

Book reviews

Joint Committee on Standards for Educational Evaluation (1994, 2003, 2009). *The program evaluation standards; The student evaluation standards & The personnel evaluation standards.*

The three sets of standards published by the Joint Committee on Standards for Educational Evaluation (JCSEE) are the result of years of dedication and effort by the leading educational institutes (both scientific and professional) of the United States (and Canada too). The consensus they embody and the vastness of their scope — a country which occupies a leading place on the global scientific and professional stage — make them an essential reference not just for those directly involved in educational evaluation but for everyone affected by the latter.

Since the first standards were published in 1981, evaluation associations all over the world have introduced orientations, principles and professional standards based on, or inspired by, them (Simons, 2006). As this author notes, the JCSEE standards are the most widely published standards of their kind in the world. In Europe, for instance, they have even been adopted / adapted by the Swiss and German Societies for Evaluation; and in Oceania they have informed the debate on professional codes of conduct in the Australasian Evaluation Society.

THE JOINT COMMITTEE ON STANDARDS FOR EDUCATIONAL EVALUATION — ORIGINS AND ACHIEVEMENTS

In 1974, a committee jointly appointed by the American Educational Research Association (AERA), the American Psychological Association (APA) and the National Council on Measurement in Education (NCME) completed its revision of the 1966 edition of the *Standards for Educational and Psychological Tests*

and *Manuals*, published by the American Psychological Association (Joint Committee on Standards for Educational Evaluation, 1994). This committee felt that the inclusion of a section on evaluation in test standards lay outwith its own remit, and recommended the creation of another committee to address this issue. The three organizations therefore appointed another committee, which met for the first time in autumn 1975, with a membership extending across twelve national organizations with an interest in the quality of evaluation in education.

The first edition of the *Standards for Evaluations of Educational Programs, Projects and Materials* (Joint Committee on Standards for Educational Evaluation, 1981) was published in 1981. Its revised version, from 1994, is the one to which we refer here. In 1988 came the first edition of *The Personnel Evaluation Standards* (Joint Committee on Standards for Educational Evaluation, 1988), which appeared in a revised edition in 2009, while *The Student Evaluation Standards*, which primarily address the internal, everyday evaluation carried out by teachers in schools, appeared in 2003. External, large-scale evaluation continued to be principally covered by the *Standards for Educational and Psychological Testing*, revised in 1999 (Joint Committee on Standards for Educational Evaluation, 2003).

All of these standards were the result of a rigorous process of elaboration and testing which drew on the input and contributions of sources including a panel of writers, review panels, field test sites, public hearings and a validation panel, and were subjected to periodical reviews which allowed them to constantly incorporate technical and scientific advances and to respond to the new challenges and requirements emerging in the field of evaluation.

THE NATURE, STRUCTURE AND IMPORTANCE OF EVALUATION STANDARDS

The standards define the principles which have been consensually accepted by evaluation professionals, and when observed they guarantee quality of evaluation and suitable professional practice. They are not, however, an instruction manual, a list of specific technical standards or rules to be applied mechanically, as a matter of course, but rather articulate a set of guidelines which are not equally important or applicable in all situations, and which may even, in certain situations, enter into conflict with one another. Their application requires qualification and professional judgement, and of themselves the standards are no replacement for a specialist bibliography, training in evaluation, professional experience, in-depth analysis and critical reflection if in each individual case we are to strike a balance which does not jeopardize any of the four major attributes they comprise. The utility of evaluation also includes an “evaluator qualification” standard, necessary for evaluation which is of good quality and recognized as such — its general acceptance is an essential condition for its inclusion in a decision plan whose agenda requires, at the outset, the presence of the evaluation factor. Consequently, although they represent the state of the art in research into evaluation and contribute to improving its quality, the standards do not of themselves guarantee this quality (Joint Committee on Standards for Educational Evaluation, 1994), as anyone who reads them will appreciate. They provide, instead, a framework of reference that reflects good practice in evaluation. Their use requires the mobilization of knowledge which they embody but which itself transcends them.

