and that is understandable for those who need it and demand it.

Also exists a trend of evaluators (Kirkhart, 2000; Henry & Mark, 2003) that defend the concept of the use of evaluation, understood as the direct utilization and focalization, that has remained narrow, so you must give way to the concept of influence, such as the idea of consequences more open and spacious, and not necessarily bounded in advance. The influence speaks also of the possible consequences and not only of the sought after, that is, of the capacity to produce changes by intangible or indirect means.

Delving deeper into the previous line of thought, Henry (2003) emphasizes the great value of the evaluation in democracies, as an instrument for social improvement, although, according to his opinion, evaluators have not known to exploit their full potential to influence in this sense, probably because research on the use of the evaluation has focused on its immediate use.

In practice, something that extraordinarily conditions the use and influence of evaluation, is the way that one communicates or informs. It is important to highlight that good evaluation research not only relies on good analytical techniques, like any other type of social research, because it must combine these with well developed communication techniques. In other words, it is not enough that the evaluation is well done, in addition, it must be well told, not only to specialists, but to all those involved and affected by it, to all audiences. It must be added that there are specialists (Perrin, 2001), that put much emphasis on convenience that evaluators use the more simple methods that can respond to the questions posed, because this not only has implications on cost, speed and acceptance of the evaluation, but also in its use.

Therefore, there are different approaches on how one should be informed of evaluations with views to optimise its visibility and, as a result, encourage their utilization. Based on a common report of the process of evaluation and its results, one can enter sub-paragraphs, appendices, specific suggestions, etc., aimed at different audiences involved. Another option is the one of producing separate reports according to the audiences, although this is a less common formula.hat is critical in all cases, is the adequacy of the technical language when presenting the results that, in evaluation research, most of the time are the product of some kind of statistical treatment, which is not familiar to many of the potential receptors of such results. The evaluator should be able to present his/her results supported by the conventional statistical jargon and, in addition, do so in a way that has significance (May, 2004) for distinct audiences.

This author tells us that a statistic result has significance for a determined audience, when it is:

- .*Comprehensible*, that is, that is understood in an easy manner by the majority of the people, with the minimum possible assumptions about the statistical preparation of the audience and avoiding the statistical jargon.

- .*Interpretable*, this is, utilizing units of measure familiar or easily explained

- .*Comparable*, that is, permitting the observation of differences with other situations, over time, other programs, other centres, etc.

The third criterion, the comparability, May tells us that it is less critical than the first two, because the conditions of need not always are given or created for comparison. When this happens, the criterion enriches a lot the previous values.

Ethics and use of the evaluation

Some years ago, House (2008) warned against the danger of infecting and endangering seriously the field of evaluation research with bad practices coming from a field having a good image in relation to the

design of their evaluations. He was referring specifically to the processes of drugs evaluation which, according to this author, with certain frequency are biased to overly emphasize the positive effects of the evaluated drug and to hide the negative ones, even manipulating the research designs in order to achieve the wished results. According to this author, such a relevant and important matter, from the social point of view, shouldn't only be entrusted nor to the beneficiary companies nor to the market. Regarding this particular one, and all social matters, public authorities and the own evaluators have to fight directly against the pseudo-evaluations and against bad evaluations and misevaluations. The political context in which an evaluation is carried out can also determine the evaluation itself and the consequent making of decisions. Therefore, strengthening the independence of the evaluators is still a worrying and study matter for specialists on evaluation research (Azzam & Levine, 2015).

As Pinkerton et al. (2002) indicate, the ethical theme is critical in studies of cost-effectiveness, cost-benefit and cost-utility of social programs that are very valuable, then they enter the scene of specific decisions of ideological basis, which condition the economic contribution with an orientation or other, benefiting some involved or others.

The same evaluator may be a source of conflict when their personal values are very opposite, or are very favorable, towards the object of evaluation. For example, someone personally involved in the defense or in criticism of a particular educational program, it is very difficult than can evaluate without generating doubts about their work.

Finally, the object of the evaluation itself can be conflicting and may pose ethical problems of the researcher's nature. A common case is when there are supporters and detractors of the program evaluated. Then, the evaluator should give audience to all parties to the conflict and articulate a harmonious system of convergence of criteria and drawing conclusions and suggestions (Abma, 2000).

A second type of conflict situation happens when social interventions that do not have the support of research are evaluated. In such a case, the evaluator must be extremely rigorous in analyzing results of it, contrasting the effects in different contexts and comparing such reality that is given in situations without intervention. In summary, in this circumstance the evaluator must record his condition of experimental researcher.

