



# “Il lato oscuro della valutazione”: The perception of evaluation and assessment in a University context

## “Il lato oscuro della valutazione”: La percezione della valutazione e dell’assessment in un contesto universitario

Antonella Nuzzaci

Università degli Studi dell’Aquila / antonella.nuzzaci@univaq.it

Anna Salerni

Università Roma 1 “La Sapienza” / anna.salerni@uniroma1.it

Maria Vittoria Isidori

Università degli Studi dell’Aquila / mariavittoria.isidori@cc.univaq.it

Alessandro Vaccarelli

Università degli Studi dell’Aquila / alessandro.vaccarelli@cc.univaq.it

### ABSTRACT

The aim of this contribution is to focus on the reasons for the rejection of evaluation practices in the field of higher education and to understand how this refusal is connected to the “mechanism of unveiling of the educational system” that the evaluation implements. It attempts to provide an interpretative framework regarding the perception of students (but also university professors) of evaluation practices adopted in the university context, proposing to give an explanation that, although provisional, can contribute to the understanding of the refusal that usually accompanies the evaluation, which becomes not only the spy of a malaise but also that of a malfunctioning of teaching.

This perception is an intervening factor, often neglected, but important if we want to fully understand the “how, when and why” the learning and the meaning of teaching are realized, and if we want to reveal the reasoning behind the educational action system with which it is possible to clarify the logics that guide the management of the interventions by the teachers.

L’obiettivo di questo contributo è quello di mettere a fuoco le ragioni legate al rifiuto delle pratiche valutative nel campo dell’istruzione superiore e di capire come tale rifiuto sia connesso al “meccanismo di svelamento del sistema didattico” che la valutazione mette in atto.

Esso tenta di fornire un quadro interpretativo riguardante la percezione degli studenti (ma anche dei docenti universitari) delle pratiche di valutazione adottate in contesto universitario, proponendosi di dare una spiegazione che, seppur provvisoria, possa offrire un contributo alla comprensione del rifiuto che generalmente accompagna la valutazione, il quale diventa non solo la spia di un malessere ma anche quella di un cattivo funzionamento dell’insegnamento.

Tale percezione è un fattore interveniente, spesso trascurato, ma importante se si vuole comprendere appieno “il come, il quando e il perché” si realizzi l’apprendimento e il senso dell’insegnamento, nonché se si desidera svelare ragionamenti che stanno alla base del sistema d’azione didattico con i quali è possibile chiarire le logiche che guidano la gestione degli interventi da parte dei docenti.

### KEYWORDS

Higher Education, Evaluation, Assessment, Perception of Students.

Istruzione Superiore, Valutazione, Valutazione, Percezione degli Studenti.

## Introduction

The growing need for qualified teaching in higher education has led educational research to increasingly deal with university teaching, design, teaching and learning and evaluation issues, especially in the field of initial and vocational training. continue teachers, where the intent to focus attention on the importance of building strong cultural profiles becomes more urgent for the consequences that derive from the school. In fact, starting from the need to reinforce the methodological skills of future teachers becomes a way to allow them to promote the quality of the training and learning of the students they will take care of, as well as to take care of their progress.

In fact, it is now known how, within this kind of skills (methodology), which constitutes the “hard core” of teaching professionalism, there is precisely the ability to evaluate and use assessment tools, which represent the instrumental kits that allow governing the teaching-learning processes at all levels.

One of the aims of university education is to promote the quality of student education and learning, as well as to take care of their progress; and assessment and assessment help to govern the teaching-learning processes and are an integral part of the educational action system, which rely on internal and external regulation and self-regulation devices, which rely on internal and external regulation devices and self-regulating devices.

In this sense, from the sixties onwards the evaluation has been the subject of a growing interest on the part of educational research (Dewey, 1939/1960; Visalberghi, 1955; 1958; Scriven, 1967; Gattullo, 1967; 1986a; 1986b; De Landsheere, 1973; Calonghi, 1976) and, later, around the 80s-90s, a demanding enhancement work within a general policy of reorganizing the system of higher education in Italy, in Europe (Edler, Berger, Dinges, & Gok, 2012; Nuzzaci & Grange, 2009; Nuzzaci, 2016) and in the rest of the world (Heywood, 2000), as it is closely connected to the central political decisions and usually associated with the concerns related to the quality of university teaching (Gibbs & Simpson, 2004, p. 3) and, therefore, to the achievement of the objectives in terms of performance, results and responsibilities (Linkon, 2005, p. 28). Evaluation is one source of helps to define problems.

In the last twenty years, following the transformation of the university system and the expansion of its social base and in relation to the growing number of students who have become part of the system, progressively more attention has been given to the value of the evaluation with reference to the quality of teaching and learning outcomes, an aspect that has progressively pushed researchers to deal more wisely with the task of verifying and ascertaining skills and knowledge within the methods of managing teaching-learning processes.

This process has been profoundly influenced by profound changes at various levels, which have impacted higher education in Europe, gradually highlighting the importance of placing evaluation at the center of teaching-learning processes as an effective tool to be used systematically and with coherence to promote learning and to increase the knowledge of the characteristics of university students and of the learning conditions, as well as qualitatively the level of the acquisitions. Furthermore, the research has shown that that quality assessment can have a greater positive impact on student learning than any other intervention (Davies et al., 2012) throughout a period of instruction, as well as achievement at the end of the learning period instruction (Wiliam, 2011).

This is so true that in 2004, in an article published by W. James Popham, the

lack of teacher literacy in evaluation is interpreted as a real “professional suicide” (2004, pp. 83-84).

Recently, the assessment was put in Italy, also at the Center of the professionalism of the university teacher, as an effective tool to substantially increase the level of training of students and their learning.

This is partly due to the widening of the social base of higher education and the increasing number of students accessing the university today that is pushing university teachers to acquire more skills in the ways of managing teaching and to systematically and effectively conduct tasks of verification and assessment of skills and knowledge.

