

School Evaluation: A Possible Dialogue between the History and Epistemology

Aldo Sena de Oliveira^{1,2*}, Linsmar Nascimento Lenartovicz²

¹Department of Exact Sciences and Education, Blumenau Center, Federal University of Santa Catarina, Blumenau, Brazil

²Postgraduate Program in Teaching Natural Sciences and Mathematics, University of Blumenau, Blumenau, Brazil

Email: *aldo.sena@ufsc.br

How to cite this paper: de Oliveira, A.S. and Lenartovicz, L.N. (2018) School Evaluation: A Possible Dialogue between the History and Epistemology. *Open Journal of Social Sciences*, 6, 179-189.

<https://doi.org/10.4236/jss.2018.68014>

Received: July 26, 2018

Accepted: August 20, 2018

Published: August 23, 2018

Copyright © 2018 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

This article approaches the school evaluation according to historical and epistemological assumptions and the strategies inserted from its genesis to its application in the context of the contemporary school. The analysis is based on a historical narrative that contextualizes the concept of evaluation and its meaning as practices established from the beginning of humanity, such as its possible origin in China. We report the construction of new meanings for evaluation and its constitution as a stage in the educational process. In addition to the qualitative assignments, the evaluative course that culminates in the development of quantification forms, explained in the docimology, content orientation psychometry and in the intellectual quotient (IQ) analysis, is reported.

Keywords

School Evaluation, Epistemology, History

1. Introduction

Epistemologically, the construction of knowledge has been investigated by several authors whose reflections have convergence in the educational field. The problematizations rely on aspects of theoretical and methodological approaches to make the debate by means of multiple approaches regarding the creation, implementation and evaluation of pedagogical proposals.

In this article, we discuss the evaluation from historical and epistemological assumptions, to subsidize discussions in the context of teaching and learning in formal educational spaces. It is important to note that this is a scan of a clipping of a subject of high complexity, subject of investigations in various areas of knowledge, but that still needs to be extensively (re)discussed and

(re)thought.

Although the evaluative practice was adopted many centuries ago, to control and select, the term “evaluation” is considered recent and was constructed in substitution of the term “exam”, until then frequently used.

The earliest accounts of the examination were in Chinese society in the year 1200 BC, which is not an educational tool but a form of social control. During this period, the exam had a mediating role between the male subjects and the public service. At that time, he had the task of selecting, among male subjects, those who would be admitted to the public service [1].

The examination was proposed for the adjustment of social positioning techniques in Chinese society, as it allowed the social mobility of male subjects, or rather, access to the administration of public functionaries. A few centuries later, particularly in the seventeenth century, there are two currents for the institutionalization of the exam: one comes from Comenius, which defends examination as a place of learning rather than a verification of learning, and, in counterpoint, La Salle defends examination as permanent supervision, aspect of continuous surveillance.

In this sense, it is possible to notice that since its origin, the assessment has been used as an instrument of control and power, in the sense of selecting people, content, processes, among others. Some of the conceptions, more striking in the international scenario about reviews beckon: the model of philosophical socratic assessment that suggested the self-knowledge as a way of determining right or wrong; the proposition of Comenius in the XVII century, focusing on the inclusion of examinations in the educational practice as a strong ally of the teaching practice and the “new” design of the evaluation model in the XVIII century.

In line with the expansion of schools and the creation of the first libraries, although the assessment was still, to represent inheritance, the role of the examination; the quantification of learning in the mid-1900s with the psychometrics of the orientation content—Henri Pierón, with the prerogatives of measurement and assignment of a note to the process of evaluating [2].

2. Brief History about the Evaluation

Practices involving evaluations had its genesis as control instruments that can develop selections in order to include some and separating the others. The logic cartesian that accompanied the creation of these practices/tools, clearly according to a linear thinking that does not constitute a causal relationship, lasted for many centuries and impacts on pedagogical tools used in post-modernity.