Crossing dimensions (use-not use) and (right-misuse use), Cousins (2004) establishes four quadrants or categories in relation to the use of evaluations. These categories are the following:

1. Ideal Utilization, that includes instrumental, conceptual and previously discussed persuasion or support

2. Bad Utilization ("misuse"), good for incompetency or incomprension of the results of the evaluation, good for the intentional manipulation of them

3. No utilization unjustifiable, irrespective of an abusive manner or removing the results of the evaluation.

4. No utilization justified, when reasons of a diverse type appear that advise to not take in account the results of the evaluation

Although it is not easy, for their own situation in the process of decision making, the evaluator should not only try to prevent their results from not being used unjustifiably (category 3), unfortunately a very common situation, but must do everything possible to prevent misuse of their work (category 2), undoubtedly the most dangerous problem.

Logically, evaluation research that is truly useful and rigorous, is done outside of these situations that endanger ethics in evaluations and their uses, which, as we have analyzed, it is fundamental for the development and credibility of a type of applied research increasingly relevant to ensure and enhance the quality of our overall social policies and, ultimately, of our society. An investigation that, as Chelinsky (2008) points out, should seek at all times the adjustment between the

necessary and scientific independence and the just demands of a democratic society.

For these reasons, the professional associations of evaluators have developed documents where they establish the ethical principles for the development of evaluations. An example is the "Guiding Principles for Evaluators" of the AEA (American Evaluation Association, 2008), aimed to promote ethical practice in the evaluation of programs, products, personnel and policies. The treaties´ principles are systematic inquiry, competence, integrity/honesty, respect for people and the responsibility with the general welfare and of the citizens.

Schweigert (2007) investigates also this terrain of ethics, indicating that justice is a priority in evaluative action, that must preserve itself with the collaboration of the evaluators and the whole of those involved and affected. This author tells of three dimensions in justice to preserve, first the ones called public justice, including obligations, principles and norms of the profession of evaluator, secondly the procedural justice, ensuring respect for the rights and treatment appropriate to all those involved and, finally, distributive justice, the search for equitable social development.

Metaevaluation and norms of quality

Metaevaluation, evaluation of evaluation itself, acquires more and more importance as a consequence of the development of evaluation research and of the expansion of evaluation projects in all terrains. It seems logical to think that the first thing one needs is confidence in something that is made precisely to have confidence when it comes to making decisions, to intervene or suggest something concerning a social project. It attempts to articulate the scientific obligation to reflect on what could be going well or poorly with regards to a determined plan (Escudero, Pino & Rodríguez, 2010).

An author of long standing tradition in this topic as Stufflebeam (2001b), in addition to indicating that metaevaluation should judge the merit, value and use of evaluation, is established as an element of scientific-professional guarantee; the evaluators themselves should make sure that evaluations are done well, that they improve systematically and that they provide a guarantee to their audiences. Therefore, metaevaluation is a profession obligation of evaluators, which connects to the ethics of revision and continuous self-questioning (Stake, 2006). Metaevaluation seeks to detect statistical bias, deviations in the interpretation of data, comprehension problems of certain aspects, etc., within a critical perspective that part of the idea that any evaluation can be improved.

Continuing to follow Stufflebeam (2001b), we can define metaevaluation as "the process of delineating, obtaining, and applying descriptive information and judgmental information—about the utility, feasibility, propriety, and accuracy of an evaluation and its systematic nature, competent conduct, integrity/honesty, respectfulness, and social responsibility—to guide the evaluation and/or report its strengths and weaknesses" (p. 185). This same author tells us that a good comparison to understand this concept is that of metaevaluation being in the field of evaluation and that auditorship is in the field of accountability which insures the external and independent scrutiny of reasonable use of financial resources.

The perspectives formative and summative also converge in metaevaluation. The former serves as a guide to help the evaluators prepare everything before the evaluation itself, while summative metaevaluation determines the merit, value, and use of evaluation once it is done. Both perspectives together determine the degree to which the evaluation is technically adequate, useful for making decisions, ethical in its influence on people and efficient regarding the use of resources. In the formative approximation the metaevaluator acts as a technical consultant and in the summative as a certified external agent.

The work in the metaevaluation can be accomplished along distinct model approaches (Escudero 2011). For an example, the revision of study is based fundamentally in the analysis of the evaluation report, of the processes and results; it is an essential procedure in the study of interpretive nature, fundamentally about qualitative data. The secondary analysis supposes the reanalysis of the original data, being a very powerful procedure if can be done correctly, so the act of arranging the data of an evaluation is a sign of its credibility. The quantitative integration or meta-analysis is a statistical procedure of a combination of results when there exist various studies about the same topic. We can also talk about internal and external metaevaluations.

Metaevaluations are supported in varied techniques, despite the fact that there exist some of growing utilization such as panels of revision, analysis of protocols, interactive seminars and the proper checklists for evaluation of distinct nature and content (Scriven, 2000; Stufflebeam, 2000). Going deeper into these ideas, Vanhoof and Van Petegen (2010) point out that metaevaluation requires multiple criteria, multiple methods, and multiple people.

