TESTS AND MEASUREMENT PROCEDURES, REVIEW AND EVALUATION.*

G.J. Matys
Curriculum Research and Development Unit
(Measurement and Evaluation Section),
Ministry of Education, Ghana.

PART I

TESTS AND MEASUREMENT PROCEDURES:

To a considerable extent modern education is characterised by the emphasis it places on adapting the educational programme to the needs of the individual child. Since these needs are governed by the child's level of ability and by the degree to which he has mastered the educational contents to which he has previously been exposed, it is important to determine these factors as accurately as possible. Once this information has been obtained, much can be done to individualize the educational process for each child, in part by grouping children into homogeneous instructional groups and in part by differentiating instruction within the classroom. It follows then that accurate educational measurement is a prime and key factor in modern educational trends.[1]

So if in modern education the emphasis is on teaching the individual it follows that we must have knowledge of, and differentiate between, individuals. Knowing the individual requires evaluating and testing. At the same time everywhere in the world the role of testing and evaluation in education is being questioned, criticized and scrutinized as never before.

Now before one goes further into the subject a definition or two should be made. The first is that Educational Evaluation is much broader than "testing". Educational Evaluation uses a variety of methods to measure and assess. These include questionnaires, surveys, cumulative records, projects, class work, oral answers, role playing and so on. Secondly the tools of educational evaluation are not a precise measure as are the tools of the engineer and physical scientist. In education one is trying to measure aptitudes, or intelligence, or content and processes - all very intangible and very difficult to assess. So it must be taken as a premise that the best of tests, under the best of conditions, provides large factors of error. When conditions (the training of personnel, administration and the tests themselves) are not ideal, results are even less reliable and less meaningful.

Testing had a traditional and fixed role for many years. It marked the end of one phase and the beginning of another. It meant passing out of one grade and into another. Successful completion of examinations allowed one to enter a profession, such as medicine. This type of testing has something in common with the initiation ceremonies characteristic of many non-Western societies and of various secret or exclusive groups within Western society. The examination or the initiation ceremony is a more or less

---

(1) Test Service Notebook - Test Bulletin, Harcourt Brace and World Inc.

difficult procedure; if the examinee reaches a certain level of performance, a level agreed upon by the elders of the society, then his status becomes altered and he becomes permitted to practice as a doctor: the high school leaver is enabled to seek employment. The characteristic of this type of examination is that it marks the end of one phase in the person's life, and it demonstrates that he is competent to enter a new phase. It is perhaps worth making another distinction between (a) the terminal examination in a particular vocation, which is intended as one indicator of the individual's competence (and the examination should not be the only indication which is used), and (b) the final examination at the end of a non-vocational course. The former should ensure that society is not plagued with incompetent doctors and other professionals, but the latter has a less clear social raison-d'etre, and would appear to be a potent agent in the development of a society in which every adult's status is determined by his scores on tests taken during his adolescence.

The second reason advanced for examining is as part of a continuous process of education, a method by which the teacher assesses what each student has and has not learned. This use of examinations has been going on for many years and it might seem hardly to deserve comment, - but the present renewed interest in "Measurement and Evaluation" seems to stem from a more precise analysis of the assessment process than used to be practised.

In modern education this first type of testing plays a less and less important role. Testing now is seen as one type of evaluative procedure and only a part of the overall educational process. There was a period recently when some advanced countries had an almost religious faith in tests, whether they were tests of aptitude, ability, achievement or intelligence. Test results were felt to provide final, definite, and reliable answers to many questions. In another, earlier, period educators ignored and discounted tests and their results as useless. It was felt that one could not measure "intelligence" and other "intangible" processes of man.

Today a middle-of-the-road approach which avoids either of the above extremes is gradually emerging. Educators are realising that tests are far from useless, and yet far from providing all absolute answers. It is realised that tests are only one factor, one piece of information which becomes valuable when combined with school marks, common-sense evaluations of teachers and a multitude of other data, some scientific, some less so. Also it must be kept in mind that measurements are only tools, a means to an end, and not an end in itself. We do not weigh or measure an article just for the sake of knowing how heavy or how long it is. We use this knowledge in some way. So it is with testing - there must be some definite purpose in the testing we do. Are the tests simply to measure achievement? Are they diagnostic tests only, to be used to diagnose teaching weaknesses and learning difficulties? Are the tests to measure a person's potential for academic or other fields, his capacity or aptitude? Do the tests try to measure interest, attitude, intelligence or personality? There are instruments available today which attempt each of these, or some combinations of these, and so the purpose must be clear from the start.