The earliest source of examination system can be traced to Xia Shang Zhou period. The Shang and Zhou Dynasties (about 1600 BC - 256 BC). In line with this concept of source selection and as an instrument of in 2205 a.C., a major Chinese emperor was testing his officers every three years, aiming to promote them or fire them [3]. The scheme of competitive examinations in ancient China

had as main purpose to provide the State with able men to work. This form of selection, providing men trained, indicates an instrument of power and exclusion, to eliminate the possibility that other, non-selected, to become official.

In line with the quantification of the assessment, there was an important movement in Europe from the contribution of science to the assignment of the notes to the evaluations, the docimologia as [3]: the docimologia comes from the Greek *dokimé*, which means note. It is the science of the systematic study of the examinations, of the system of assignment of notes and of the behaviors of examiners and examined, taking as a basis the works of [4] [5].

The assessment originated as a form of classification and selection of the evaluation processes after several criticisms of the methods of “examination” hitherto used, thus the systematic study of them (Allocation of notes, interindividual variability and intraindividual at examiners, subjective factors, etc.) [5].

The experiences in docimology focusing on the adequacy of the evidence to the pedagogical objectives, the classification procedures, the preparation of the examiners for the evaluation task, and the use of objective methods of assessing knowledge [5].

Psychometric tests, based on quantification and measurement through psychological scales, can be understood according to the technique of measurement of mental processes, especially applied in psychology and education. It is based on the theory of measurement in science in general, that is, the quantitative method that has, as the main characteristic and advantage, the fact of representing the knowledge of nature with greater precision than the use of the common language to describe the observation of natural phenomena [6]. Historically, psychometrics tests have its origins in the psychophysics of German psychologists Ernst Heinrich Weber and Gustav Fechner. Francis Galton, considered the creator of Psychometrics, also contributed to his development, creating tests to measure mental processes. It was, however, Leon Louis Thurstone, the creator of multiple factorial analysis, which gave the tone to psychometricism, differentiating it from psychophysics. This was defined as the measure of directly observable processes, that is, the stimulus and response of the organism, while the psychometricism consisted in the measure of the behavior of the body through mental processes (law of the judgmental judgement).

The concept of measure in science has provoked debates among researchers, particularly in social sciences. However, the most accepted definition of measure was given by Stanley Smith Stevens in 1946, when he said that: measuring consists of ticking numbers to objects and events according to some rule [6]. To minimize the controlling power, the term “evaluation” was used after assessment and psychometricism, although the control and values of the dominant classes continue to be followed and taught. The pedagogy of the exam is articulated in support of the certification and promotion of the subjects, placing the examination as an element inherent to all educational action. Even in the twentieth century the test of intellectual quotient (QI) is used to measure human intelligence,

using criteria such as cognitive development and the age of the evaluated [7]. This test was considered a scientific, valid and objective instrument that could determine a multitude of psychological factors of an individual, among them are the intelligence, attitudes, interests and learning [8].

For many years, it seems that the prevailing view of education was that, provided instruction was of reasonable quality, it need not be adaptive to the needs of learners [2]. It was assumed either that well-designed instruction would be effective for most students for whom it was intended (with others being assigned to remedial activities) or that the causes of any failures to learn lay within the individual learner (the material was just too hard for them, and they should instead pursue other, and generally less academic, avenues). However, in the 1960, Benjamin Bloom and his graduate students at the University of Chicago began to explore the idea that the normal distribution of student outcomes was not a “natural” outcome but caused by the failure of the instruction to recognize differences in learners [2].

The “Individual System” often regarded as the first truly individualized system of instruction, was developed by Frederic Burk, from 1912 to 1913, for use in the elementary school associated with the San Francisco Normal State School, an institution providing pre-service education for teachers [2].

One of the main reasons that one-to-one tutoring is so effective, according to Bloom, is that the tutor can identify errors in the student’s work immediately, and then to provide clarification, and further follow-up if necessary [9]. Bloom de-scribed these two processes as “feedback” and “correctives” and this language has be-come part of the standard way of talking about assessment ever since. However, in a very important sense, Bloom’s distinction between “feedback” and “correctives” has been counterproductive and has served to distort the original meaning of the term “feedback” in a particularly unfortunate manner [2].