In support of these positions, in fact, the most representative meta-analyses showed how the evaluation produces decisive effects on the student learning process and contributes to “raise the cultural fate”, as well as the results (Hattie, 2009), indicating in particular, the high quality of “evaluation practices” and their virtuous integration in “teaching practices” are able to increase the students’ level of learning (Black & Wiliam, 1998; 2010) with evident repercussions on the internalization of acquisitions (Yeh, 2006; 2007), of their retention, integration and transfer (De Ketele, 1993a; 1993b); even if still open some questions remain regarding the conditions of success of certain forms and evaluation models with respect to others and their beneficial effects on different kinds of learning (Benett, 2011; Briggs et al., 2012; Kingston & Nash, 2011).

Evaluation is therefore important to understand the learning processes, but equally important is the perception that students have of it (Sambell et al., 1997, p. 335), as well as the way in which they approach it (Struyven et al., 2005, p. 329) (just as important would seem to understand how teachers perceive and use it).

In understanding the mechanisms that stimulate learning, taking into account the conceptions that students have of it and the evaluation are to be considered important factors for the explanation of the refusal that generally occurs with regard to the evaluation itself.

The results of the research in fact reveal how the perceptions that the students have of the evaluation significantly influence the approaches they have towards learning and the behaviors of study and the latter in turn affect the ways in which they perceive the assessment.

They suggest that students with positive opinions about the assessment and the different forms of assessment are more available to learn and achieve more positive results than those who are governed by negative concepts, as well as less fearful to face more innovative assessment methods and more urged to stress the problems of “correctness” about the evaluation methods used.

### **1. The “dark side” of the evaluation**

If it is therefore true that evaluation is now considered a key element of the educational system, especially because it is aimed at enabling all teachers, including university teachers, to understand, regulate and implement the teaching-learning processes, it is equally true as even today there are still some doubts about the concrete uses that are made of it and its actual usefulness, but also a little practical use of its most advanced forms, in particular, the role played by the formative assessment and feedback should be used to empower students as self-regulated learners. The self-regulation refers to the degree to which students can regulate aspects of their thinking, motivation and behaviour during

learning (Nicol & MacFarlane-Dick, 2016; Pintrich & Zusho, 2002). It is manifested in the active monitoring and regulation of a number of different learning processes: e.g. the setting of, and orientation towards, learning goals; the strategies used to achieve goals; the management of resources; the effort exerted; reactions to external feedback; the products produced.

Si osservi, cioè, il perdurare della presenza di un “lato oscuro”, tuttora poco esplorato, che, nella realtà, pare induca gran parte degli individui a rifiutarla. In contesto universitario, tale aspetto merita di essere indagato a partire dall’analisi delle rappresentazioni che docenti e studenti si fanno di essa (Sambell et al., 1997, p. 357) e dal modo in cui la percepiscono e spontaneamente l’approcciano (Struyven et al., 2005, p. 329).

It is observed, that is, the persistence of the presence of a “dark side”, still little explored, which, in reality, seems to induce a large part of individuals to reject it. In a university context, this aspect deserves to be investigated starting from the analysis of the representations that teachers and students make of it (Sambell et al., 1997, p. 335) and the way in which they perceive it and spontaneously approach it (Struyven et al. 2005, p. 329).

Taking into account perceptions, as well as conceptions / misconceptions, which students and teachers have of evaluation also means trying to understand which factors hinder or stimulate learning, since the theory of “conceptual change” is able to provide a robust basis for to frame the questions on the evaluation of students’ learning (Wandersee et al., 1994), being able to say the result of the interaction between what the students think, or what ideas or concepts they possess (Posner et al., 1982), and what is taught to them.

Concepts, as well as perceptions, develop and change with education only if students are really actively involved in the learning process with appropriate and appropriate interventions. In this direction, many studies clearly show how it is possible to develop and leverage forms of educational innovation aimed at strengthening feedback systems in terms of control and correction to obtain significant benefits in terms of knowledge, as the frequent answers that students receive about their learning helps them to recalibrate actions leading them progressively towards training success.

For example, the repercussions that the different categories of feedback (constructed and based on the data applicable to the different abilities and behaviours) have on the performance of the students and the benefits that they derive from their appropriate use, such as to provide a solid basis for the development of theoretical models regarding formative evaluation and for the affirmation of perspectives aimed at improving educational practices (Black & Wiliam, 1998).

The students’ perceptions and the role they played in self-control processes, alongside the strategies used by teachers and those embedded in specific systemic approaches, such as “learning for mastery”, are now considered an integral part of that educational system, design (instructional design) and action, which allows the achievement of training success.

The results of the research reveal in fact how the conceptions and perceptions that students have of the evaluation significantly influence the approaches to learning and study behaviors, which in turn seem to influence the ways in which they perceive the assessment.

Some studies have shown that it was extremely useful for research (Sambell et al., 1997; Segers et al., 2001) to ask students what kind of study behaviours they adopt in relation to the forms of evaluation used by the teachers and how the information (of different nature) obtained become valuable to be able to

reconstruct a rich and contextualized framework of the widespread and generalized attitude of refusal towards evaluation.

Specifically, these studies suggest that students who have positive opinions about the assessment have a higher emotional disposition that is conducive to learning and achieve higher results than those who tend to reject it, showing even less fearful in coping with methods and more innovative and more critical evaluation techniques to highlight problematic aspects linked to the evaluation methods adopted in terms of “correctness”, “consistency” and “validity” of the tests (Nuzzaci, 2007a; 2007b).

In this direction, then attempting to provide an explanation of the refusal that frequently accompanies the evaluation and its practices means trying to understand what “mechanisms” it sets in motion and what effects it produces in terms of teaching-learning processes.

The perception of the experience lived by the student in the evaluation is a very important reality that in a university context often appears as a “forgotten dimension”; on the contrary, it can be called a decisive intervening variable, which must not be neglected if one wants to fully understand “how, when and why” the learning is realized and it is desired to reveal reasoning that is at the basis of the didactic action system with which it is possible to clarify the logics that guide the management of the interventions by the teachers.

The refusal and discomfort with regard to evaluation, which too often accompany evaluation practices in the field of education, are not only a sign of a deeper general malaise, but also and above all the sign of a “malfunctioning of teaching”, fed the latter by verification systems aimed at “judging” rather than “forming” and which end up slowly turning into instruments of “condemning the differences” of the students.