Generally the types of tests listed above are Standardized Tests - tests built by experts over long periods of time and with carefully selected norms and so on. However, both these and teacher-constructed tests have a role or purpose in the classroom. What are some of the purposes of classroom testing?

Let us examine some of the purposes and uses of such testing (this applies whether the tests are "teacher constructed" or "standardized"):

(1) To test pupils' achievement. This is probably the most common purpose of testing. The teacher should have constant feed-back on how well the skills taught have been mastered and how well the concepts and understandings can be applied. However, the diagnostic aspects of achievement tests should not be overlooked at any time.

(2) To assess the effectiveness of instruction. Educators are so prone to say, when looking at the results of a test, that the pupils have done "well" or "poorly". Frequently the results of a test are more an indication of how well the teaching has been done. If the results of a test show weaknesses, the teacher has the opportunity to re-teach, change the method of approach, or seek for other methods of increasing the effectiveness of instruction.

(3) To motivate pupils to improve in their work. Test results will encourage most pupils to put forth their best efforts. A word of caution is in order here. Every class has pupils of varying abilities. We do not expect all of them to run at the same pace when they are racing. In the same way it would be wrong to expect all pupils, the bright and slow ones, to achieve the same standards on a test. It would be wrong, therefore, to compare the mark of a pupil with lower ability with that of a pupil with higher ability. It is sound, however, to stimulate pupils to improve their own marks on successive tests rather than comparing them with the brightest pupils.

(4) To discover individual problems and weaknesses. The test results will identify pupils who have particular problems and the teacher then has the opportunity to provide individual help and instruction to such pupils.

(5) To provide a sound basis for keeping parents informed regarding pupils' progress. Parents are usually interested in knowing how their children are performing. If full records of test results are kept by the teacher, these form a good basis of communicating pupil progress to the parents.

(6) To locate or identify weak areas in the teaching-learning situation. If test results are analysed carefully "gaps" or weaknesses may be discovered and necessary steps taken to deal with them. Methods of doing this are mentioned later.

(7) To gain information for grouping pupils for instructional purposes. It has already been mentioned that pupils within a class vary greatly in their innate ability to learn. The slower pupils require simpler explanations and more instruction and drill. This frequently becomes boring to the brighter ones and causes them to lose interest. Test results would indicate which pupils might be grouped to provide the most suitable instruction.

(8)   <u>To gain knowledge about individual pupils for guidance</u>
<u>purposes</u>. If full records of test results are kept pupils'
strengths and weaknesses as well as their special interests
will be discovered. This information can be used in guiding
pupils to make proper choices when they go on to further
education or when choosing a vocation.

The above suggestions apply particularly to teacher-made classroom
tests. Standardized Tests that are prepared for a wider use, such as
throughout a school system or country, would serve other purposes as well.
Such tests should provide even more help and information to teachers,
parents, administrators, curriculum makers and educational planners and
policy makers. This additional information would help:

(1)   to evaluate courses or syllabuses for the purposes of
revision, etc.;

(2)   to compare different methods of instruction and assess teaching
methods;

(3)   to ascertain standards of classes within a school, a district,
a region or the country as a whole;

(4)   to assess the work of individual teachers;

(5)   to assess pupils and recognise the individual characteristics
of each pupil (we should know: (i) his difficulties and
weaknesses
(ii) his strength and present
knowledge);

(6)   to provide a means for pupil and teacher <u>review</u>, an integral
aspect of learning;

(7)   to provide information for educational planning and policy
making.

