

**THE ECONOMIC AND SOCIAL
RESEARCH INSTITUTE**

Memorandum Series No. 84.

NOTES ON DEVELOPMENTS IN THE TECHNOLOGY
OF EDUCATIONAL EVALUATION

(For Department of Education Inter Cert Committee, 1971)

John Raven

November 1970

**Confidential: Not to be quoted
until the permission of the Author
and the Institute is obtained.**

These notes are concerned with developments in the technology of educational evaluation. They are, however, written in the context of some other considerations which, if not accepted, mean that much of what is said is irrelevant.

These considerations are: first, that the main objective in educational evaluation is to provide feedback to teachers and pupils concerning how well they are doing and how they should change their behaviour in order to do better in the future. Educational evaluation is not mainly concerned with providing employers or others with information. The latter goal often conflicts with that of providing feedback to schools and pupils; schools and pupils need specific and detailed information if they are to use it as basic data on which to base future activities; potential employers probably need fairly global assessments supplemented by information on specific aptitudes which are relevant to their own particular needs in their own particular field.

Detailed information is needed by schools to guide policy not only in the areas of cognitive development, but also in areas of non-cognitive functioning: how is the pupil's personality developing? And what guidance does he need? How do pupils perceive their future? In what ways are their self-images developing and changing? How do the pupils perceive their school and society? How do they think the latter works and what do they see as their place in it?

Educationalists, teachers, and pupils require two different types of feedback: one is a broad picture of the human resources located in the youth of the nation. The objective of this assessment is to monitor, over a very wide range of attitudes, values, knowledge, and social and intellectual skills, what is happening over time and with improved, or at least changed, methods of education, and to assess the quality of the current output from our educational institutions. This monitoring would enable one to detect whether there was a drift in standards in any one area of educational endeavour, and whether some areas were being neglected or over-emphasised in education as a whole. The feed-back would necessarily be concerned with all pupils and providing an objective picture of what their knowledge, skills and attitudes are. It would be possible to see from this whether the percentage of pupils reaching any particular standard, in any area of social, intellectual, or attitudinal competency, was increasing or decreasing. Although the assessments would have to relate to all

pupils this does not mean that the assessments could not be carried out on random samples of pupils rather than the total population. In fact fairly reliable estimates of what was happening in the total population could be obtained from well constructed samples.

The other type of feedback required is that aimed at providing detailed information to individual teachers and pupils concerning their attainment of the objectives they set themselves. It would involve providing detailed feedback over a wide range of educational outcomes so that corrective action could be taken in relation to the needs of individual pupils or methods of instruction. The feedback would have to include feedback concerning the effect that the educational inputs currently being used was having on the attitudes of the pupils, their feelings of competence, their motivation to apply themselves, etc. It would have to include feedback of this broad type since it is possible to envisage instructional techniques which facilitate the attainment of one goal (e.g., imparting knowledge of routine procedures) having negative effects in other areas (e.g., encouragement of initiative). And the feedback from these detailed assessments would have to include feedback concerning the effects that the present educational inputs were having on detailed competencies within subject matter areas, these would include: understanding of concepts, ability to perform manipulative operations, ability to apply knowledge to new problems, etc.

The second consideration which has guided the selection of material presented here and the way in which it is viewed is the belief that one of the main problems in education is to develop systems of instruction and evaluation which enable schools to run a wide range of courses suited to the needs of pupils who come from different backgrounds, have different values, have different interests, have mastered different bodies of knowledge and different skills, have acquired different attitudes, have different aptitudes and abilities, and will enter different walks of life in which they will require different knowledge, skills and attitudes and from which they will make different types of contribution to society.

If this is true - that is, if it is true that the specific objectives of education are different for different pupils, then it is clearly inappropriate for them all to be assessed on a common examination syllabus and certificate, whatever its form. Which is not to say that one does not need some general educational mapping device in order to monitor what the general state of the human resources located in youth is and how it is developing and changing.

The third consideration which has provided guidelines and a context for the selection of the material that is to follow is that many teachers feel that many educational objectives are not pursued sufficiently at the present time owing to the fact that there is so much pressure for adequate examination performance in the areas which are evaluated. The view of these teachers is that education has become lop-sided and lost sight of many of its main objectives. One way of remedying this situation is to evaluate the attainment of some of the other objectives. Hence a considerable amount of attention is paid to this problem in these notes although it has not traditionally figured as a major objective of examinations.

Finally, in order to cover developments in the technology of evaluations adequately, it is necessary to include here a discussion of some developments which would normally be considered developments in teaching techniques, (rather than evaluation techniques). This is because these techniques either contain within themselves developments in evaluative techniques, or have different objectives (which therefore require different evaluative techniques) or permit of, and encourage, wider evaluation than that normally attempted in external school examinations.