This refusal to take charge of the characteristics of the students leads them progressively towards failure instead of favoring their success, which can be considered one of the priority purposes of education.

Therefore, within a precise frame of interpretation, it is a question of focusing on the reasons, albeit provisionally, how refusal and discomfort are connected to the complex “mechanism of unveiling of the didactic system”, which brings out those self-referential and a-critical components, as well as poorly coherent with the didactic action, which lead to the lack of the central purpose of the training, ie the training success of all students, based primarily on the ability to be able to value differences rather than penalize them.

## 2. Evaluate to design

Evaluation is a term that includes different methods and methods for collecting information on the nature of student performance both in terms of learning processes and their progress in studies; effective evaluation includes gathering any information that may help to understand the level of learners’ acquisition of skills and knowledge.

It involves the analysis and interpretation of information, as well as the evaluation of such interpretations, which are processes that can be influenced by external and subjective factors.

This explanation, however, becomes unlikely if the evaluation techniques used, both quantitative and qualitative, that the tools (ie multiple choice structured tests, extended answers, problems, projects, laboratory investigations and interviews, etc.), did not make use of valid learning measures.

However, it is often remembered by many parties, for example, how evaluation and assessment, in the students' idea, are not only conceived as totally overlapping terms (this in Italy, in particular, occurred due to the absence of a widespread evaluative culture), but are generally attributable to negative experiences and perceptions, as well as

- "poorly correct" practices that are expressed at different levels (types of delivery, type of tests used, times, etc.);
- to figures who assume a position of authority exerting on others a certain power (the teacher-student asymmetry so strongly recalled from the history of education) precisely in relation to the examination of the performance results, frequently used to return a "judgment on the person" rather than being interpreted as signs to "read" the quality of a performance (think of the effects now known to all the "classical" educational literature, such as halo effect, pigmalion, etc.);
- feelings of subjection, unmodifiable judgment, misunderstanding, etc., which, in some cases, favor the emergence of states of anxiety or fear in the students (and in teachers) in evaluative or certification moments and, in others, provoke in their disturbances for any component or perspective regards the evaluation without being able to make him understand the positive aspects of it or even to make him understand how it can be objectively employed (Donaldson, Gooler, & Scriven, 2002, p. 226) to learn or to teach;
- to states of difficulty, uncertainty or impotence, commonly related to situations of incapacity, subjection and conditions undergoing;
- to sources of stress or conditioning related to the taking of crucial decisions, to forms of decision making processes, to significant cultural passages that concern determining factors such as success, promotion, expectations, orientation, employment, etc. (by way of example, we recall the role played by the contamination between the concept of general self and the concept of school self in the training process, with significant repercussions on both university and professional success);
- unpleasant forms of thought or perception that lead, in higher education processes and beyond them, to diminish the ability to give the "best of themselves" with negative repercussions also on the feeling of self-efficacy and satisfaction with knowledge and on the personal systems of reference, as well as on the way we learn, study and socialize.

Emblematic situations are those that we often witness: students pervaded by anxiety, gripped in a grip of fear, fear, apprehension, concern of "failing" (accompanied by a sense of insecurity and apprehension and physical symptoms such as sweat of the hands, increased heart rate, insomnia, general excitement, etc.), sitting nervously in a classroom to support a test through which someone will be asked to evaluate their skills concerning a certain cultural journey.

This is a test, which can be entry (as, for example, the access test) or final (such as, for example, the case of the high school graduation exam that every year is made to rise at the central moment of the process of education shared by the rest of the community), capable of generating anxiety and anxiety, which the students will try to receive the result obtained at that same test.

It is therefore legitimate to ask why it is not possible to undermine this intrinsic link between evaluation and anxiety and for what reason assessment and assessment cannot be authentically "interpreted" as tools to support confidence in the ability to learn.

These are processes integrated into the regular practice of a university study course, based on meaningful learning objectives, valid evidence and criteria aimed at providing solid information to the teacher to increase students' success and to provide them with useful answers to support them. on the path towards understanding and motivation, improving performance and, in the specific case of future teachers, training for teacher professionalism, avoiding repercussions on the plan of educational particle that will interest them when they enter the school.

It is a not inconsiderable fact that almost always the evaluative experiences are connected to traumatic situations experienced in the formative path (Donaldson, Gooler, & Scriven, 2002), which proceed as a cumulative effect and put most of the people later in the condition of to show great reservations about valuation practices of any kind, which are perceived as a threat or a hazard even before they manifest themselves concretely and are clearly identified, causing individuals to flee more than to "fight" or adopt "strategies" defensive "because of the fear of exposing oneself to forms of failure.

Teachers and students are rebuffed to talk about evaluation; and even when this happens the answers provided include a wide range of terms whose meanings are almost always negatively connoted and which, in the case of teachers, range from common expressions such as "form of control", "budget cuts", "loss of time "to those as" "artificiality", "uselessness", and, in the case of students, are linked to feelings of "fear" and states of anxiety up to the use of emblematic phrases, such as that known to everyone, of a teacher conceived as one who has "the knife on the side of the handle".

These negative connotations always prevail over positive ones, conveying a conception of learning evaluation linked to success rather than one *for learning* focused on *how to learn* (Torres & Preskill, 2001; Preskill & Boyle, 2008; Preskill, 2008).

The negative semantic value of the etymon seems to have contaminated the concept of "performance evaluation", in terms of results achieved and understanding of how these performances have been achieved or have been only exceptionally "good" or "bad", as well as factors have had a positive or negative influence on them and who should be responsible for it.

It is clear that if the teachers about how they evaluate the learnings of the students are appealed in various ways by the learners (the examples that can be done in this regard are endless) we cannot think that the evaluation is being carried out appropriately and that his action can be considered independently of these feelings of opposition.

Therefore, when we try to understand the role and function that evaluation has within the teaching-learning processes we cannot limit ourselves to aseptically considering its intrinsic qualities, but we must also take into account how much we rotate around the failure to affirm an evaluative culture in the university, which is not yet able to become a current practice, capable of helping students more than "promoting them".

This is likely to be due to a lack of incorporation of university evaluation into the formative process which ends up requiring continuous adjustments rather than following the perspective and evolution of the teaching-learning process itself, as well noted in that literature (Ginsburg et al., 1993; Shepard et al., 1998) which argues that education and evaluation are inseparable and interconnected dimensions.