SOME OTHER GUIDELINES

It is worth emphasizing that testing should be done only when we
know <u>how</u> and <u>by whom</u> the results will be used. Of all the functions for
which tests may be used, the least valuable function educationally a test can
perform is when it is used <u>only</u> by administrators and <u>only</u> in passing,
failing, admitting or screening students. Yet examinations with this function
alone are still quite common. The most valuable function of tests is in
helping the pupil and teacher communicate, and derive benefits from the
learning process. In addition tests should be constructed with a clear
knowledge of all their <u>educational objectives</u> and should be critically
evaluated to see whether they are valid (measure what they are supposed to
measure) and are reliable (consistently measure the same thing in the same
way). The <u>interpretation</u> of course must be <u>logical</u>, attributing no more, or
less, to a particular result than it deserves.

It is worth noting also that if testing and evaluation are to become
an integral part of the educational system, teachers must know something
about the field. Generally today teachers learn how to demonstrate, explain
and put a point across. But little or nothing is given them on how to evaluate,

get feedback and measure what changes have taken place in the pupil. And yet to be truly effective a teacher must know (a) what the pupil has already; (b) what he has failed to learn and, if possible, why; (c) what the pupil is capable of learning.

There are, then, certain problems and pre-conditions to good testing that must be attacked simultaneously with any attempt to enlarge the role of evaluation in schools. These include:

(1) A FOUNDATION OF TEACHER TRAINING in the understanding, interpretation and use of tests. Programmes must be developed to improve this in the colleges and through in-service training. A basic record-keeping system (cumulative records) is needed together with teachers who can use it properly.

(2) DEFINED OBJECTIVES. Good evaluation programmes can help to show how far the school programme meets the objectives of education. This presupposes that there are measureable objectives set forth for general education as a whole and also detailed objectives for each subject. It is only against some objectives, however simple, that one can evaluate.

(3) OTHER GENERAL PROBLEMS TO TEST DEVELOPMENT IN A DEVELOPING COUNTRY.

   (a) Difficulty with control groups due to seemingly high turnover among pupils and teachers, very different levels of teacher training, lack of records of ages and other data etc.

   (b) Lack of pupil and teacher familiarity with the notion of (i) carefully timed tests (ii) objective tests.

   (c) Greater differences than in developed countries between urban and rural cultural factors.

   (d) Administration problems (developing the effective machinery necessary).

   (e) Language problems. Literature and other evidence suggests that any standardized test meant to measure "anything". In countries where English is not the first language it will, in fact, measure largely facility with the English language.

THE INFLUENCE OF TESTS

In a modern technological world there is bound to be a great concern with accurate measurement. Scientists can calculate an exact point and time for a moon landing 240,000 miles away. It is inevitable, then, that this desire to evaluate accurately should spill over into education. And, as is pointed out elsewhere in this paper, the two extremes (a) of attributing too much, and (b) too little, significance to the role and value of testing, both exist.

One author says "Measurement touches upon and influences every phase of education. Whether it is marking, promotion, guidance and

counselling, curriculum development, instruction or some other aspect of the work, measurement plays an important part."[2] Examinations and marks can be called the currency of education. By these marks, or value assigned, people are passed, granted certificates, promoted, given degrees and so on. We often judge a man's worth by his academic percentages!

There is general agreement then that testing can and does have a profound effect on the educational system of a country. The methods of teaching, the emphasis in the curriculum, the attitude of teachers and students, are all affected or sometimes dominated by the examinations. The types of things stressed in examinations largely determine what happens in the classroom. It matters little what teaching notes or syllabuses are prepared unless the examinations reflect the same spirit and aims. This is especially true where there are large scale and important external examinations.

EXAMINATIONS

Although examinations should measure what is being taught in the classroom, it is very easy for the situation to develop where we teach what is tested rather than test what is taught. Curriculum development and examinations cannot be separated and should be developed in close harmony at all points. Persons sitting on Curriculum Panels or Examination Panels should both be familiar with the general national aims and objectives of education as well as the specific spirit and aims of a given sullabus.

All this does not mean that examinations are the only determining factor in education nor is this a criticism of external examinations. The important thing is that these things should be in the right order and priority; tests should serve the educational goals and needs, not determine them.

The author recently sat on a committee the members of which were drawn from the Ghana Ministry of Education and the West African Examinations Council. A paper produced as a result of these meetings had in part this comment on examinations:

"In spite of inherent weaknesses external examinations are useful and necessary in many situations. In Ghana, for instance, some common measure is needed to provide objective norms and maintain a common standard owing to great disparities in:

    i) staff;

    ii) training facilities;

    iii) libraries and supply of textbooks, etc."