An important feature of Ramaprasad’s definition is that information about the gap between actual and reference levels is considered as feedback only when it is used to alter the gap. If the information is simply recorded, passed to a third party who lacks either the knowledge or the power to change the outcome, or is too deeply coded (for example, as a summary grade given by the teacher) to lead to appropriate action, the control loop cannot be closed, and “dangling data” substituted for effective feedback [10].

Therefore, Bloom’s formulation is unhelpful in describing the information generated about the gap between current and desired performance as “feedback” Bloom separated the information from its instructional consequences. For Wiener, Ramaprasad, and Sadler, feedback is more than just information. It is information generated within a system, for a purpose [2].

Ralph Tyler in 1934, Proposed the use of the term “educational evaluation”, while education was created by objectives, which has the principle to formulate objectives and to verify that these were fulfilled [11].

Another review [12] proposed a model of the evaluation process as consisting

of eight stages:

- 1) Establishing the purpose of the evaluation;
- 2) Assigning tasks to students;
- 3) Setting criteria for student performance;
- 4) Settings standards for student performance;
- 5) Sampling information on student performance;
- 6) Appraising student performance;
- 7) Providing feedback to student performers; and
- 8) Monitoring outcomes of the evaluation of students.

3. The Evaluation Strategies Used as Testers of Pedagogical Skills

A Learning Assessment is key in Educational process, which requires a monitoring of the appropriation of the thematic and the construction of the knowledge of the student. In this sense, it is necessary to use pedagogical tools designed intentionally to follow the paths of multiple learning. The evaluation of learning presents three basic functions: diagnose (investigate), control (monitor) and classify (value). These functions are directly related to their modalities: diagnostic, formative and somative [13].

The somative evaluation, with qualifying function, students at the end of the unit, semester or school year, according to levels of use presented. The somative evaluation has a qualifying function at the end of the teaching process, with the bias of checking if there was learning, and with the intention of grading the student and is linked to approving or failing [13].

This relates to the product demonstrated by the student in situations previously stipulated and defined by the teacher, and materializes in the note, object of desire and suffering of the students, their families and even the teacher himself. This logic predominates the bureaucratic bias that impoverishes learning, stimulating didactic actions focused on the control of the activities carried out by the student, but not necessarily generating knowledge [14].

Units of measure that are expressed in numbers in general are used justifying their objectivity and accuracy, thus referring to the quantitative aspect of the phenomenon described [13]. As the quantitative conception of the evaluation, there are numerous criticisms about the models and practices of evaluation in our schools, with a rapid development of alternative evaluation approaches, with ethical assumptions, epistemological and very different theorists [15].

With the increase in interest in the evaluation and the need for adequate information about the appropriation of knowledge by the students, there was an interest in the insertion of other qualitative aspects, beyond the quantitative. This insertion was influenced by the epistemological movements against criticism of the traditional method and the understanding of the evaluation as something much more complex than approving or failing. The qualitative assessment aims to overcome the quantitative evaluation without dispensing with

this. It believes that in the educational area the processes are more relevant than the products, not living up to the reality, if reduced only to empirically measurable manifestations. These are easier to manipulate methodologically, because the scientific tradition has always privileged the measured treatment of reality, sometimes advancing in an incisive way in some social disciplines, such as economics and psychology. However, it is not possible to transfer the methodological limitation to the pretension reduction of the real. This is more complex and comprehensive than its empirical face. The qualitative assessment would like to reach the qualitative face of reality, or at least to approach it [16].

Traditionally associated with the school, the evaluation creates hierarchies of excellence. Students are compared and then classified by a standard of excellence, defined in absolute or embodied by the teacher and the best students [14].

A certification provides few details of the knowledge and skills acquired and the level of dominance precisely acquired in each field covered. It ensures, above all, that a student knows globally “what is needed to know” to move on to the next series in the course, be admitted into a qualification or start a profession. The benefit of an established certification is precisely that it does not need to be controlled point by point, to serve as a passport for employment or for further training [14].

The qualitative evaluation model is set up as a transitional model for having as centrality the understanding of the processes of subjects and of learning, which produces a rupture with the primacy of the characteristic result of the quantitative process, says [8].