There are diverse criteria offered to judge the quality of evaluations, generally tied to merit, value, utility, efficiency of design, etc. It is generally talked about the criteria of: a) rigor (precise methods correctly applied), b) utility (value, merit, importance of results) and c) efficiency and effectiveness of the design and the evaluative process. Ethical criteria are also discussed, in general as protection of the rights of the people implicated in the evaluation, preserving the truth of what is evaluated in response to certain distorted pressures.

In relation to the development of metaevaluation in a context of professional growth of evaluation research is the establishment of norms under those which develop evaluation with visions of securing its quality and avoiding the maximum number of potential defects. For example, during the last few decades, norms of educational research and measurement have been elaborated by distinct scientific associations such as the American Psychological Association (APA), the American Educational Research Association (AERA) and the National Council on Measurement in Education (NCME).

But no one can escape that the academic and social prestige that evaluation research as a discipline and the work of professional evaluators has come to acquire and growing the last few decades, is in good measure marked by the leadership and responsibility of the ¨Joint Committee on Standards for Educational Evaluation¨, created by the three aforementioned scientific associations, which started to work in 1975. Nowadays, this committee is backed by more than fifteen academic-professional organizations, principally American and Canadian, from the field of educational evaluation (Joint Committee, 2003), and its major support has been the production and establishment of some ethical-scientific norms of certain standards to regulate the educational evaluation in the United States and Canada. Nevertheless, its application has gone much further than the geographical space alluded to and the field of education, although they have provoked some problems of adaptation in certain contexts (Stufflebeam, 2004).

In the African field, for example, the ¨African Evaluation Associations and Networks¨ has adopted these norms, producing a combination of 30 criteria of quality control which are called ¨African Evaluation Guidelines¨ (Patel 2002a; Patel, 2002b).

The basic objective of the Joint Committee and of its norms, which are understood as dynamic and which should revise and adjust themselves periodically, is one of promoting conscience concerning the necessity of the evaluations being solid, of quality, useful and which respond to the needs of evaluation. To obtain this, the Committee establishes norms of the following four types: of utility, of feasibility, of legitimacy and of precision.

It is very important to observe the order in which the distinct norms are considered, as they act as a key or previous condition for the following. In particular, in laying out an evaluation project, the first condition there is to secure is that it is useful, that is, that it makes sense to carry out in order to improve the system, to make summative decisions, to plan, etc. In other words, the Joint Committee comes to suggest in some measure that an evaluation should be useful to start with, and if this is not the case, it doesn't make sense to consider.

Ensuring the utility of an evaluation is when one needs to analyze if it is viable, if it is able to be done, if it is legitimate, if it doesn´t violate principles of legitimacy, the rights of the implicated, etc., and finally, it is analyzed, designed and carried out with the maximum scientific rigor following the most adequate and powerful principles and methodologies.

Nowadays, the bases for this metaevaluative process, not forgetting the natural context of its definition being the United States and Canada, with which its application in other contexts requires contextualization (Jang, 2000), is in the alluded to standards of evaluation of the Joint Committee and in the principle guides of the American Evaluation Association, defined for all fields, not only that of education. Without these ethical-scientific principles which regulate the work of evaluators, it is indubitable that evaluation research would not have reached the prestige and relevance that presently has (Escudero 2011).

It seems indubitable that the great importance and growing influence that evaluative investigation has had in the last few decades in education and other social fields would have been more questioned if the professional associations more directly implied had not have established this ethical-scientific norms to regulate the work of evaluators.

Reflection to finish

In this work we present and appreciate the consolidation of evaluation research as a cross-discipline of great strength, especially in the field of social sciences, as well as its growing utilization as a field of research in many fields of knowledge and social development.

Based on the current definition of evaluation research, we have emphasized the its function in the service of social change and, especially, in the service of social improvement. We have seen that the evaluation is an essential element of planning, management and decision-making, with the major objective of improving society and, more specifically, the quality of life of our citizens, but we know that we are talking about something complex, dynamic, multidimensional, needing contextualization, which requires analysis and rigorous and restful treatments.

Analysis of the utility and the use of evaluation constitutes a central block in this work because it is about one of the biggest problems that the evaluators as professionals encounter today, because it is not always given the attention to the results and suggestions of its works and because the future of the evaluation research as an academic discipline and as a relevant profession, is directly related and conditioned as a good use and a correct utilization of the evaluations and of its results.

If the evaluations are not useful for change and social improvement and if are not well used, our thesis is that they are one effort of little relevance and, therefore, are very little justifiable. But we know and argue that if used well, they are tools of a special strategic relevance. In the work we present different models of use, use of results, as well as distinct formulas and strategies enabling this utilization. Also we entered the debate that has recently arisen around the concepts of use and influence. We also make a clear defense of participatory evaluation strategies as a stimulus and guarantee of use and good utilization of their results, pointing out the distinct formulas of participation that arise or may arise.

The work also provides a wake-up call about the ethical utilization of evaluations,

with the presentation of the principles that must be followed, and highlights, likewise, the relevance of the meta-evaluations as instruments of guarantee of quality of the evaluations.