Programmed Instructions

Objectives:

- To provide detailed feedback to the pupil concerning his progress, so that mistakes can be rectified immediately.
- To ensure mastery of that which has gone before so that time is not wasted trying to teach a pupil something new when he has not mastered essential previous work and is, as a result, completely at a loss to know what is going on.
- To ensure adequate motivation. To encourage the pupil to believe that he could master new material without reference to a teacher.
- To free the teacher from nearly all routine instruction and correction, and hence allow him to concentrate on more important aspects of education.
- To allow pupils to know that they can quickly master the necessary knowledge, skills and attitudes to retrain for a new occupation, move into a new field, or sample the contents of some field of special interest, both now and later in life.
- To allow different pupils to master different fields of knowledge, rather than have them all master the same one.
- To change the teacher's role from that of disciplinarian and goad to activity to that of helpmate to the pupil in performing a task which he has himself chosen and into which he is thus more likely to put effort.

Queries:

- . What range of programmes are available?
- . What are their costs?
- . How are they working in other countries?
- . Can programmes be developed to teach and assess general understanding and ability to apply the facts and techniques that have been learned to new problems?
- . To what extent can skills such as willingness to recognise that one has a problem and willingness to seek out relevant information be programmed?
- . To what extent can attitude development be programmed?
- . To what extent can such skills as ability to trace themes to previous authors be programmed?

Project WorkObjectives:

- . To motivate pupils to wish to learn, and hence develop a positive attitude toward education and re-education.
- . To develop skills of working on one's own, learning to find one's own references, learning to use libraries and reference sources.
- . To encourage pupils to make contact with the world outside school.
- . To change the teacher's role from that of disciplinarian, instructor, and goad to activity, to that of helpmate and counsellor.
- . To develop skills of working with others.
- . To develop skills of communication and expression.
- . To encourage pupils to develop self-motivation.
- . To enable a wider range of attitudes and abilities to be assessed.
- . To avoid having all pupils measure themselves against the same standards.
- . To output students having a much wider range of knowledge, skills, attitudes, and interests.
- . To help pupils develop more appropriate images of many roles in society.
- . To help pupils develop more appropriate self-images, and, particularly vocational self-images.
- . To avoid pupils being given a sense of failure by being unable to do that which others do.
- . To encourage pupils to recognise problems, and realise that they are capable of doing something to further their understanding of these problems.
- . To encourage pupils to be able to analyse material, develop an argument well, and present material well.

- . To encourage pupils to break down subject boundaries and relate what they are learning to their everyday lives.
- . To familiarize pupils with sources of information.

Queries:

- . Costs? (including strain on teachers, school buildings, etc.).
- . Availability of assembled materials.
- . Range of skills that can be assessed?
- . Problems?
- . How effective in relation to each of the objectives listed above?
- . How reliable and valid are the school based externally moderated assessments which are necessarily involved if assessments of any useful sort are to be passed on to employers and others?
- . Problems of evaluation?

Objective Tests

Objectives:

- . To reduce unreliability of marking.
- . To permit better sampling of content, than is possible with essay-type examinations, that is to make sure that the pupil has adequate knowledge of all aspects of the course.
- . To provide detailed feedback to teachers and pupils by giving a number of scores for different areas of content within subject areas so that they know what their strong and weak points are, and can improve on what they are doing.
- . To assess a wide range of intellectual and non-intellectual characteristics such as social skills, attitudes and values, attitudes to school, attitudes to education, attitudes toward society, development of vocational and personal self-images.
- . To reduce the marking load on teachers through having the materials machine scored.
- . To allow national norms to be constructed within which schools can reasonably make their own assessments.

Queries:

- . Costs?
- . Range of characteristics that can be assessed?
- . How does it work elsewhere?
- . Effects of too narrow a scope of assessments?
- . Can it be developed to assess, e.g., ability to develop an argument: if not what are the effects on these characteristics?

- . What special skills are needed to generate good objective test items?
How costly is it to train sufficient people in such skills? Thereafter what is the cost of the system?
- . What additional administrative and technical machinery is required for such a system?
- . What savings could be achieved compared with the costs of the present system?

Essay-Type Questions

Objectives:

- . To encourage pupils to express themselves and be able to develop an argument.
- . To encourage pupils to link material together.
- . To encourage pupils to select from their knowledge that which is most relevant to bring to bear in answering any particular question.
- . To assess a pupil's best performance, rather than his average over all possible areas of content connected with the course. (This, of course, implies that pupils should be given a wide range of topics from which to choose.)

Disadvantages:

- . Different pupils perceive different questions in the same set topic, that is, they feel that different types of answer and content is required; very often if it occurred to them that a set of facts and ideas that they do know was relevant, and should be produced in answer to the question they would be able to provide it.
- . It often confines pupils and students to the very child-like activity of having to recite what they know in order to show that they know it, rather than use it. And if they do use their knowledge then it must be soul-destroying to have the product disappear for ever into the limbo of the Department of Education.

Queries:

- . How can assessment of essay-type work be made both more reliable and valid?
- . Can its objectives usefully be clarified and a system developed to avoid the disadvantages whilst retaining the advantages?

Self-Image and "Attitude" Measurement

Objectives:

- . To enable the schools to monitor the effect that they are having on the development of their pupils in areas quite outside those with which external examinations have traditionally been concerned, that is to

monitor the effect that they are having on the pupils' self-images, that is on the areas in which the pupils regard themselves as competent and incompetent, the areas they feel they know about and do not know about, their feelings of ability to master new things etc., to monitor the effect that they are having on the pupils' views of school and society, the way that society works and their possible roles within it. To monitor the development of pupils' abilities to relate effectively to others, their willingness to take initiative, their willingness to listen to others and take what they say into account, their willingness to respect values other than their own and tolerate other ways of life etc., to monitor educational experiments etc.