This is because when the evaluation is separated from the teaching, the learning loses its value, at the same time, formal and substantial.

The separation and / or disconnection of the teaching from the evaluation creates subtle reverberations on the students' knowledge process, since the acquisition of information and the adoption of strategies to better respond to certain learning needs are dissociated by reducing the interpretive space of need and, therefore, of the effectiveness of didactic intervention.

This situation has involved different scholars in the search for an idea of "integrated and interactive evaluation" that corresponds to a systemic logic of grafting in the teaching and learning process (Cardinet, 1986) with clear benefits in terms of the quality of the acquisitions in a optics of educational continuity.

### 3. Evaluation and assessment: two sides of the same coin

So what does evoke or hide the evaluation? What is it able to hide or unveil? Why do you fear so much? Even in the best case it is forgotten that the evaluation is, as Maria Lucia Giovanni underlines "the attribution of a value judgment" (2014), which can be traced in the general sense of a tradition (Visalberghi, 1955; Gattullo, 1967; Giovannini, 2014).

It is above all a measure of the design and, more generally, methodological capacity of the teacher, which relies on concrete aspects of teaching to promote and implement learning, directly attributable to the educational action system in use (Nuzzaci, 2017), but of which it is able to reveal its functioning, that is to "bare" its essential components, in terms of defects and virtuosity, strengths and weaknesses, and, above all, to measure its quality in reference to the degree of consistency and congruence between objectives pursued and evaluation strategies adopted.

That system which is the mirror of the design system that sees techniques and tools of instructional design acting within the field of intervention.

In this regard, De Ketele (1993a), to support the value that evaluation takes in teaching and to show how it has the capacity to reveal the teacher's thought about teaching, at all levels, and how to manage it, as well as to reveal the functioning from the point of view of the logic that governs his action, recovers an expression of common sense linked to tradition ("Tell me who you go and I'll tell you who you are...", "Tell me how you eat and I'll tell you ...") and apply it to the evaluation: "Tell me how you evaluate and I will tell you what your pupils or students really learn [...] and your real conception of learning" (De Ketele, 1993, p.15).

The attempt is to make people understand how the function of evaluation takes on a much more important character than what is commonly believed, since it is able to relate objectives, contents and knowledge of learning; it is that relationship that has its roots in critical analysis and in classical studies on the effectiveness and efficiency of teaching processes.

The evaluation intended as "unveiling" therefore basically recalls the awareness and intentionality of teaching and learning as capable of unmasking the functioning of the design and action educational system employed by the university teacher.

This happens because the evaluation, even before being "measured", is a comparison between a series of functions and activities conducted by the teacher to "govern the teaching practices" (programming actions, planning and so on) and the effects produced by them on learning of the students (Crooks, 1988).

But it also becomes an "indirect measure" of teaching when, evaluating students' abilities, it reveals information on teaching effectiveness or, better said,

when, recording the student's performance, it measures the teacher's effectiveness.

University professors, however, usually surprisingly ignore the fact that their students' performance is an indicator capable of measuring the quality of their education, just as they often forget to investigate and interpret the results of a test appropriately.

Stiggins (2002), in this regard, clarifies the need for evaluation in teaching and defines a "assessment literate teacher", a teacher who knows what evaluation methods to use to gather information on student results, how to conduct an effective dialogue on evaluation results using scores, reports, portfolios, etc., how to use them to increase student motivation and actively involve them in the learning process.

Basically, the evaluation then only reveals "how to proceed in teaching", putting into play the very meaning of teaching and assessment (why teach / evaluate, what to teach / evaluate, with which methodologies, how to ensure a qualitatively valuable teaching and an assessment suitable for the purpose, how to use the data obtained, etc.).

It responds, in its "complete" meaning, to a "evidence-gathering process" of the students' comprehension skills, assuming an adaptation of the instruction to their needs in order to increase the level and quality of learning.

This process must be integrated into classroom practice by providing at least three key steps:

- set precise and meaningful learning objectives and provide learners with criteria to define the conditions for success;
- use valid and reliable tests that allow students to understand the meaning of these surveys, providing them with useful answers throughout the learning process;
- motivate students to improve their performance.

In this regard, it is necessary to digress and to clarify firstly the difference between "evaluation" and "assessment", terms that in Italian tradition have not always been well defined, as the first evaluation has often ended up incorporating the second one of measurement.

In the full and complete meaning of giving value to learning and teaching, the assessment, whose main meaning is precisely that of estimating the quantity and quality of the skills and knowledge acquired by a student, and the evaluation, whose meaning is to give value to the skills and to what a subject is able to do, are understood as concrete "tools" that require the use of criteria, measures and evidence with the intent to manage the interventions, to improve them in "real time", to guarantee the conditions for the training success, using reliable tests for the comprehension of the skills and knowledge corresponding to precise parameters (adequacy, coherence, etc.).

However, it is also necessary to explain how in the process of "unveiling" three dimensions take on a different role:

- assessment = the process of measuring something (object) with the aim of assigning a numerical value;
- scoring = the procedure for assigning a numerical value to evaluate a test (performance);

- evaluation = the process of determining the value of something in relation to established benchmarks using the information obtained from the assessment, or the process of analysis, reflection and synthesis of the information gathered through the measurement, so as to be able to express judgments and / or make appropriate decisions based on the information collected.

