COMMISSION INTERNATIONALE DE L'ENSEIGNEMENT MATHÉMATIQUE (THE INTERNATIONAL COMMISSION ON MATHEMATICAL INSTRUCTION)

L'Enseignement Mathématique, t. 36 (1990), p. 197-206

COMMISSION INTERNATIONALE DE L'ENSEIGNEMENT MATHÉMATIQUE (THE INTERNATIONAL COMMISSION ON MATHEMATICAL INSTRUCTION)

ASSESSMENT IN MATHEMATICS EDUCATION AND ITS EFFECTS

DISCUSSION DOCUMENT

The International Program Committee appointed by the International Commission on Mathematics Instruction (ICMI) announces a study on "Assessment in Mathematics Education and its Effects".

The study consists of two components, a *conference,* to be held in Spain, 11-16 April, 1991, and a *volume,* to be published in the ICMI Studies series based on (the contributions to and products of the outcomes of) the conference. This discussion document: (1) provides background on the study, and outlines its aims, scope, and issues; (2) announces a call for papers; and (3) provides preliminary practical and organizational information to potential contributors to the study.

BACKGROUND AND OUTLINE OF THE STUDY

WHY AN INTERNATIONAL STUDY ON ASSESSMENT?

Each ICMI study is conducted to contribute to the understanding and tackling of a specific topic or problem area which is of current importance to mathematics education in countries in different parts of the world.

Why is *assessment* of current importance to mathematics education in an international context?

The Program Committee believes that in many countries difficult questions, serious problems, and major challenges exist with respect to assessment in mathematics education. The committee also believes that many current assessment practices are counter-productive, the products of (and supports for) out-

moded educational traditions which fail to meet today's views about mathematical and societal needs.

These challenges would *not* exist:

— if society were not experiencing rapid and substantial changes, and if these changes were not in turn bringing about changes in mathematics instruction in ways that call for new approaches to assessment;

— if the roles, functions, and effects of contemporary methods of assessment were completely clear and well understood;

— if there were no divergent aims or conflicting interests and no unintended, undesired, or sometimes even dangerous side- effects involved in the current methods of assessment;

— if we had devised and implemented methods of assessing, in a valid and reliable way, the essential knowledge, insights, abilities, and skills related to mathematics and its place in the world; and

— if current assessment procedures could provide genuine assistance to (a) the individual learner in monitoring and improving his or her acquisition of mathematical insight and power, (b) the individual teacher in monitoring and improving his or her teaching, guidance, supervision, and (c) textbook authors, curriculum planners and authorities, in adequately shaping the framework of mathematics instruction.

Unfortunately, the Program Committee is convinced that most of these conditions are problematical in schools, colleges and universities throughout the world today. Failure to confront these issues raises important and significant questions which must now be addressed.

In brief: if, in mathematics, it were easy to employ effective, harmonious, assessment procedures free from serious internal or external problems, an ICMI study on assessment would not be relevant. Nor would it be relevant if the difficulties were present only in a few isolated nations, for then an international perspective would not be justified.

It has become clear, however, that serious questions about mathematics assessment are being raised in countries throughout the world. Although we recognize that the purposes, roles, functions, and practices of assessment may be viewed very differently in different educational systems and in different societies, there is no doubt that the crucial questions, problems, and challenges now being addressed are very similar worldwide. Thus, the 1991 conference and volume should be of common interest to researchers and practitioners of mathematics education all over the world. What faces mathematics educators are

matters and issues about assessment of a fundamental rather than of an incidental nature. Thus, the goals of the study are: to present and examine current assessment practices in many nations; and to identify examples, practices, and ideas that will enable assessment to become universally a positive influence on instruction by contributing to link together the purposes, implementation and outcomes of any mathematical program.