The point here, then, is not to weigh the advantages and disadvantages of external examinations not to debate how much external examinations can affect classroom practice, for such argument or debate has limited value.

---

(2) V.H. Holl, Introduction to Educational Measurement, Haughton Mifflin Co., Boston, 1965.

INTEGRATING EXAMINATIONS AND CURRICULUM

The important thing, then, is to recognise

(1)   that there is interaction between examinations and curriculum;

(2)   that examinations are not simply passive instruments of assessment but an integral and vital part of the educational process;

(3)   that both examinations and curriculum are important and powerful forces for change;

(4)   that both form part and parcel of the educational process; and

(5)   that both should be under constant review in terms of relevance to changing needs.

The central problem - which must be true of every educational system - is to find the most effective ways of ensuring that curriculum planning and examinations complement each other and work towards the same end. In other words, what should be done is to make sure that the examinations used (a) reflect the same goals (b) promote the same spirit, objectives, emphasis and priorities that the curriculum planners had in mind. Without the proper integration with curriculum, a tester starting from the same written syllabus could build several examinations, each one providing a different emphasis and different educational objectives and goals. It is for these reasons that the contacts between curriculum builders and the examiners must be continuous, and at all stages of development.

DEFINING OBJECTIVES

To establish contact between curriculum planners and examiners it is essential that the objectives and goals of education, both general and specific, should be clearly defined and clearly set out. Without clear direction as to the goals and objectives in education it can follow that there can be the situation where the main emphasis will be teaching what is tested rather than testing what is taught. Examinations can either lead or follow in education. When examinations become the key determiners of curriculum and education, it is usually by default, because the curriculum planners and syllabus writers have not been clear enough in their directions and objectives. Similarly if objectives are clearly defined but examiners are not properly informed about these and cannot translate them into the examination material the same unhappy situation may occur.

In order to meet today's needs, curriculum panels or testing panels must be conversant with:

(1)   modern testing ideas;

(2)   general aims and objectives of education for the country;

(3)   the desirable objectives, spirit, and emphasis for that particular syllabus or subject.

CONCLUSIONS

Curriculum and examinations are two sides of the same coin and it is only when they operate together that the goals and objectives of education can be adequately reached.

The pre-requisites, then, are:

(1)   that curriculum makers must build into the original curriculum evaluation and testing goals and objectives, and

(2)   see to it that they reach the examiners who actually make up the tests;

(3)   that the examiners (i.e. the examining body) must keep themselves fully informed at all stages of curriculum planning and defining objectives, and

(4)   become conscious of the spirit, aims and implications of the written syllabus;

(5)   that both curriculum planners and examiners, working as a team, appreciate where they are leading and heading from the earliest stages in terms of what will be measured and how.

PART ll

REVIEW AND EVALUATION:

As indicated in Part 1, the focus today in progressive education is on individualizing education, focussing on the individual. This means getting to know the pupil. Knowing the pupil in turn requires a number of practices including the necessity of measuring and evaluating each pupil in order to recognize individual differences. Hopefully the day of considering examinations as something separate and apart in education is over. The curriculum, the teaching and learning process and test and evaluation procedures should all be part and parcel of one complete and integrated process. Evaluation procedures turn education from a teacher-to-pupil 'monologue' into an effective 'dialogue' and communication. Continuous evaluation makes the process truly fruitful and meaningful.

Two main reasons have been given earlier for having examinations: the first is the need to provide evidence of an individual's competence to move from one social status to another; the second is to provide, as part of a continuous process of education, a method by which the teacher can assess what each student has and has not learned.

An analogy to educational assessment could be the study of cybernetics, a comparatively new science which is concerned with the behaviour of control systems in the physical and biological worlds. Perhaps the basic law of this science is that goal-seeking systems are error-actuated. What does this mean? Here is an example: a missile which "homes" on a target does not, in fact, go straight to it. Its course is constantly changing and its direction is modified in accordance with "feedback" information about its errors. The missile receives feedback which tells it how it is off target, and it then makes appropriate compensatory movements, though it can never be said to be exactly "on target". The same process can be recognised in

many human goal-seeking activities, and it seems directly applicable to the educational process. The teacher can apply this process consciously if he draws up a list of objectives which he hopes to achieve (he hopes to effect certain changes in his students as a result of the course he is teaching) and then by frequently evaluating his pupils' progress. He then has the feedback information which is necessary to reduce the errors inherent in progress towards any goal.