The books which comprise the standards, and each standard in itself, are however arranged according to a structure which significantly facilitates the comprehension of the principles presented (and which has improved in clarity if we compare the 1994 edition with those of 2003 and 2009). After a general presentation of their institutional context, development process, method of application and potential uses, the various standards are presented in detail (by groups), with a brief statement/description/definition/explanation preceding the exposition of the guidelines (which are neither exhaustive nor to be unquestioningly applied to all situations) for implementation of the standard. This is followed by an examination of the errors most frequently committed and which can compromise the quality of evaluation in particular situations, and by accounts of cases taken from real-life situations which illustrate how the standard was or was not applied, the conditions in which the case occurred, the consequences and how they could have been avoided, and the actions which could have been taken to improve the application of the standard.

The books therefore provide a broad, rich, diversified, realistic and insightful perspective on the complexity, intensity and difficulties inherent to the evaluation activity in all its phases, planes and dimensions, and of the constraints, determining factors, challenges and requirements which the standards help to address, articulate, weigh up and balance in the decisions made in the evaluator’s work and in the evaluation process.

The standards help evaluators to plan their evaluation, to select from alternative evaluation plans (in accordance with the available and necessary resources), and to monitor, regulate and control the implementation of evaluation. They also provide pointers on the appraisal of evaluation reports and the suitability and global quality of evaluation, and meta-evaluation; and together they constitute a sourcebook for research into evaluation which undoubtedly contributes to a general improvement in evaluation as an activity. Consequently, they are also a valuable resource for evaluator training and qualification.

THE FOUR CATEGORIES OF EVALUATION STANDARD

The three sets of standards, which largely share the same attributes (as can be seen in Table 1) across the three primary domains of evaluation practice — 1) personnel; 2) student; 3) program — are predicated on four major categories for a high-quality evaluation: 1) propriety; 2) utility; 3) feasibility and 4) accuracy. These are the expression of a unified and consistent perspective on educational evaluation and they vindicate a joint analysis of the three publications of the JCSEE. And although directed at different audiences, they should be used together (Joint Committee on Standards for Educational Evaluation, 2003). There is no shortage of examples, after all, of the interpenetration of evaluation of learning, evaluation of the performance of education professionals, and evaluation of programs (projects and policies).

Furthermore, each standard may encompass more than one of the four attributes (despite its being categorized according to its primary emphasis), while the different standards often overlap as they are inter-related and complementary: for evaluation itself is by nature a complex and system-based activity. Despite the apparent omissions in Table 1, a global reading of the data it presents reveals every standard is present in the guidelines for high-quality evaluation (even if the emphasis is placed on the four major attributes of evaluation and not on the standards considered in isolation).

Although they were developed in the USA and are based on the ideas, laws, education system and circumstances proper to that country, these standards articulate a practical philosophy of evaluation which has gained

TABLE 1
THE THREE SETS OF STANDARDS

STUDENT EVALUATION STANDARDS	PERSONNEL EVALUATION STANDARDS	PROGRAM EVALUATION STANDARDS
Propriety Standards		
Service to Students	Service Orientation	Service Orientation
Appropriate Policies and Procedures	Appropriate Policies and Procedures	Formal Agreement
Access to Evaluation Information	Access to Evaluation Information	Disclosure of Findings
Treatment of Students	Interaction with Evaluatees	Human Interaction
Rights of Students	Legal Viability	Rights of Human Subjects
Balanced Evaluation	Comprehensive Evaluation	Complete and Fair Assessment
Conflict of Interest	Conflict of Interest	Conflict of Interests
		Fiscal Responsibility
Utility Standards		
Constructive Orientation	Constructive Orientation	
Defined Users and Uses	Defined Uses	Stakeholder Identification
Information Scope		Information Scope and Selection
Evaluator Qualifications	Evaluator Qualifications	Evaluator Credibility
Explicit Values	Explicit Criteria	Values Identification
Effective Reporting	Functional Reporting	Report Clarity
		Report Timeliness and Dissemination
Follow-up	Follow-up and Professional Development	Evaluation Impact
Feasibility Standards		
Practical Orientation	Practical Procedures	Practical Procedures
Political Viability	Political Viability	Political Viability
Evaluation Support	Fiscal Viability	Cost Effectiveness
Accuracy Standards		
Validity Orientation	Valid Judgments	Valid Information
Defined Expectations for Students	Defined Expectations	Program Documentation
Context Analysis	Context Analysis	Context Analysis
Documented Procedures	Documented Purposes and Procedures	Described Purposes and Procedures
Defensible Information	Defensible Information	Defensible Information Sources
Reliable Information	Reliable Information	Reliable Information
Bias Identification and Management	Bias Identification and Management	Impartial Reporting
Handling Information and Quality Control	Systematic Data Control	Systematic Information
Analysis of Information	Analysis of Information	Analysis of Quantitative Information
		Analysis of Qualitative Information
Justified Conclusions	Justified Conclusions	Justified Conclusions
Meta-evaluation	Meta-evaluation	Meta-evaluation