As a final idea, we want to point out the growing importance of evaluation research in all scientific fields, university departments, institutes and other organisms of research, etc. coexisting with other research focuses of basic and applied research, both in the field of training and in the development of projects

References

Abelson, J., Forest, P-G., Eyles, J., Smith, P., Martin, E., & Gauvin, F.P. (2003). Deliberations about deliberative methods: issues in the design and evaluation of public participation processes. *Social Science & Medicine 57*, 239–251. DOI: http://dx.doi.org/10.1016/S0277-9536(02)00343-X

Abma, T. A. (2000). Stakeholder conflict: a case study. *Evaluation and Program Planning* 23, 199-210. DOI: http://dx.doi.org/10.1016/S0149-7189(00)00006-9

Aguilar, M. (2001). La evaluación institucional de las universidades. Tendencias y desafíos. *Revista de Ciencias Sociales* (Cr), II-III, 93-92, 23-34.

American Evaluation Association (2008). Guiding Principles for Evaluators. *American Journal of Evaluation*, *29* (4), 397-398.

Askew, K., Green Beverly, M. & Jay. M. L. (2012). Aligning collaborative and culturally responsive evaluation approaches. *Evaluation and Program Planning, 35*, 552–557. DOI: http://dx.doi.org/10.1016/j.evalprogplan.2011.12.011

Azzam, T. & Levine, B. (2015). Politics in evaluation: Politically responsive evaluation in high stakes environments *Evaluation and Program Planning, 53*, 44-56. DOI: http://dx.doi.org/10.1016/j.evalprogplan.2015.07.002

Betzner, A., Lawrenz, F. P. & Thao. M. (2016). Examining mixing methods in an evaluation of a smoking cessation program. *Evaluation and Program Planning, 54*, 94-101. DOI: http://dx.doi.org/10.1016/j.evalprogplan.2015.06.004

Brandon, P. R. & Fukunaga, L. L. (2014). The state of the empirical research literature on stakeholder involvement in program evaluation. *American Journal of Evaluation*, *35* (1), 26–44. DOI: http://dx.doi.org/10.1177/1098214013503699

Bredo, E. (2006). Philosophies of Educational Research. In J. L. Green, G. Camilli & P. B. Elmore, *Handbook of Complementary Methods in Education Research.* London: Lawrence Erlbaum Associates Publishers AERA 3-31.

Calderon, A. J. (2004). Institutional Research at RMIT. A case study. Ponencia presentada en el *26th EAIR Forum*, Barcelona, September 5-8.

Chelimsky, E. (2008). A Clash of Cultures: Improving the "Fit" Between Evaluative Independence and the Political Requeriments of a Democretic Society. *American Journal of Evaluation*, 29, 4, 400-415. DOI: http://dx.doi.org/10.1177/1098214008324465

Chouinard, J. A., & Milley. P. (2016). Mapping the spatial dimensions of participatory practice: A discussion of context in evaluation. *Evaluation and Program Planning,* 54, 1-10. http://dx.doi.org/10.1016/j.evalprogplan.2015.09.003

Christie, C. A. (2003). The practice-theory relayionship in evaluation. *New Directions for Program Evaluation*, 97, Jossey-Bass, San Francisco, Ca.

Christie, C. A. (2007). Reported influence of evaluation data on decision makers'actions: An empirical examination. *American Journal of Evaluation*, 28, 1, 8–25. DOI: http://dx.doi.org/10.1177/1098214006298065

Christie, C. A. & Fleischer, D. N. (2010). Insight Into Evaluation Practice: A Content Analysis of Designs and Methods Used in

Evaluation Studies Published in North American Evaluation-Focused Journals. *American Journal of Evaluation*, *31*(3), 326-346. DOI: http://dx.doi.org/10.1177/1098214010369170

Christie, C. A., Ross, R. M. & Klein, B. M. (2004). Moving toward collaboration by creating a participatory internal-external evaluation team: a case study. *Studies in Educational Evaluation*, *36*(2), 107-117. DOI: http://dx.doi.org/10.1016/j.stueduc.2004.06.002

Claverie, J., Gonzalez, G. & Perez, L. (2008), El Sistema de Evaluación de la Calidad de la Educación Superior en la Argentina: El Modelo de la CONEAU. Alcances y Límites para Pensar la Mejora. *Revista Iberoamericana de Evaluación Educativa*, *1*(2), 149-164.