THE FUNCTION OF ASSESSMENT

In every educational system, the purpose of assessment in mathematics education is to provide information, gathered in a specific manner, for a specific constituency, about the mathematical performance of a student or a group of students. The objective of gathering this information is to assist "someone" to make decisions about students, teachers, or programs. The information sought may be relative (i.e., it may compare one outcome with a larger sample of outcomes), or absolute (i.e., it may determine the quality of the particular outcome). Finally, the information may be reported in either a qualitative or a quantitative form.

What varies in different assessment approaches is:

(1) the kinds and forms (relative, absolute; qualitative, quantitative) of the information to be collected;

(2) the way of gathering such information;

(3) the professional position of those who gather the information;

(4) the "someone" receiving and using the information;

(5) the types of outcome performances;

(6) the unit of aggregation of the information (individual, group, class, cohort, nation);

(7) the types of decisions or actions that might be taken as a result of the information; and

(8) the students, teachers or programs that are the objects of the decisions or actions to be taken.

The possible variations in each of those components resulting from combining the diverse approaches encountered in actual practice is, of course, tremendous. Yet, we are dealing only with variations on a theme. And it should not be forgotten that the salience of certain elements tends to set the entire character of the assessment approach adopted.

Traditionally, the purpose of student assessment in mathematics education has been to determine whether a given individual should be granted access to certain privileges. A typical privilege is *either* a "license" to practice a profession or vocation based on mathematics — for instance, as a surveyor, an actuary, or a teacher of mathematics in a school or university — or, more importantly, a "ticket" to further education in the next grade, in a new subject, in a new institution, for which achievements in mathematics have been made part of the entrance requirement. An examination results in a verdict: passed or failed, admit or deny admission. It is typical of this conception of assessment that an evaluation is not being made of the assessment procedures themselves, nor of the curriculum, the institution, the textbooks, the instruction, the teachers, the assessors, and so forth. Only the individual student, viewed as an *object,* is being judged, and decisions or actions concerning the student are the focus of attention.

This function of assessment as a tool for the selection or placement of people continues to be strong, even dominant, in most educational systems throughout the world. However, a broader conception of assessment has begun to emerge in recent decades. Information about student performance is now being used to judge other components.

Information about the outcomes of mathematics instruction is now being sought for the general purpose of informing and guiding teaching practice, teacher performance, and curriculum development. Given this development, an important question then is: *Does mathematics instruction function satisfactorily in relation to different groups of learners?*

If not, where are the problems? With the students? With their family backgrounds? With the teacher(s)? With the institutional environment? With the textbooks? With the curriculum or program? So, we are led to an ensuing four-fold question: What instructional procedures, under what conditions, and for what students, are effective in achieving what types of learning in mathematics?

If the answer is yes, are further improvements possible? The "someone" who is, or should be, interested in knowing the answers to these questions varies, ranging from the individual student to local or national curriculum authorities and administrative and government agencies. The system which attempts to obtain the answers varies in a similar manner. The decisions or actions to be taken as a result of the information gathered are not limited to addressing the individual student. They may address any element in the spectrum of components which constitute mathematics instruction.

First of all, the learner is no longer considered as merely an object of assessment, but as an *autonomous individual* too, as an individual with his or her own

dynamic to grow and develop, a person who has the right to demand a certain quality from the mathematics instruction he or she is receiving. In modern society, mathematical competence is essential not only to obtaining access to careers (whether educational or vocational/professional), but it is also crucial in exercising active, responsible citizenship. The learner may want to insist that mathematics instruction should provide such competence as well. When considering the individual student as an independent person, there is a growing trend in assessment to respect the integrity of what the student knows, how the student knows, and why he or she seeks to known. Mathematical knowledge is being increasingly viewed as a working force in the life and being of the student and has to be assessed as such.

The outcomes of mathematics instruction are now regarded as having much wider implications than before. More emphasis has been put on students' ability to actively and creatively deal with mathematical ideas, concepts, topics, problems, and issues within mathematics itself as well as in extra-mathematical contexts. Problem posing and problem solving, modeling and applications, open-ended situations, investigations, scientific debate, and so forth, have been introduced in mathematics instruction in many areas and at many levels. Briefly, it has become important to give students as much opportunity as possible to engage in the same *kind* of activities and processes, although not at the same level, as mathematical professionals in a broad meaning of that term: thinking and acting mathematically. So, mathematics instruction is being given many more dimensions than before.