Note: The sequencing of the standards has been adjusted to allow better alignment of the three sets. In the list of program evaluation standards utility precedes feasibility, probably because those involved are not so central to evaluation, and also because the evaluation of programs is more problematic (perhaps because the relationship between evaluation and decision-making is less structured on the institutional level than in the other two cases, characterized by regular and perennial decision-making cycles).

universal acceptance in the western world — with the promotion and safeguarding of the quality of educational services as the ultimate objective. The first two standards presented and examined by the JCSEE, therefore, are those relating to the propriety and the utility of evaluation.

With the propriety standard the aim is to ensure that evaluation is conducted ethically, legally and with respect for the well-being of all those involved in it and affected by it. “Service orientation” is the key concept, according to which the satisfaction of the educational needs of the student, and by extension of the community and society in general, is paramount. “Conflicts of interest” must therefore be avoided, or managed in such a fashion that the evaluator is independent and impartial, neither benefiting nor suffering from any result which evaluation may produce.

The aim of the utility standards is to ensure that evaluation is applied in a clear and timely fashion (providing a response to the needs for information of the users), and as a “constructive guideline” which informs the recommendation, planning (including implementation), supervision and evaluation of follow-up actions designed to consolidate or develop strengths while eliminating, correcting or improving weaknesses — the “impact of evaluation”.

Feasibility standards are designed to ensure that the political and material conditions exist for evaluation to be implemented as intended. This requires diplomacy, and procedures which do not interfere with educational activity, which are practical/practicable and which can mobilize the necessary resources.

Accuracy standards address the production of reliable and representative information which permits valid interpretations, justified conclusions, and appropriate follow-up actions. In this context “meta-evaluation” is of prime importance.

As we can see, each group of standards addresses an essential aspect of evaluation. But all of these aspects are strongly inter-dependent, and this means we have to take all aspects into consideration in each particular evaluation.

CONCLUSIONS: THREE OBLIGATORY WORKS FOR EVALUATION IN EDUCATION

Even when properly situated in the context (educational, legal, social, political and economic) from which they emerged, the three works we have examined above are an

indispensable source of reference for all those interested in evaluation, especially those whose professions require them to address the issue.

BIBLIOGRAPHICAL REFERENCES

JOINT COMMITTEE ON STANDARDS FOR EDUCATIONAL EVALUATION (1981). *Standards for evaluations of educational programs, projects and materials*. New York: McGraw-Hill.

JOINT COMMITTEE ON STANDARDS FOR EDUCATIONAL EVALUATION (1988). *The personnel evaluation standards. How to assess systems for evaluating educators*. Newbury Park: Sage Publications.

JOINT COMMITTEE ON STANDARDS FOR EDUCATIONAL EVALUATION (1994). *The program evaluation standards, 2nd edition. How to assess evaluations of educational programs*. Thousand Oaks: Sage Publications.

JOINT COMMITTEE ON STANDARDS FOR EDUCATIONAL EVALUATION (2003). *The student evaluation standards. How to improve evaluations of students*. Thousand Oaks: Corwin Press, a Sage Company & Educational Policy Leadership Institute.

JOINT COMMITTEE ON STANDARDS FOR EDUCATIONAL EVALUATION (2009). *The personnel evaluation standards. How to assess systems for evaluating educators, 2nd edition*. Thousand Oaks: Corwin Press, a Sage Company.

SIMONS, H. (2006). Ethics in Evaluation. In I. F. SHAW; J. C. GREENE & M. M. MARK (eds.), *The Sage Handbook of Evaluation*. London: Sage, pp. 243-265.

PEDRO RODRIGUES

prodrigs@ie.ul.pt

Institut of Education of the University of Lisbon

Translated by Mark Ayton

Rodrigues, Pedro (2009). Review of “The program evaluation standards”; “The student evaluation standards” & “The personnel evaluation standards”, by the Joint Committee on Standards for Educational Evaluation [1994, 2003, 2009]. *Sísifo. Revista de Ciências da Educação*, 09, pp. 113-116.

Consultado em [mês, ano], em: <http://sisifo.fpce.ul.pt>