Cook, J. R. (2015). Using Evaluation to Effect Social Change: Looking Through a Community Psychology Lens. *American Journal of Evaluation*, *28*(1), 107-117. DOI: http://dx.doi.org/10.1177/1098214014558504

Cousins, J. B. (2004). Commentary: Minimizing evaluation misuse as principled practice. *American Journal of Evaluation*, 25(3), 391-397. DOI: http://dx.doi.org/10.1177/10982140040250031 1

Cousins, J. B.; Goh, S.C.; Elliot, C.J. & Bourgeois, I. (2014). Framing the Capacity to Do and Use Evaluation. *New Directions for Evaluation*, *141* (1), 7-23. DOI: http://dx.doi.org/10.1002/ev.20076

Daigneault, P. (2014). Taking stock of four decades of quantitative research on stakeholder participation and evaluation use: A systematic map. *Evaluation and Program Planning, 45*, 171–181. DOI: http://dx.doi.org/10.1016/j.evalprogplan.2014.04.003

Donaldson, S. I. (2007). *Program theory-driven evaluation science*. Mahwah, NJ: Erlbaum

Donaldson, S. I. & Gooler, L. E. (2003). Theory-driven evaluation in action: lessons from a $20 million statewide Work and Health Initiative. *Evaluation and Program Planning,* 26, 355–366. DOI: http://dx.doi.org/10.1016/S0149-7189(03)00052-1

Donaldson, S. I. & Lipsey, M. W. (2006). Roles for theory in contemporary evaluation practice: Developing practical knowledge. In I. Shaw, J. C. Greene & M. M. Mark (Eds.), *The Handbook of Evaluation: Policies, programs, and Practices* (56-75*).* London: Sage

Donaldson, S. I. & Scriven, M. (2003). Diverse visions for evaluation in the new millennium. In S. I. Donaldson & M. Scriven (Eds.) *Evaluating social programs and problems: Vision for the new millenium* (pp. 3-16), Mahwah, NJ: Erlbaum.

Escudero, T. (2000). Evaluación de centros e instituciones educativas: las perspectivas del evaluador. In D. González, E. Hidalgo & J. Gutiérrez (Eds.), *Innovación en la escuela y mejora de la calidad educativa* (pp. 57-76). Granada: Grupo Editorial Universitario

Escudero, T. (2002). Evaluación institucional: algunos fundamentos y razones. En V. Álvarez & A. Lázaro, *Calidad de las Universidades y Orientación Universitaria* (pp. 103-138). Málaga: Ediciones Aljibe.

Escudero, T. (2003). From tests to current evaluative research. One century, the XXth, of intense development of evaluation in education. *RELIEVE*, 9(1). DOI: http://dx.doi.org/10.7203/relieve.9.1.4348

Escudero, T. (2005-2006). Claves identificativas de la investigación evaluativa: análisis desde la práctica. *Contextos Educativos. Revista de Educación, 8-9*, 179-199.

Escudero, T. (2006). Evaluación y mejora de la calidad en educación. In T. Escudero & A. D. Correa, *Investigación en innovación educativa: algunos ámbitos relevantes* (pp. 269-325). Madrid: La Muralla, S. A.

Escudero, T. (2007). Evaluación institucional de la calidad universitaria en España: Breve pero interesante historia. *Anuario de Pedagogía*, *9*, 103-115.

Escudero, T. (2009). Some relevant topics in educational evaluation research. In M. Asorey,

J. V García Esteve, M. Rañada, & J. Sesma, *Mathematical Physics and Field Theory. Julio Abad, in Memoriam*, (pp. 223-230). Prensas Universitarias de Zaragoza.

Escudero, T., Pino, J. L. & Rodríguez, C. (2010). Evaluación del profesorado universitario para incentivos individuales: revisión metaevaluativa. *Revista de Educación*, *351*, 513-537.

Escudero, T. (2011). *La construcción de la investigación evaluativa. El aporte desde la educación*. Prensas Universitarias-Universidad de Zaragoza.

Escudero, T. (2013). Utilidad y uso de las evaluaciones. Un asunto relevante. *Revista de evaluación educativa, 2* (1). Retrieved from http://revalue.mx/revista/index.php/revalue/issue/current

European Commission/EACEA/EURYDICE (2015). *Assuring Quality in Education: Policy and Approaches to School Evaluation in Europe,* Eurydice Report. Luxembourg: Publication Office of the European Union. Retrieved from http://eacea.ec.europa.eu/education/eurydice/

Exposito, J., Olmedo, E. & Fernandez-Cano, A. (2004). Patrones metodológicos en la investigación española sobre evaluación de programas educativos. *RELIEVE*, *10*(2). DOI: http://dx.doi.org/10.7203/relieve.10.2.4319

Ferrandez, R. (2008). Programas de Auditoría Institucional Universitaria. Comparación de la Propuesta Española con el Sistema Británico. *Revista Iberoamericana de Evaluación Educativa*, *(1)*1, 156-170.

Fetterman, D. M. (2001a). The Transformation of Evaluation into a Collaboration: A Vision of Evaluation in the 21st Century. *American Journal of Evaluation*, *(22)* 3, 381-385. DOI: http://dx.doi.org/10.1177/109821400102200315

Fetterman, D. M. (2001b). *Foundations of empowerment evaluation*. Thousand Oaks, CA: Sage.