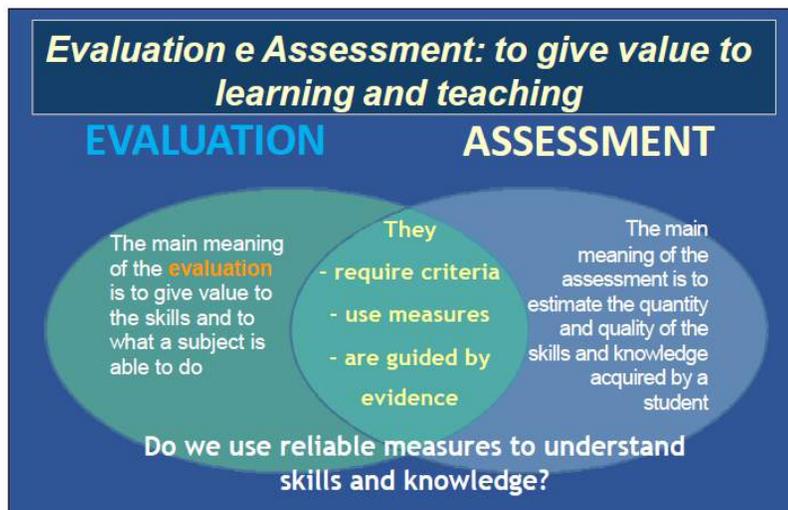


Fig. 1 – Evaluation e Assessment: to give value to learning and teaching

This happens because the evaluation, even before being “measured”, is a comparison between a series of functions and activities conducted by the teacher to “govern the teaching practices” (programming actions, planning and so on) and the effects they produce on the student learning (Crooks, 1988).

Although the evaluation in this sense appears, in its different forms and its different models, a tool for the enhancement of knowledge (Scriven, 1967), a device able to facilitate decision-making (Stufflebeam, 1967; 1997; 2008), responsibility, development, etc., in reality it is almost never perceived as a way

- to guarantee the quality of teaching and the enhancement of teaching activities, the use of specific skills by the teacher, enabling him to increase his pleasure in teaching;
- to ensure the quality of learning, which requires the use of reliable and valid tools to ensure the understanding of what is being measured;
- to increase the effectiveness of communication between the various players in the university system: internal stakeholders (teachers, students, administrative staff) and external stakeholders (social partners, etc.);
- to achieve cultural and social improvements (Henry & Mark, 2003; 2004), capable of affirming a deliberative (House & Howe, 2000) and participatory democracy.

On the contrary, the idea prevails of a “stepmother” evaluation, which ends up guiding the evaluative behaviors towards a “tacit zone”, imbued with implicit and ambiguous, which is revealed through its negative effects and that does not does that hide the ineffectiveness of the educational system in use, whose

planning and intentionality are realized in the inability to finalize learning to understand and to measure it adequately.

This prevents the evaluation of:

- be unique;
- be objective (in terms of observable behavior and performance), but also intersubjective (with respect to the results of the performance itself);
- meet certain conditions;
- respect certain criteria;
- be representative.

The criteria are the most interesting aspect of this discourse, since the “political and procedural correctness” of the assessment, revealing the characteristics and the functioning of the approach to education and the consistency of the system of action in use, help not to forget nothing and to respect the concerns of the recipients of training; furthermore, a “total” rating of all the criteria reinforces the logical conclusions and the cultural power of evaluation as a tool for “government of action”.

It follows that the refusal to measure and use measuring instruments regards the incapacity (and / or renunciation) of being able to obtain organized information and to use specific criteria, which are nothing more than intermediate levels that integrate the information provided from indicators and where an interpretable evaluation crystallizes.

The word “criterion” indicates a term of comparison that allows one to express a judgment subtracted from simple intuitive appreciation and referred to the whole population that presents the same characteristics. In the evaluation, the criterion is the reference on the basis of which a certain performance is evaluated operatively.

Having a criterion means having specified the objectives to be achieved through the training activity, having prepared a test (tool) capable of detecting the behaviors that affect the positive verification of these objectives and having ascertained on a group (in reality a “sample”) the distribution of the results. In the various teaching situations, the appreciation of a student’s performance is therefore based on the position that he / she occupies in the model of distribution detected.

Failing or not wanting to use criteria means putting in place a real “systemic hypocrisy”, which affirms the emergence of a form of “educational ideology” (naturalistic) that is built in relation to the mastery of learning and the determination of a kind of misalignment and inconsistency between discourses and practices.

Thus the unveiling has to do with this “ideology”, unmasking the capacity or inability to use criteria like

- *relevance*: the degree to which the training provided is useful to the recipients to whom it is addressed (especially those with difficulties);
- *effectiveness*: the ability of training to adapt to the needs of the recipients and to comply with the objectives pursued, as well as adequate to the training project outlined by the institution and the teachers;
- *efficiency*: the ability to carry out training at the lowest possible cost and with the least resources (in this sense there is much to be said!);
- *impact*: the ability to measure to what extent the objective has been achieved

- and the result obtained from the measurement is directly determined by the training proposal;
- sustainability: the probability that a teaching activity will be able to pursue its effects even after the implementation of the training intervention.

The criteria are linked to the construction of the evaluation plan to ensure its efficient use and to ensure that it is integrated effectively into the educational system, or harmoniously becomes part of the education design, becoming a concrete tool for managing the teaching-learning processes. (Nuzzaci, 2015).

This calls into question the nature of the evaluation in terms of the description of the coordination actions within the educational system, the articulation of the mechanisms, the devices, the methods and tools used, the use and the moment in which they will be used, as well as the management of information and the use of results.

The criteria in the evaluation are related to the construction of the relationship between discourse and practices as in Figure 1 below.

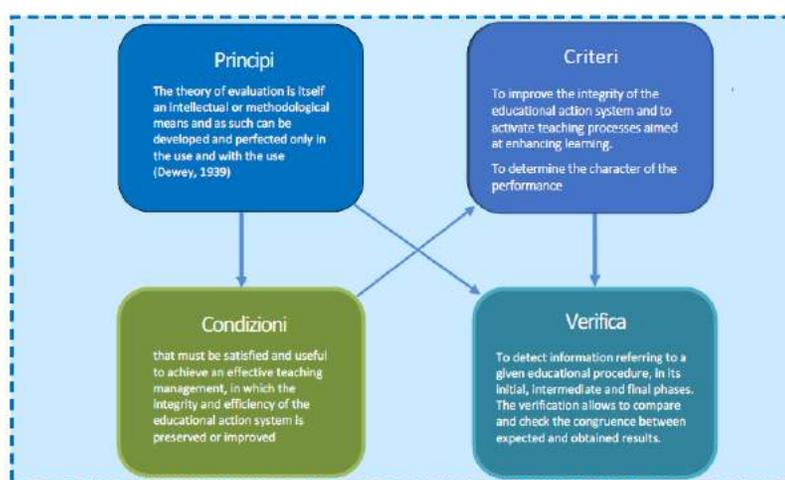


Fig. 2 - Criteria and conditions

Therefore, when a teacher decides to renounce to use the criteria or to use them in an appropriate manner, it only hides his inability to fulfill the objectives to which the evaluation is linked. This refusal is expressed in the incapacity of

- identifying, recognizing and appreciating the “differences” through their “positive discrimination” that leads to an adequate design or educational re-planning aimed at enhancing their particularities, opting for an average student that does not exist;
- planning, planning and informing the educational action addressed to those who most need specific support actions;
- meet the needs of all students, and especially those with special needs, gathering information about the functional characteristics of disability / difficulty from the point of view of learning, determining whether or not it is related to the construct involved in the measurement phase (this is which often makes interpretation of test results difficult);
- take account of differences and be able to deal with them at the level of

educational intervention (individualization and personalization strategies, compensatory and dispensing tools and so on).