When the outcomes of mathematics instruction are perceived in a wider way, the information about them also must be perceived in this wider way. The crucial question here is: To what extent do we possess the means to obtain valid, authentic information about student performance? Although in recent years much work has been done to acquire more valid information about the outcomes of mathematics instruction, the methods of gathering such information lag far behind the need for it. This is due not only to inertia in educational systems but also to insufficient ingenuity, research, and development in the field, and to insufficient resources for creating new methods of assessment.

Finally, we should not forget that any system of assessment strongly influences, for better and for worse, the educational system in which it is embedded. The way mathematics instruction functions, as well as the entire spirit in which it takes place, is strongly influenced by assessment methods. Assessment is not just a separate appendix to mathematics instruction; it is one of its crucial components.

Scope of the Study

The changing perception of assessment outlined above encompasses, in principle, each agent and component of mathematics instruction as an actual, and potential, object of assessment. In this context, the terms *assessment* and *evaluation* are often used interchangeably without a clear difference in meaning being made between them. The Program Committee suggests that *assessment* be used to refer to the outcome of mathematics instruction as reflected in the performance of *students* as individuals or in groups, whereas *evaluation* should deal chiefly with the use of such performance information to make judgements about instructional *programs, curricula,* and appraisal of teachers. In order to obtain a clear focus of the present study, and in order to limit it to a tractable size, the dominant emphasis will be on *assessment* as just defined, but within the wide conception of outcomes outlined in the previous section. This limitation does not imply that issues related to evaluation are to be left out of consideration in the study. However, the evaluation of programs, curricula, and teachers will primarily be addressed as they are reflected in the assessment of students' performance. To avoid possible misunderstandings of this: it is the firm view of the Program Committee that programs, curricula, and teachers should not be judged solely on the basis of student performances.

Problems and Issues

The problems and issues to be addressed in the study conference concentrate on the various purposes, roles, and functions of assessment of students' mathematical performance. In particular, to contrast practices among countries, the following fundamental questions about assessment procedures and their uses need to be addressed:

— What are the significant historical developments in the philosophy and evolution of assessment and evaluation?

— For what purposes is information about students' mathematical performances being gathered? (To help teachers make instructional decisions? To assist students in monitoring and controlling their own learning process? To select or place students? To evaluate the effects of new programs?

— What are the units on which information is being aggregated - the individual student, group, class, teacher, institution, program, system?

— For what kinds of mathematical tasks are students' perfomances being assessed (short technical exercices, long tasks, extended problems, portfolios, project reports)? And what kind of information is being gathered

(examination of written items, of oral responses or responses to oral questions; observation of perfomance)?

— Who gathers the information?

— How is information gathered, coded, and recorded?

— How is the coded data aggregated and analyzed?

— What kinds of decisions or actions are taken on the basis of the information gathered?

— Are new procedures being developed/trid out?

— Are there conflicting views or interests between different segments of the educational system in relation to assessment and evaluation (e.g., between government authorities and mathematics educators)?

— What are the important differences in the assessment practices of different countries?

— How *useful* are international perfomance comparisons?

While the questions above serve primarily descriptive purposes, the following questions focus on the analysis of different modes of assessment with particular regard to their influences and effects:

— What are the theoretical and empirical foundations of current assessment procedures, and to what extent are these procedures valid, reliable, efficient?

— What are the known influences of external assessment practices on mathematics instruction?

— Are there examples of assessment practices which are known to influence instruction positively? What aspects should be maintained and encouraged?

— Are there examples of assessment practices which negatively influence instruction; for example, by focusing instruction on assessment and tests rather than on more general goals?

— How do different assessment modes influence the social environment in the classroom?

— In what respects are teachers good or bad judges of student perfomance? And to what extent cant they be trained to be good judges?