Most control systems have an optimum frequency for receiving feedback, they will swerve this way and that in their progress towards their target, but if they receive feedback too frequently, they may be unable to process the data at sufficient speed. (e.g. the feeling when one set of essays is due for collection before you have finished marking the previous set?) So the first question is: How frequently should one obtain such information? As we shall see later the answer is, as often as possible, in fact, continuously (but systematically).

It is a common occurrence to find the classroom teacher most surprised at the results of a testing programme. Assuming the test is a good and valid one, it shows how often teachers know little or nothing about the progress and capabilities of individuals, or indeed of even a whole class. Perhaps the most vital point in sound educational evaluation is the fact that to be effective evaluation must be both diverse and continuous. Part I of this paper, on the role of testing, emphasizes that tests are not precise instruments and that even the best of tests under the best conditions leaves large margins for error. For the sake of accuracy and reliability alone, evaluation must be continuous and not just a periodic event for selection or admission. However, as has been pointed out already, testing should be the other side of the educational coin, the means whereby communication and feedback to the educators is established. So, in order that evaluation procedures may provide accurate information to (1) pupil, (2) teacher, (3) parent and (4) administrator, and in order that evaluation may become an integral and useful part of the educational process, it must be continuous.

As a periodic event applied at certain times for screening and selecting pupils, testing has a limited educational function. And this is mainly an administrative function since it allows those in charge to pick people, pass and fail people, and assign a certificate or value to the person in question. It has limited educational value because it does not necessarily help the teacher to teach better, or the pupil to learn more effectively, which is, after all, the core of the educational process. Evaluation taken in the many forms discussed later, and used for diagnostic and remedial purposes on a day-to-day basis is the type of evaluation which really is a valuable part of the teaching-learning process. In this sense, written tests, oral questioning, quizzes, projects, reports, classwork, homework, etc., etc., are all considered as evaluative measures. Records are kept and results are analysed first in the classroom to help the teacher teach better and pupil to learn better, and secondly in the larger area of planning, curriculum, etc.

This is not really something dramatically new. It simply involves a consciousness on the part of teacher and other educators of the need for constant feedback and communication from the pupil. It shows a mature realization of the weaknesses and deficiencies of any one test or group of tests. It simply means a greater emphasis on gaining more and better information about the pupil and what is happening in the classroom and then using this in the next stages. It also involves a recognition of the importance of individual differences and the realization that to know pupils well and

accurately, we must assess (1) very frequently and (2) in as many ways as possible. It means that the focus is taken off the class as a single entity which must absorb a certain amount of material and be able to regurgitate this in an examination. Instead, education with evaluation as an integral part focusses on:

(1) developing the full potential of each individual to his or her capacity;

(2) seeing education as a dialogue between the teacher and the learner, where both, communicating effectively and constantly, 'grow' together;

(3) providing more accurate information as a result of continuous, diverse, and multi-faceted evaluation to all the people who need it - first of all the teacher and pupil, secondly parents, thirdly curriculum builders, planners and administrators.

But to be successful, a wide-scale use of testing and evaluation as a continuous element in the schools requires certain basics and pre-requisites. The main one is in the training of the teachers both in training institutions and through in-service work. Before testing can play the role described above, an understanding of some testing theory along with enough technical knowledge to understand, use and interpret tests and evaluation procedures must be basic to all education officers and teachers. They must be able to make effective use of objective as well as essay questions and to be able to use all the other evaluation techniques. In the field of objective tests they should have a practical classroom knowledge about the construction and interpretations of the various types of objective tests such as multiple choice, fill-in blanks, true-false, matching and so on. In addition a "guidance approach" to the child and test results and cumulative records should be a part of the teacher's equipment.