The qualitative evaluation tries to respond to the imposition of the qualitative assessment to apprehend the dynamics and intensity of the learning-teaching relationship but articulated by principles that support the knowledge-regulation-market, state and community. Based on the logic qualifying and excluding in the classroom and in the proposals arriving at the school, the maintenance of the practice of evaluation based on the qualification logic is excluding, even if the Practice acquires an innovative appearance and that the concept of school evaluation associated with the quantification of the income of the student/A is subject to numerous and profound criticisms [17].

Had a rather narrower focus the impact of classroom evaluation practices on students. His review covered formal classroom-based assessments such as tests, informal evaluation processes such as adjunct questions in texts, and oral questioning by teachers in class. His main conclusion was that too much emphasis has been placed on the grading function of evaluation and too little on its role in assisting students to learn [18].

With the intention of the teacher to play the role of facilitator and not of appraiser [2], teachers engaged in this kind of feedback conveyed a sense of work in progress, heightening awareness of what was being under-taken and reflecting on it [19].

4. The Use of the Formative Evaluation on the International Stage as the Main Verification Tool in the Contemporary

The formative evaluation considers that the student learns throughout the process, which is restructuring his knowledge through the activities that he performs. From the cognitive point of view, formative evaluation focuses on understanding the functioning of knowledge construction [20]. The information sought in the evaluation refers to the mental representations of the student and the strategies used, to arrive at a certain result. The errors are objects of study, because they reveal the nature of the representations or strategies elaborated by the student. Formative assessment has been on policy agendas internationally for decades, but implementation has proven to be challenging. Although many researchers acknowledge that formative assessment can have a positive effect on learning, the proof for this is based on limited scientific evidence [20].

Formative assessment provides information about the learning process that teachers can use for instructional decisions and students can use in improving their performance. It provides too information about the learning process that teachers can use for instructional decisions and students can use in improving their performance, which motivates students [2].

Much of this strategies for implementing identified [21]:

- Clarifying and sharing learning intentions or goals and success criteria;
- Generating opportunities to effectively gather evidence of student learning through in-formal and formal assessment, through classroom discussions, questioning or learning tasks;
- Providing formative feedback to students to support their learning;
- Supporting students in acting as instructional partners through discussion and peer assessment; and
- Activating students as agents in their own learning through self-assessment and self-regulation.

The notion of formative assessment as being central to the regulation of learning processes has been adopted by some writers in the Anglophone community (see, for example, [22], and the broadening of the conception of formative assessment in the English-language literature [23]. Her review of the literature on “formative classroom assessment” charted the development of the conception of formative assessment as a series of nested formulations [2]. As points out, in an important critical review of the field, one cannot be sure about the effects of such changes in practice unless one has an adequate definition of what the terms formative assessment and assessment for learning mean, and a close reading of the definitions that are provided suggests that there is no clear consensus about the meanings of the terms formative assessment and assessment for learning [20].

Such assessment does not necessarily have all the characteristics just identified as helping to learn. It may be formative in helping the teacher to identify areas where more explanation or practice is needed. But for the pupils, the marks or

remarks on their work may tell them about their success or failure but not about how to make progress towards further learning [24].

Boekaerts has proposed a deceptively simple, but powerful, model for understanding self-regulated learning, termed the dual processing theory [25]. It is assumed that students who are invited to participate in a learning activity use three sources of information to form a mental representation of the task-in-context and to appraise it: 1) current perceptions of the task and the physical, social, and instructional con-text within which it is embedded; 2) activated domain-specific knowledge and (meta) cognitive strategies related to the task; and 3) motivational beliefs, including domain-specific capacity, interest and effort beliefs [26].

Depending on the outcome of the appraisal, the student activates attention along one of two pathways: the “growth pathway” where the goal is to increase competence or the “wellbeing pathway” where attention is focused on preventing threat, harm or loss. While the former is obviously preferable, the latter is not necessarily counter-productive by attending to the wellbeing pathway, the student may find a way to re-store well-being (for example by lowering the cost of failure) that allows a shift of energy and attention to the growth pathway [2]. To summarize, because learning is unpredictable, assessment is necessary to make adaptive adjustments to instruction, but assessment processes themselves impact the learner’s willingness, desire, and capacity to learn [27].