Fetterman, D.M., Kaftarian, S.J. & Wandersman A. (Eds.) (2015). *Empowerment Evaluation: Knowledge and Tools for Self- Assessment, Evaluation Capacity Building, and Accountability* (2nd edition). Thousand Oaks, CA.: Sage Publications. DOI: http://dx.doi.org/10.1016/b978-0-08-097086-8.10572-0

Fitzpatrick, J. L. (2012). Commentary-Collaborative evaluation within the larger evaluation context, *Evaluation and Program Planning*, *35*, 558-563. DOI: http://dx.doi.org/10.1016/j.evalprogplan.2011.12.012

Geist, M. R. (2010). Using the Delphi method to engage stakeholders: A comparison of two studies. *Evaluation and Program Planning, 33*, 147-154. DOI: http://dx.doi.org/10.1016/j.evalprogplan.2009.06.006

Henry, G. T. (2003). Influential evaluations. *American Journal of evaluation*, *24* (4), 515-524. DOI: http://dx.doi.org/10.1177/109821400302400409

Henry, G. T. & Mark, M. M. (2003). Beyond use: Understanding evaluation´s influence on attitudes and actions, *American Journal of Evaluation*, *24* (3), 293-314. DOI: http://dx.doi.org/10.1177/109821400302400302

House, E. R. (2008). Blowback. Consequences of Evaluation for Evaluation. *American Journal of Evaluation*, *29*(4), 416-426. DOI: http://dx.doi.org/10.1177/1098214008322640

Jacob, S. (2008). Cross-Disciplinarization: A New Talisman for Evaluation? *American Journal of Evaluation*, *29* (2), 175-194. DOI: http://dx.doi.org/10.1177/1098214008316655

Johson, K. (2009). Research on Evaluation Use: A Review of the Empirical Literature From 1986 to 2005. *American Journal of Evaluation*, *30*(3), 377-410. DOI: doi: http://dx.doi.org/10.1177/1098214009341660

Joint Committee On Standards Of Educational Evaluation (2003). *The student evaluation standards.* Thousand Oaks, Ca.: Corwin

Kirkhart, K. (2000). Reconceptualizing evaluation use: An integrated theory of influence. In V. Caracelli, & H. Preskill (Eds.), *The expanding scope of evaluation use. New*

Directions for Evaluation, Nº 88, San Francisco, CA: Jossey-Bass.

La Velle, J. M. & Donaldson, S. I. (2010). University-Based Evaluation Training Programs in the United States 1980-2008: An Empirical Examination**.** *American Journal of Evaluation*, *31*(1), 9-23. DOI: http://dx.doi.org/10.1177/1098214009356022

Leviton, L. C. (2003). Evaluation use: Advances, challenges and applications. *American Journal of Evaluation*, *33* (2), 159-178.

Ledermann, S. (2012). Exploring the Necessary Conditions for Evaluation Use in Program Change. *American Journal of Evaluation*, *24* (4), 525-535. DOI: http://dx.doi.org/10.1177/1098214011411573

Makrakis, V. & Kostoulas-Makrakis, N. (2016). Bridging the qualitative–quantitative divide: Experiences from conducting a mixed methods evaluation in the RUCAS programme. *Evaluation and Program Planning* 54, 144-151. DOI: http://dx.doi.org/10.1016/j.evalprogplan.2015.07.008

Mark, M. M. (2003). Toward a integrative view of the theory and practice of program and policy evaluation. In S. I. Donaldson & M. Scriven (Eds.) *Evaluating social programs and problems: Vision for the new millenium*, (pp. 183-204). Mahwah, NJ: Erlbaum.

Maxcy, S. J. (2003). Pragmatic threads in mixed methods research in social sciences: An emerging theory in support of practice. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social anb behavioral research*, (pp. 51-89). Thousand Oaks, CA: Sage

May, H. (2004). Making statistics more meaningful for policy research and program evaluation. *American Journal of Evaluation*, 25(4), 525-540. DOI: http://dx.doi.org/10.1177/109821400402500408

Mcclintock, C. (2003). Commentary: The evaluator as scholar/practicioner/change agent. *American Journal of Evaluation*, *24*(1), 91-96.

DOI: http://dx.doi.org/10.1177/109821400302400110

Mira, G. E., Meneses, R. M. & Rincón, D. A. (2012). La Investigación Evaluativa y su perspectiva en la Acreditación y Evaluación de Programas e Instituciones en Educación Superior, *XIII Asamblea General de la Asociación Latinoamericana de Facultades y Escuelas de Contaduría y Administración (ALAFEC)* (pp. 1-26). Buenos Aires, Argentina.

Muñoz, A., Perez Zabaleta, A., Muñoz, A. & Sanchez, C. (2013). La evaluación de políticas públicas: una creciente necesidad en la unión europea. *Revista de Evaluación de Programas y Políticas Públicas*, *1*, 1-30.