It should be emphasized as well that testing and evaluation does not mean only large scale sophisticated standardized tests. Widespread use of classroom testing combined with effective records and use of results, could provide much of the evaluation data now lacking. "Measurement devices and techniques prepared by the teacher are often the best and sometimes the only means of determining how well a class or individual pupils are progressing towards the objectives of instruction."[3]

But in such continuous classroom testing the teacher must develop a certain level of sophistication. For instance, the ability to analyse test results (do a simple item analysis) can add a great deal to the teacher's knowledge of the effectiveness of his teaching and the extent of the learning. It is also a powerful tool for test improvement. Item analysis also indicates which items are too easy or too difficult to discriminate between better and poorer examinees and it can be done simply and with little loss of time in the classroom. Then, too, teachers must understand validity in testing - that is, that the test measures what it is supposed to measure.

---

(3) Victor H. Holl, Introduction to Educational Measurement, Haughton Mifflin & Co. Boston, 1965.

The classroom teacher who has a sound knowledge of testing and who has available cumulative test records is in a far better position to understand the learning problems and difficulties of individual children. He is able to identify the most capable youngsters, who need enriched learning experiences, as well as the slow learners who may need special help and modified assignments. The slow learner who achieves less because he is slow mentally is a perfectly normal child; he should in no sense be considered a failure simply because he does not reach the average level of achievement of children of his own age or grade. On the other hand, the child with high ability who does mediocre work, is, in a truer sense of the word, a school learning problem. The concept of failure in school is one with which we could very easily dispense since it is never possible to determine with certainty who is failing - it may be the school quite as much as the child.

It should be stated emphatically that standardized testing here is no complete substitute for an effective evaluation programme on the part of the classroom teacher. Such an evaluation programme includes the teacher's own locally constructed tests as well as ratings on specially assigned projects and daily classroom recitations. Nevertheless, the professional technicians who develop standardized tests can offer the classroom teacher many suggestions for evaluating the results of classroom instruction. Indeed, much in-service training is needed in this important area.

At several points it has been mentioned that evaluation must (1) be continuous (2) employ a wide variety of evaluative procedures. What are some of these ways of evaluating? Here is a list of some of the many common-sense methods for diverse and multi-faceted evaluation:

(1)   Tests - there are many types including:

    (a) Achievement:

        (i) informal teacher-made

        (ii) standardized

    (b) Mental ability

    (c) Personality

    (d) Aptitude

    (e) Interest;

(2)   Rating scales;

(3)   Checklists, surveys, inventories and questionnaires;

(4)   Observation;

(5)   Records and reports:

    (a) cumulative folders,

    (b) anecdotal reports,

    (c) diaries and logs;

(6)　Interview;

(7)　Sociometry;

(8)　Role-playing:

    (a)　sociodrama,

    (b)　psychodrama;

(9)　Situational or performance tests;

(10)　Student papers and projects:

    (a)　papers,

    (b)　notebooks,

    (c)　reports,

    (d)　autobiographies,

    (e)　personal data sheets;

(11)　Case studies;

(12)　Case conferences.

Tests must measure all the important outcomes of instruction such as course objectives, factual knowledge, understanding of human nature, the proper weight for each topic and so on. Benjamin Bloom[4] in his taxonomy lists six main objectives or outcomes of learning in the cognitive domain that should be measured. These include:

    (a)　Knowledge of specifics, ways and means of dealing with specific universals, abstraction from specifics;

    (b)　Comprehension, involving abstraction, interpretation, extrapolation of communication;

    (c)　Application of knowledge;

    (d)　Analysis of elements, relative principles;

    (e)　Synthesis;

    (f)　Evaluation.

Before continuous evaluation becomes a full and integral element in education there must be clearly defined educational objectives. Good evaluation shows how far school progress meets the objectives set out. These must be clear both to the policy makers, examination bodies and teachers. Unless these are clarified, testing cannot play its proper role.

---

(4) Bloom, B.S. (ed) - Taxonomy of Educational Objectives, Handbook I: Cognitive Domain, David McKay & Co., Inc., New York, 1956.

Then too in newly developing countries there are other problems which often require close attention before testing can be effective; these include:

(a) Keeping records. Records of pupils' ages are basic to many kinds of testing. Continuous evaluation is only useful when a cumulative record of each individual is kept and used.

(b) Both teachers and pupils must become familiar and at ease with such things as objective tests, the concept of "time tests" where every minute counts, and so on.