In a syncretic way, one can highlight the six stages observed in a formative assessment [28].

1) Choose. Ensure that teachers have some autonomy to decide which formative assessment strategies and practices to try to implement or how to approach learning accurately. By providing options, we can better respond to different levels of teacher readiness.

2) Flexibility. Encourage teachers to make changes to strategies to make them their own so they are applicable and relevant to their environment and students.

3) Small steps. Learning is incremental and takes time to change practice. To make lasting changes, support teachers with the time, resources, and training they need as they transfer new learning into their daily routines.

4) Responsiveness. The information we collect is nothing until we act on it. Help teachers not only extract evidence of learning, but also make responsive adjustments to their instruction based on that data. It is also important that the teacher teaches students to be responsive in using their own data.

5) Collegial support. Provide teachers with a space to collaborate with peers around formative assessment practices and time to meet. This gives teachers the opportunity to develop personal action plans, report to a group of colleagues on the outcome of implementing those plans and reflect and receive feedback from colleagues who are facing similar challenges.

6) Support responsibility. Teachers, like any professional, need to be held accountable for results. They need to be given the time and resources to make a meaningful change.

For assessment to support learning, it must provide guidance about the next steps in instruction and must be provided in way that encourages the learner to direct energy towards growth, rather than well-being [2].

5. Final Considerations

The evaluation process since the beginnings makes clear its role as a tool of selection and power by the ruling classes, starting with the Chinese emperors. The evaluation was historically constituted as an instrument dedicated to promotions, control and quantification. Despite numerous scientific and technological advances, in some episodes today, they still reaffirm their, albeit on paper picker and excluding. According to the epistemological historical development, a numerical quantification of the individual's learning capacity was required through docimology, psychometric and intellectual quoeficient.

The evaluation began to demonstrate pre-established goals in schools, based on the interest and the mounds of the dominant classes, that is, a set of previously set knowledge that should be learned and demonstrated, in a way Quantitative, and thus allow the classification as approved or reproved.

The evaluation of learning is something complex and we usually reduce to two polarities: the result and the learning process, the previous one referring to the achievements of the apprentice, the latter to the process by which they are attained. Evaluating the formative product means judging the results of the teaching and learning of the integrated process, whose effects can be controlled by considering the specific performances of the subject.

This research was developed based on a review of data in the literature, without the pretension of a strict historical deepening. As a future perspective, the formative evaluation will be discussed using a post-structuralist framework, based on the contributions of the philosopher Michel Foucault.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Esteban, M.T. (2002) Who Knows Who Is Wrong? Reflections on Evaluation and School Failure. 3rd Edition, DP&A, Rio de Janeiro.
- [2] Wiliam, D. (2011) What Is Assessment for Learning? *Studies in Educational Evaluation*, **37**, 3-14. <https://doi.org/10.1016/j.stueduc.2011.03.001>
- [3] Depresbiteris, L. (1989) The Challenge of Learning Assessment; From the Foundations to an Innovative Proposal. EPU, São Paulo.
- [4] de Landsheere, G. (1976) Continuous Evaluation and Exams: Notions of Docimology. 3rd Edition (Revised and Increased), Almedina, Coimbra, 330 p.
- [5] Piéron, H. (1973) Dictionary of Psychology. Globo Publishing House, Porto Alegre.
- [6] Pasquali, L. (2009) Psychometry. *Journal of Nursing School, University of São Paulo*, **43**, 992-999.