To do this it is necessary that assessment and evaluation

- are directed to obtain benefits for all the students;
- pursue a specific purpose;
- use reliable, valid and correctly used tools;
- lead us to reflect on the policies that recognize how the learning of the student is at the center of the training process:
- is aimed at increasing the reliability and validity of the evaluation;
- is methodologically adequate;
- is linguistically appropriate because all assessments are affected by the cultural variable;
- enhance the sources of information.

It is therefore a question of recovering here the formative and formative dimension of evaluation, in the proactive sense of terms, to collect, organize, share and integrate information and interpret it in order to prepare a new sequence of education that takes into account the characteristics of the students and direct more and more towards the use of “differentiated” teaching strategies to try to respond adequately to inequalities and to satisfy the learning needs of all students, especially those who need them most and not only those who do not need them of further support.

This orientation argues that the aim of evaluation is not to demonstrate anything but to strive for continuous improvement through the use of rapid feedback cycles that lead to appropriate decisions (Stufflebeam & Shinkfield, 1985, page 151) and to make the essential changes to the route, where necessary.

It is therefore a matter of understanding the evaluation as “interactive” of progress and of the indispensable identification of students’ learning needs to regulate teaching in a manner suitable to the desired purpose.

It is indeed in fact the use of approaches and techniques of formative evaluation, which helps to satisfy the growing and diversified needs of training and professionalization through the differentiation and adaptation of university teaching, which can be reached to really increase the degree and the quality of acquisitions and greater fairness of results.

However, in reality, the frequent presence of different barriers, including the perceived tensions between the training assessment and the summative assessment, as well as their possible lack of connection, makes it difficult to respond adequately to the various problems that occur interior of the university classrooms.

It is for this reason that a wider dissemination of the principles of formative evaluation and their more consistent use could help university teachers and trainer trainers to identify areas of improvement with greater precision, progressively reducing those of disadvantage, recognizing individual differences (cultural, linguistics, etc.) of the students and taking charge of them (Bruner 1996, Bishop & Glynn, 1999) with the aim of promoting an effective, constructive and authentic culture of evaluation within the educational system.

An assessment capable of taking on the value of both the assessment and the assessment itself and of explaining its relationship with the educational action system in use becomes a tool for promoting knowledge for all students and for

all teachers, stimulating and nurturing a “meaningful learning”, the depth of which determines the quality of the resulting thought (Dewey, 1994/1916), and an “effective teaching”, whose quality succeeds in penetrating the real problems of formation.

On the contrary, an assessment that does not attach adequate importance to both the evaluation and the assessment is transformed, arbitrarily and without criteria, into a “highly limited vacuum device”, which nourishes cultural and social discrimination and highlights the differences by stigmatizing them and relegating them to a zone of “educational darkness”, where ambiguity reigns, understood as non-probabilistic uncertainty (Wakker, 2010), that is to say total absence of available information, to which individuals react with attitudes and behaviors of fear, opposition and rejection.

This renunciation of learning with consequent renunciation of teaching is based on inadequate forms of educational planning, aimed at reinforcing inequality through inappropriate and scarcely ethical evaluation processes, which prevent the fight against and the elimination of indifference due to differences (Perrenoud, 1995; 1998).

An instruction founded on the denial of differences transforms initial inequalities into learning inequalities (Bourdieu, 1966), thus recalling the idea of a university as a “factory of iniquity”, unable to combine the need for “success” with that of the “Equality of opportunities” and to create a “sur mesure” university, in the perspective indicated by Edouard Claparède.

## References

- Bennett, R. (2011). Formative assessment: a critical review. *Assessment in Education*, 18(1), 5-25.
- Bishop, R., & Glynn, T. (1999). *Culture counts: changing power relations in education*. Palmerston North, New Zealand: Dunmore Press.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7-74.
- Black, P., & Wiliam, D. (2010). Inside the black box: raising standards through classroom assessment. *Phi Delta Kappan*, 92(1), 81-90.
- Bourdieu, P. (1966). Condition de classe et position de classe. *European Journal of Sociology*, 7(2), 201-223
- Briggs, D. C., Ruiz-Primo, M. A., Furtak, E., & Shepard, L. (2012). Meta-analytic methodology and inferences about the efficacy of formative assessment. *Educational Measurement: Issues and Practice*, 31(4), 13-17.
- Bruner, J. (1996/1997). *La cultura dell'educazione*. Milano: Feltrinelli.
- Calonghi, L. (1976). *Valutazione*. Brescia: La Scuola.
- Cardinet, J. (1986a). *Evaluation scolaire et mesure*. Bruxelles: De Boeck-Wesmael.
- Cardinet, J. (1986b). *Evaluation scolaire et pratique*. Bruxelles: De Boeck-Wesmael.
- Cardinet, J. (1986c). *Pour apprécier le travail des élèves*. Bruxelles: De Boeck-Wesmael.
- Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research*, 58(4), 438-481.
- Davies, A. Herbst, S. & Parrott-Reynolds, B. (2012<sup>2</sup>). *Leading the way to assessment for learning: a practical guide*. Courtenay, BC: Connections Publishing and Bloomington, IN: Solution Tree Press.
- De Ketele, J.-M. (1993a). Objectifs terminaux d'intégration et transfert des connaissances. In R. Hivon (éd.), *L'évaluation des apprentissages. Réflexions, nouvelles tendances et formation* (pp. 15-26). Sherbrooke (Canada): CRP.
- De Ketele, J.-M. (1993b). L'évaluation conjugulée en paradigme. *Revue française de pédagogie*, 103, 59-80.