(c) Often in such countries there is a less settled population of teachers and pupils, both of whom move about and leave schools frequently. This makes for difficulties in establishing norms, control groups and experimental groups.

In newly developing countries differences between urban and rural groups tend to be bigger. Urban groups quickly become sophisticated in a variety of aspects and ideas while rural pupils remain almost totally unaffected. Language problems, administrative problems and others must all be tackled in a special way for rural areas.

It was mentioned earlier that in addition to evaluation being continuous it must include a wide variety of techniques and methods. In testing, we are measuring people and their responses and knowledge and not bricks or bridges. Physical things can be measured accurately and completely with a ruler or scale. Because of the complexity of man and the complexity of the facets we wish to measure, there is more chance of accurate assessment if a variety of techniques are used and used often. And by accumulating and combining results, we are more likely to measure accurately the many processes and facets that we see as the goals of education.

In summary, then, worthwhile testing should meet a number of criteria and requirements. An attempt is made here to include the key elements in a graphic form:
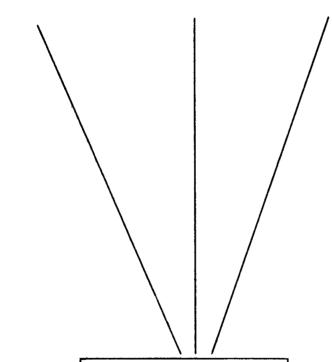
GOOD WORTHWHILE TESTING

Proper Administration

**Timing, testing condi-
tions, uniformity,etc.,
are all properly done.**

Good Sampling, Clear Objectives

Questions measure different
aspects and levels (Bloom) - valid
and reliable tests - the tests
measure all the objectives of edu-
cation and measure it always in
the same way.

Continuous Evaluation

A variety of types of ways of
testing are used continually as
an integral part of the teaching
and learning process. Testing
must be continuous:
(1) to provide communication
and dialogue between teacher
and pupil;
(2) to make evaluation an
effective teaching and learning
tool by providing pupils and
educators with feedback infor-
mation on what is happening;
(3) to allow pupil and teacher
to assess progress, assess
their work and make adjust-
ments;
(4) to provide up-to-date infor-
mation to teachers, pupils,
parents, administrators,
curriculum builders and policy
makers;
(5) to compensate for the inherent
weaknesses in the results of any
one test. The average results of
a great many (continuous) assign-
ments are much more valid,
reliable, accurate and meaningful
than the results of any one or two
major tests however carefully
constructed.

PROPER USE OF TESTS

Interpretation

**Test results are
interpreted for what
they are. No more or
less evidence** or value
is attached to them
than they deserve.
Test results are not
regarded as "god-like"
nor as "useless".
They are given their
proper due and right and
used as one piece of
information along with
all other information
available.

Well used information

A truly useful test provides
information to all of:
1) pupils
2) teachers
3) administrators and
   inspectors
4) parents
5) school curriculum
   builders and policy
   makers.

Variety of Technique

Evaluation must take a variety of
forms and include as many
different kinds of evaluative
techniques as possible. These
should range from carefully con-
structed standardized terms to
common-sense evaluative
observations and ratings of
teachers. This is because
1) we are assessing complex
human beings;
2) the qualities being measured
are tangible, abstract and
difficult to measure;
3) more kinds of measurements
taken more often ensure more
accuracy, reliability and
validity.

Part I of this paper presented some points to stimulate discussion on the topic of the nature, place and influence of tests and measurement procedures including examinations; Part II has focussed on continuous review and evaluation. As was pointed out in the opening statements of Part I, the author feels that the two topics are indeed only one. Evaluation is a continuous and integral part of education. The frequent repetition in Part II of points from Part I are meant to emphasize this point.

However, for organizational purposes, the paper was divided into two parts. It is hoped, however, that the overlapping and repetitions of similar points in the two parts may lead to a line of thinking that combines the two ideas.

This marks the end then of a few brief ideas in the field of measurement. Not all points are covered, nor are those that are mentioned covered adequately or completely. However, hopefully, enough has been said to provide the raw material so that discussion can whittle away the rough edges and produce a refined and finished product on this important and contraversial subject. Perhaps because of its deficiencies, this paper will serve the better to stimulate discussion, which is, after all, its purpose.