- [7] dos Santos, J.G. (2008) History of Evaluation: From the Examination to the Diagnostic Evaluation Federal University of Uberlândia. 4th Week of Service and 5th Academic Week.
- [8] Esteban, M.T. (Org.) (2003) School, Curriculum and Assessment. *Culture Series Memory and Curriculum*, Vol. 5, Cortez, São Paulo.
- [9] Guskey, T.R. (2010) Formative Assessment: The Contributions of Benjamin S. Bloom. In: Andrade, H.L. and Cizek, G.J., Eds., *Handbook of Formative Assessment*, Taylor & Francis, New York, NY, 106-124.
- [10] Sadler, D.R. (1989) Formative Assessment and the Design of Instructional Systems. *Instructional Science*, **18**, 119-144. <https://doi.org/10.1007/BF00117714>
- [11] Hoffmann, J.M.L. and Lerch, M. (1991) Evaluation and Construction of Knowledge. *Education and Reality*, **16**, 67.
- [12] Natriello, G. (1987) The Impact of Evaluation Processes on Students. *Educational Psychologist*, **22**, 155-175. https://doi.org/10.1207/s15326985ep2202_4
- [13] Haydt, R.C. (2000) Evaluation of the Teaching-Learning Process. Attica, São Paulo.
- [14] Perrenoud, P. (1999) Evaluation: From Excellence to Regulation of Learning—Between Two Logics. ArtesMédicasSul, Porto Alegre.
- [15] Saul, A.M. (1988) Emancipatory Evaluation: Challenge to Theory and Practice of Curriculum Evaluation and Reformulation. Cortez, São Paulo.
- [16] Demo, P. (2004) Theory and Practice of Qualitative Evaluation. *Themes of the 2nd International Congress on Education Assessment*, Curitiba, 156-166.
- [17] Esteban, M.T. (2001) What Do You Know Who Is Wrong? Reflections on School Failure and Evaluation. DP & A, Rio de Janeiro.
- [18] Crooks, T.J. (1988) The Impact of Classroom Evaluation Practices on Students. *Review of Educational Research*, **58**, 438-481. <https://doi.org/10.3102/00346543058004438>
- [19] Tunstall, P. and Gipps, C.V. (1996) Teacher Feedback to Young Children in Formative Assessment: A Typology. *British Educational Research Journal*, **22**, 389-404. <https://doi.org/10.1080/0141192960220402>
- [20] Bennett, R.E. (2009) A Critical Look at the Meaning and Basis of Formative Assessment (ETS RM-09-06). Educational Testing Service, Princeton.
- [21] William, D. and Thompson, M. (2008) Integrating Assessment with Instruction: What Will It Take to Make It Work? In: Dwyer, C.A., Ed., *The Future of Assessment: Shaping Teaching and Learning*, Lawrence Erlbaum Associates, Mahwah, 53-82.
- [22] William, D. (2007) Keep Learning on Track: Classroom Assessment and Learning Regulation. In: Lester, F.K., Ed., *2nd Manual of the Teaching and Learning of Mathematics*, Information age Publishing, Greenwich, 1053-1098.
- [23] Brookhart, S.M. (2007) Expanding Views about Formative Class-Room Assessment: A Review of the Literature. In: McMillan, J.H., Ed., *Formative Classroom Assessment: Theory into Practice*, Teachers College Press, New York, 43-62.
- [24] Broadfoot, P.M., Daugherty, R., Gardner, J., Gipps, C.V., Harlen, W., James, M., et al. (1999) Assessment for Learning: Beyond the Black Box. University of Cambridge School of Education, Cambridge.
- [25] Brookhart, S.M. (2004) Classroom Assessment: Tensions and Intersections in Theory and Practice. *Teachers College Record*, **106**, 429-458. <https://doi.org/10.1111/j.1467-9620.2004.00346.x>

- [26] Eccles, J.S and Wigfield, A. (2002) Motivational Beliefs, Values, and Goals. *Annual Review of Psychology*, **53**, 109-132.
<https://doi.org/10.1146/annurev.psych.53.100901.135153>
- [27] Harlen, W. and Deakin-Crick, R. (2002) A Systematic Review of the Impact of Summative Assessment and Tests on Students' Motivation for Learning. In: EPPI-Centre, Ed., *Research Evidence in Education Library*; University of London Institute of Education Social Science Research Unit, London, 153.
- [28] Dyer, K. (2016) Six Steps to Formative Assessment Success in the Classroom. Teach. Learn. Grow. The Education Blog.