Neuman, A., Shahor, N., Shina, I., Sarid, A. & Saar, Z. (2013). Evaluation utilization research -Developing a theory and putting it to use. *Evaluation and Program Planning, 36*, 64–70. DOI: http://dx.doi.org/10.1016/j.evalprogplan.2012.06.001

Nicoletti, J. A. (2013). La evaluación de la calidad educativa. Investigación de base evaluativa en centros de educación superior. *Revista Argentina de Educación Superior*, *6*, 189-202.

Nitsch, M., Waldherr, K., Denk, E., Griebler, U., Marent, B. & Forster, R. (2013). Participation by different stakeholders in participatory evaluation of health promotion: A literature review. *Evaluation and Program Planning*, *40* (1), 42–54. DOI: http://dx.doi.org/10.1016/j.evalprogplan.2013.04.006

Paricio, J. (2015). *Análisis de los modelos de calidad de la educación superior. Diseño de una metodología de análisis multidimensional*. Tesis Doctoral, Universidad de Zaragoza.

Patel, M. (2002a). A meta-evaluation, or quality assessment, of the evaluations in this issue, based on the African Evaluation Guidelines: 2002. *Evaluation and Program Planning, 25*, 329-332. DOI: http://dx.doi.org/10.1016/S0149-7189(02)00043-5

Patel, M. (2002b). The African Evaluation Guidelines: 2002. A checklist to assist in planning evaluations, negotiating clear contracts, reviewing progress and ensuring adequate completion of an evaluation. *Evaluation and Program Planning, 25*, 481–492.

Patton, M. Q. (2012). A utilization-focused approach to contribution analysis. *Evaluation, 18*, 364–377. DOI: http://dx.doi.org/10.1177/1356389012449523

Perassi, Z. (2009). Commentary: Evaluar un Programa Educativo: Una Experiencia Formativa Compleja. *Revista Iberoamericana de Evaluación Educativa, 2*(2), 172-195.

Perez Juste, R. (2002). La evaluación de programas en el marco de la educación de calidad. *XXI Revista de Educación*, *4*, 43-76.

Perrin, B. (2001). Commentary: Making yoursel -and evaluation- useful. *American Journal of Evaluation*, *22* (2), 252-259. DOI: http://dx.doi.org/10.1177/109821400102200209

Pinkerton, S. D., Johnson-Massoti, A. P., Derse, A. & Layde, P. M. (2002). Ethical issues in cost-effectiveness analysis. *Evaluation and Program Planning, 25*, 71-83. DOI: http://dx.doi.org/10.1016/S0149-7189(01)00050-7

Preskill, H., & Boyle, S. (2008). A multidisciplinary model of evaluation capacity building. *American Journal of Evaluation*, *29*, 443–459. DOI: http://dx.doi.org/10.1177/1098214008324182

Renger, R & Hurley, C. (2006). From theory to practice: Lessons learned in the application of the ATM approach to developing logic models. *Evaluation and Program Planning, 29*, 106–119. DOI: http://dx.doi.org/10.1016/j.evalprogplan.2006.01.004

Rodriguez-Campos, L. (2012). Stakeholder involvement in evaluation: Three decades of the American Journal of Evaluation. *Journal of Multidisciplinary Evaluation*, *17*, 57– 79.

Rodríguez-Campos, L. (2012). Advances in collaborative evaluation. *Evaluation and Program Planning*, 35, 523–528. DOI: http://dx.doi.org/10.1016/j.evalprogplan.2011.12.006

Roseland, D., Lawrenz, F. & Thao, M. (2015). The relationship between involment in and use of evaluation in multi-site evaluations. *Evaluation and Program Planning*, *48*, 75-82. DOI: http://dx.doi.org/10.1016/j.evalprogplan.2014.10.003

Ryan, K. E. (2004). Serving public interests in educational accountability: Alternative approaches to democratic evaluation. *American Journal of Evaluation*, *25*(4), 443-460. DOI: http://dx.doi.org/10.1177/109821400402500403

Scheerens, J. (2004). The evaluation culture. *Studies in Educational Evaluation*, *30*(2), 105-124.

Schwandt, T.A. (2002). *Evaluation Practice Reconsidered*. New York NY. Peter Lang Publishing

Schwartz, R. & Mayne, J. (2005). Assuring the quality of evaluative information: theory and practice. *Evaluation and Program Planning*, *28*(1), 1-14. DOI: http://dx.doi.org/10.1016/j.evalprogplan.2004.10.001

Schweigert, F., J. (2007). The priority of justice: A framework approach to ethics in program evaluation. *Evaluation and Program Planning, 30*, 394–399. DOI: http://dx.doi.org/10.1016/j.evalprogplan.2007.06.007

Scriven, M. (2000). *The logic and methodology of checklists*. Retrieved from www.wmich.edu/evalctr/checklists/