- De Landsheere, G. (1973). *Elementi di docimologia. Valutazione continua ed esami*. Firenze: La Nuova Italia.
- Dewey, J. (1960/1939). *Teoria della valutazione*, con un saggio introduttivo di Aldo Visalberghi. Firenze: La Nuova Italia.
- Dewey, J. (1994/2016). *Democrazia ed educazione*. Firenze: La Nuova Italia.
- Donaldson, S. I., Gooler, L. E., & Scriven, M. (2002). Strategies for managing evaluation anxiety: toward a psychology of program evaluation. *American Journal of Evaluation*, 23(3), 261-273.
- Edler, J., Berger, M., F., Dinges, M., & Gok, A. (2012). The practice of evaluation in innovation policy in Europe. *Research Evaluation*, 21(3), 167-182.
- Gattullo, M. (1967). *Didattica e docimologia. Misurazione e valutazione nella scuola*. Roma: Armando.
- Gattullo, M. (1986a). Misurare, valutare, decidere. Un modello e tre adattamenti. *La Scuola Se*, 17-18, 55-58.
- Gattullo, M. (1986b). *Quale università? Proposte per il cambiamento*. Scandicci: La Nuova Italia.
- Gibbs, G., & Simpson, C. (2004). Conditions under which assessment supports student learning. *Learning and Teaching in Higher Education*, 1(1), 3-31.
- Ginsburg, H. P., Jacobs, S. F., & Lopez, L. S. (1993). Assessing mathematical thinking and learning potential in primary grade children. In M. Niss (Ed.), *Investigations into assessment in mathematics education* (pp. 157-167). Boston: Kluwer Academic Publishers.
- Giovannini, M. L. (2014). Il dibattito su valutazione ed evidenze: per un processo valutativo credibile e trasparente. *Journal of Educational, Cultural and Psychological Studies* (ECPS Journal), 9(5), 101-126.
- Gitlin, A., & Smyth, J. (1989). *Teacher evaluation. Critical education and transformative alternatives*. Lewes: Falmer Press.
- Hattie, J. (2009). *Visible learning. A synthesis of over 800 meta-analyses relating to achievement*. London & New York: Routledge.
- 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.
- Henry, G. T., & Mark, M. M. (2004).
- Heywood, J. (2000), Preface. In J. Heywood (Ed.), *Assessment in higher education. Student learning, teaching, programmes and institutions*. London and Philadelphia: Jessica Kingsley Publishers.
- House, E., & Howe, K. (2000). The deliberative democratic view. In K. Ryan and L. DeStefano (eds), *Evaluation as a democratic process: promoting inclusion, dialogue, and deliberation, new directions for evaluation* (Vol. 85, pp. 3-12). San Francisco, CA: Jossey-Bass
- Kingston, N., & Nash, B. (2011). Formative assessment: a meta-analysis and a call for research. *Educational Measurement: Issues and Practice*, 30(4), 28-37.
- Linkon, S. L. (2005). How can assessment work for us? *Academe*, 91(4), 28-32.
- Mark, M. M., & Henry, G. T. (2004). The mechanisms and outcomes of evaluation influence. *Evaluation*, 10(1), 35-57.
- Nicol, D., & MacFarlane-Dick, D. (2006). *Formative assessment and self-regulated learning: a model and seven principles of good feedback practice*. *Studies in Higher Education*, 31(2), 199-218, DOI: 10.1080/03075070600572090.
- Nuzzaci, A. (2006). Evaluation practices and primary school. A research on how teachers use evaluation at school and students perceive it. *7th EES Biennial Conference Evaluation in Society: critical connections*, London, 4-6 ottobre 2006. Abstract 103. From: <http://www.europeanevaluation.org/conferences/ees-conferences-and-events.htm>
- Nuzzaci, A. (2007a). Quando la valutazione diventa elemento strategico del processo di insegnamento-apprendimento. In F. Costabile (a cura di), *Didattica e didattiche disciplinari* (pp. 9-23). Quaderni per la nuova secondaria. Cosenza: Luigi Pellegrini Editore.
- Nuzzaci, A. (2007b). Usare e regolare l'errore: un'occasione per imparare. In F. Costabile (a