— How does the teacher's assessment role conflict with his/her supportive role?

— In many countries, university professors are considered professionally capable of assessing students justly, validly, and accurately, whereas school teachers are not; thus, external examinations are judged necessary. Does this make any sense?

— It is widely recognized that most current assessment practices deal mainly with independent facts and skills. Many of these practices have a high degree of reliability but a low degree of validity. To what extent is it possible to devise assessment modes which are both valid and reliable? How well do we assess authentic abilities such as the capacity for scientific debate, problem formulation, and problem solving, modelling, application, etc?

— What assessment modes are suitable in relation to different types of tasks, such as short technical exercices, long tasks, extended problems, project work, etc?

— How can assessment be embedded harmoniously into instructional practices as an instrument to serve the needs of both teachers and students in everyday instruction perfoming tasks in natural situations and in contexts psychologically close to the learner, not in isolation; respecting the cultural setting of mathematics; making use of a variety of ways of accomplishing specific tasks?

— What are the main obstacles to devising and implementing innovations in assessment, and what can be done to overcome these obstacles?

Structure of the Study

To accommodate a thorough treatment of the issues listed above, the study will contain *four* sections.

1. A *descriptive* section:

In this section, the most important conceptions and modes of assessment practiced in different countries or educational systems will be identified and described. Emphasis will be on *archetypes* rather than on peculiarities.

2. An *analytic* section:

This section will establish a framework for analyzing goals, functions, effects, consequences, limitations, possibilities, difficulties, and problems, related to the assessment of students' mathematical capabilities. By means of this framework, specific analyses regarding the issues listed above will be carried out. Furthermore, important empirical or theoretical research contributions to the field will be presented.

3. A selection on the presentation and discussion of *innovative/experimental* cases: In many places around the world, very interesting innovative/experimental work on new modes of assessment and evaluation has been or is being done. This section will present and discuss a number of the most interesting examples.

4. A *statement* section:

This section will present a formulation of policy statements and recommendations if appropriate. At the least, it will serve to identify important open issues as well as objectives for future research and development.

CALL FOR PAPERS

Given this background, the International Program Committee invites individuals and groups to propose or submit contributions to the study for consideration by the Committee. Contributions should be related to the problems and issues identified in the present document and fit into the least one of the four sections of the study just described. Participation in the conference is *only at the invitation* of the Program Committee, but those who submit a contribution are encouraged to apply for an invitation

In addition to calling for papers, the Program Committee will solicit contributors to address specific questions, to present research contributions, and to share examples of innovative work at the conference. We invite suggestions on topics and names of potential contributors. Also, comments regarding the structure of the conference will be welcome

The Program Committee will meet in November 1990 to make major decisions about the conference program. For matters regarding the program, please contact:

Professor Mogens Niss
IMFUFA, Roskilde University
P.O. Box 260
DK 4000 Denmark
Tel: +45 46757711 ext. 2266
Fax: +45 46755065

PRACTICAL INFORMATION

The study conference will take place in:

Calonge (Costa Brava), Spain, April 11-16, 1991.

The number of participants will be limited to about 75. The local organization of the conference will be taken care of by the *Federación Española de Sociedades de Profesores de Matemàticas* and the local organizer will be:

Professor Claudi Alsina
Secció Matemàtiques, ETSAB
Universitat Politècnica de Catalunya
Diagonal 649
E 08028 Barcelona
Spain

The International Program Committee consists of:

Claudi ALSINA, Local Organizer, Universitat Politècnica de Catalunya, Barcelona, Spain

Desmond BROOMES, University of the West Indies, Bridgetown, Barbados

Hugh BURKHARDT, Shell Centre for Mathematical Education, University of Nottingham, UK

Mogens NISS, Chairman, Roskilde University, Denmark

Thomas A. ROMBERG, National Center for Research in Mathematical Sciences Education, University of Wisconsin, Madison, USA

David ROBITAILLE, University of British Columbia, Canada

Julia SZENDREI, O.P.I., Budapest, Hungary