Scriven, M. (2003). Evaluation in the new millenium: The transdisciplinary vision. In S. I. Donaldson & M. Scriven (Eds.), Evaluating *Social Programs and Problems. Visions for the New Millennium*, (pp. 19-42). Mahwah, N. J.: Lawrence Earlbaum Associates

Smith, M. J. (2010). *Handbook of Program Evaluation for Social Work and Health*

Professionals. New York: Oxford University Press

Sondergeld, T. & Koskey, K. (2011). Evaluating the impact of an urban comprehensive school reform: An illustration of the need for mixed methods. *Studies in educational Evaluation*, *37*, 94-107. DOI: http://dx.doi.org/10.1016/j.stueduc.2011.08.001

Stake, R. (2006). *Evaluación comprensiva y evaluación basada en estándares*. Editorial Graó, Barcelona.

Stufflebeam, D. L. (2000). *Guidelines for developing evaluation checklists*. Retrieved from www.wmich.edu/evalctr/checklists/

Stufflebeam, D. L. (2001a). Interdisciplinary PHd Programming in Evaluation. *American Journal of Evaluation*, *22*(3), 445-455. DOI: http://dx.doi.org/10.1177/109821400102200323

Stufflebeam, D. L. (2001b). The metaevaluation imperative. *American Journal of Evaluation*, *22*(2), 183-209. DOI: http://dx.doi.org/10.1177/109821400102200204

Stufflebeam, D. L. (2004). A note on the purposes, development, and applicability of the Joint Committee Evaluation Standards. *American Journal of Evaluation*, *25*, 1, 99-102. DOI: http://dx.doi.org/10.1177/109821400402500107

Taut, S. (2008). What have we learned about stakeholders involvement in program evaluation. *Studies in Educational Evaluation*, *34*, 224-230. DOI: http://dx.doi.org/10.1016/j.stueduc.2008.10.007

Thomas, V. G. & Madison, A. (2010). Integration of Social Justice Into the Teaching of Evaluation. *American Journal of Evaluation*, *31* (4), *570-583*. DOI: http://dx.doi.org/10.1177/1098214010368426

Urban, J. B., Hargraves, M. & Trochim, W. M. (2014). Evolutionary Evaluation: Implications for evaluators, researchers, practitioners, funders and the evidence-based program mandate. *Evaluation and Program Planning*, *45*, 127-139. DOI: http://dx.doi.org/10.1016/j.evalprogplan.2014.03.011

Urban, J. B. & Trochim, W. M. (2009). The role of evaluation in research-practice integration: Working toward the "Golden spike". *American Journal of Evaluation*, *30* (4), 538-553. DOI: http://dx.doi.org/10.1177/1098214009348327

Vanhoof, J. & Van Petegem, P. (2010). Evaluating the quality of self-evaluations: The (mis)match between internal and external meta-evaluation. *Studies in Educational Evaluation*, *36*, 20–26. DOI: http://dx.doi.org/10.1016/j.stueduc.2010.10.001

Walton, M. (2014). Applying complexity theory: A review to inform evaluation design. *Evaluation and Program Panning, 45*, 119-126. DOI: http://dx.doi.org/10.1016/j.evalprogplan.2014.04.002

Wasserman, D. L. (2010). Using a systems orientation and foundational theory to enhance theory-driven human service program evaluations. *Evaluation and Program Planning, 33*, 67-80. DOI: http://dx.doi.org/10.1016/j.evalprogplan.2009.06.005

Weiss, C. H. (2004). On theory-based evaluation: Winning friends and influencing people. *The Evaluation Exchange*, 9(4), 1-5.

White, H. (2013). The Use of Mixed Methods in Randomized Control Trials. *New Directions for Evaluation*, *138*(2), 61-73. DOI: http://dx.doi.org/10.1002/ev.20058

Youker, B. W., Ingraham, A. & Bayer, N. (2014). An assessment of goal-free evaluation: Case studies of four goal-free evaluations. *Evaluation and Program Planning*, *46*, 10-16. DOI: http://dx.doi.org/10.1016/j.evalprogplan.2014.05.002

Yusa, A., Hynie, M. & Mitchell, S. (2016). Utilization of internal evaluation results by community mental health organizations: Credibility in different forms. *Evaluation and Program Planning*, *54*, 11-18. DOI: http://dx.doi.org/10.1016/j.evalprogplan.2015.09.006

Author / Autor	To know more / Saber más

Escudero, Tomás (tescuder@unizar.es).

Chair of Educational Research at the University of Zaragoza. He has authored a large number of Works and publications in the evaluative education field, particularly institutional evaluation. His postal address is: Universidad de Zaragoza. Instituto de Ciencias de la Educación. C/ Pedro Cerbuna, 12. 50009 Zaragoza (España)

Revista **EL**ectrónica de **I**nvestigación y **EV**aluación **E**ducativa
E-Journal of Educational Research, Assessment and Evaluation

[ISSN: 1134-4032]