- cura di), *Didattica e didattiche disciplinari* (pp. 25-39). Quaderni per la nuova secondaria, 4, Cosenza, Luigi Pellegrini Editore.
- Nuzzaci, A. (2014). Valutare per conoscere, per pianificare e per cambiare, ovvero per innovare il sistema d'azione didattico. In M. V. Isidori, *I disturbi specifici dell'apprendimento* (pp. 197-235). Roma: Anicia. ISBN: 978-88-6709-143-0
- Nuzzaci, A. (2015). L'apprendimento della 'scienza dell'insegnamento': il test di accesso a Scienze della Formazione Primaria predice il successo nella progettazione didattica? - Learning the 'science of teaching': Does the access test of Degree Course in Primary Education Sciences predicts the success in instructional design? *Giornale Italiano della Ricerca Educativa*, 8(14), 227-247.
- Nuzzaci, A. (2016). Promoting and supporting the methodological skills of teachers and trainers for the successful of the teaching and the quality of the training - Scienza dell'insegnamento, formazione e competenze metodologiche dei docenti e dei formatori: dalla progettazione alla valutazione. *Formazione & Insegnamento*, 14(3), 17-36. pISSN: 1973-4778 – eISSN: 2279-7505.
- Nuzzaci, A., & Grange, T. (Eds) (2009). *Qualità, ricerca, didattica. Quale sistema europeo per l'istruzione superiore?* Milano: FrancoAngeli. ISBN: 978-88-568-1320-3
- Nuzzaci, A. (2017). L'esperienza del Corso di laurea in Scienze della Formazione primaria aquilano tra intersezioni curriculari, attività formative e processi di qualità. In G. Domenici (a cura di), *La formazione degli insegnanti iniziale e in servizio* (pp. 148-153). Roma: Armando.
- Patton, M. Q. (2008<sup>4</sup>). *Utilization-focused evaluation*. Thousand Oaks, CA: Sage.
- Perrenoud, Ph. (1995). *La pédagogie à l'école des différences*. Paris: ESF.
- Perrenoud, Ph. (1998). From formative evaluation to a controlled regulation of learning processes. Towards a wider conceptual field. *Assessment in education: principles, policy and practice*, 5(1), 85-102. doi:10.1080/0969595980050105
- Pintrich, P. R., & Zusho, A. (2002). The development of academic self-regulation: the role of cognitive and motivational factors. In A. Wigfield & J. S. Eccles (Eds.), *Development of achievement motivation* (pp. 249-284). San Diego, CA: Academic.
- Popham, W. J. (2004). Why assessment illiteracy is professional suicide. *Educational Leadership*, 62(1), 82-83.
- Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: toward a theory of conceptual change. *Science Education*, 66(2), 211-227.
- Preskill, H. (2008). Evaluation's second act: a spotlight on learning. *American Journal of Evaluation*, 29(2), 127-138.
- Preskill, H., & Boyle, S. (2008). A multidisciplinary model of evaluation capacity building. *American Journal of Evaluation*, 29(4), 443-459.
- Sambell, K., McDowell, L., & Brown, S. (1997). But is it fair? an exploratory study of student perceptions of the consequential validity of assessment. *Studies in educational evaluation*, 23(4), 349-371.
- Scriven, M. (1967). The methodology of evaluation. In R. E. Stake (ed.) *Curriculum evaluation. AERA Monograph Series on Curriculum Evaluation* (Vol. 1, pp. 39-83). Chicago: Rand McNally.
- Scriven, M. (1974). Evaluation perspectives and procedures (pp. 68-84). In W. J. Popham (ed.), *Evaluation in education: Current applications*. Berkeley, CA: McCutchan.
- Scriven, M. (1991). *Evaluation thesaurus*. Newbury Park, CA: Sage.
- Scriven, M. (1993). Hard-won lessons in program evaluation. *New Directions for Program Evaluation*, 58, 1-107.
- Scriven, M. (1994a). Evaluation as a discipline. *Studies in Educational Evaluation*, 20(1), 147-166.
- Scriven, M. (1994b). The final synthesis. *Evaluation Practice*, 15(3), 367-382.
- Scriven, M. (1994c). Product evaluation: the state of the art. *Evaluation Practice*, 15(1), 45-62.
- Segers, M., Dierick, S., & Dochy, F. (2001). Quality standards for new modes of assessment. An exploratory study of the consequential validity of the OverAll test. *European Journal of Psychology of Education*, 16(4), 569-586.

- Segers, M., Dochy, F., & Cascallar, E. (Eds.) (2003). *Optimising new modes of assessment: In search of qualities and standards*. Dordrecht: Kluwer Academic Press.
- Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 4-14.
- Shepard, L. A., Kagan, S. L., & Wurtz, E. (Eds.). (1998). *Principles and recommendations for early childhood assessments*. Washington, DC: National Goals Panel.
- Stiggins, R. (1999). Are you assessment literate? *The High School Journal*, 6(5), 20-23.
- Stiggins, R. (2002) Assessment crisis: the absence of assessment for learning. *Phi Delata Kappan* 83(10) 758-65
- Struyven, K., Dochy, F., & Janssens, S. (2005). Students' perceptions about evaluation and assessment in higher education: A review. *Assessment and Evaluation in Higher Education*, 30(4), 325-341
- Stufflebeam, D. L., & Shinkfield, A. J. (2007). Evaluation theory, models, & applications. San Francisco, CA: Jossey-Bass
- Stufflebeam, D. (2001). Evaluation models. *New Directions in Evaluation*, 89(Spring), 7-98.
- Stufflebeam, D. L. (1967). The use of and abuse of evaluation in Title III. *Theory Into Practice*, 6(3), 126-33.
- Stufflebeam, D. L. (1997). A standards-based perspective on evaluation. In R. L. Stake, *Advances in program evaluation* (Vol. 3): Evaluation and the postmodern dilemma (pp. 61-88). Greenwich, CT: JAI Press.
- Stufflebeam, D. L., & Shinkfield, A. J. (1985). *Systematic evaluation*. Boston: Kluwer-Nijhoff.
- Stufflebeam, D. L., Foley, W. J., Gephart, W. J., Guba, E. G., Hammond, R. L., Merriman, H. O., & Provus, M. M. (1971). *Educational evaluation and decision making*. Itasca, IL: Peacock.
- Stufflebeam, D. L., Madaus, G. F., & Kellaghan, T. (2000). *Evaluation models* (Rev. ed.). Boston: Kluwer.
- Torres, R. T., & Preskill, H. (2001). Evaluation and organizational learning: Past, present and future. *American Journal of Evaluation*, 22(3), 387-395.
- Tyler, R. W., Gagne, R. M., & M. Scriven, M. (Eds.) (1967). *Perspectives of curriculum evaluation*. Chicago, IL: Rand McNally.
- Visalberghi, A. (1955). *Misurazione e valutazione nel processo educativo*. Milano: Edizioni di Comunità.
- Visalberghi, A. (1958). *Esperienza e valutazione*. Torino: Taylor.
- Wakker, P. P. (2010). *Prospect theory: for risk and ambiguity*. Cambridge, UK: Cambridge University Press.
- Wandersee, J. H., Mintzes, J. J., & Novak, J. D. (1994). Research on alternative conceptions in science (pp. 177-210). In D. Gabel (ed.), *Handbook of research on science teaching and learning*. New York: MacMillan.
- Wyatt-Smith, C., Klenowski, V., & Colbert, P. (Eds.) (2014). *Designing assessment for quality learning*. London: Springer.
- Wiliam, D. (2011). *Embedded formative assessment*. Bloomington, IN: Solution Tree.
- Yeh, S. S. (2006). *Raising student achievement through rapid assessment and test reform*. New York, NY: Teachers College Press.
- Yeh, S. S. (2007). The cost-effectiveness of five policies for improving student achievement. *American Journal of Evaluation*, 28(4), 416-436.