- . To enable schools to provide a much more effective system of guidance and placement for their pupils so that pupils receive an education that is suited to the background from which they come and their interests, values, abilities and aptitudes.

Queries:

- . To what extent can these characteristics really be assessed?
- . What is the cost?
- . What further development work needs to be carried out on these measuring instruments?

Abolition of External Examinations and Substitution of School Based Examining

Objectives:

- . To allow different schools and teachers to pursue different policies.
- . To allow a wider range of educational objectives to be included in the assessments. This can be done not only by introducing attitude measurement methods, as discussed above, but also by training teachers to make use of their own observations - to be sensitive to behavioural indicators previously neglected and to structure special situations in which additional information can be collected. This, of course, involves training teachers much more in the technology of assessment of people.
- . To encourage the belief among both pupils and teachers that evaluation is a friend and guide rather than a judgement on work done.
- . To avoid examination 'fever' and allow continuous assessment.
- . To avoid forcing all pupils to plod through the same system regardless of the differences in their background, abilities, values, interests, and the positions they will enter in society.
- . To encourage more teachers, through participation in drawing up syllabuses and assessing pupils in different types of schools to become involved in curriculum development.
- . To save money on the costs of administering a traditional external examination.

Queries:

- . What does the traditional external system cost?
- . Costs of a school-based system?
- . Problems?
- . Efficacy?
- . Problems of moderation?
- . Are teachers subjected to pressures from some parents to upgrade their children's grades?
- . Do more teachers in fact become involved in re-thinking the goals of education?

General Query

One of the main queries one is left with after glancing over these notes is "How can these various techniques best be combined?". It is quite clear that many of them do not conflict. Indeed they perform complementary functions: for example, project work combined with programmed learning of necessary material would seem to be a highly desirable way of ensuring that pupils master relevant knowledge, skills and attitudes. Simultaneously, the overall effect that schools were having on basic social and intellectual skills, and on images of self and society, could be assessed by means of objective tests, and attitude and image measurement. Furthermore appropriate guidance could be given to individual pupils on the basis of some of this material.

However, it should also be noted that much of the feedback to schools does not require, as an essential ingredient, that all pupils in the schools be tested on all measures; one can obtain general information on the effect that one is having from studies of samples of pupils. As a result the evaluation programme could involve assessing different sub-samples of pupils on different measures so that all important areas would be studied but no pupil would have to complete all the assessments. By no means all assessments would involve paper and pencil tests.

THE QUERY

Supposing that more detailed investigation revealed that steps to utilize some of these procedures was desirable, even if, at present, too costly to implement in toto, how could one set about producing such a major change in education and evaluation?

Several steps could be taken:

1. Before doing anything else one could put in hand the necessary desk-research to answer, from material available in other countries, the questions that have been raised and to ascertain procedures, costs, feasibility and advantages and disadvantages.
2. One could broaden and generalize the assessments made in the Intermediate examination in order to provide the detailed information needed by the schools: change its format to approximate more closely to that of objective tests and include attitude measurement. Diffuse the knowledge that it is a mapping and research device intended to find out what the present situation is (that is how pupils are reacting to the educational system, how they perceive society, what social and intellectual knowledge, beliefs, attitudes, and skills or competencies they possess) and what trends are discernable with changing exposure to different types of education, rather than a device intended to pass judgement on individual pupils. Emphasize that it is designed to give information to schools rather than pupils, and that the assessments are designed to give an overall but detailed picture of the current state of the human resources located in the nation's youth. Educate the public to recognize how much more valuable this overall picture of the human resources located in youth would be to educationalists and policy makers compared with the current rather black and white picture painted in terms of whether or not pupils reach particular standards in education. Emphasize that the new system would enable one to ensure that the present standards of attainment were maintained or increased, but that it would also enable attention to be paid to the standards prevailing in many other areas of concern to educationalists. Emphasize that it would enable pupils to develop more realistic vocational self-images - to recognize, perhaps, that they could not attain the highest academic grades, but simultaneously to recognize, in detail, what their strengths are and pay attention to developing them. Link the new evaluation system to other testing and guidance programmes run by the Department of Education. In due course, as the number and adequacy of the elective objective tests available from the Department increases, change the base of the national monitoring system to samples instead of the total

population (sampling of both pupils and school subjects being implied) and extend the population studied to other age groups. Encourage everyone to recognize that the system in operation at any one time is in the process of developing and changing, and that they have a part to play in continuously re-shaping it to meet their needs better.

3. One could develop more project work, programmed instruction, and teaching for mastery in schools, and set up a system to provide more extensive (i.e., including more attitudinal etc., data) and more detailed feedback to pupils based on programmed learning and internally assessed externally moderated project assessments.
4. One could move toward offering a much wider range of short (e.g., half year) courses in schools so that pupils can sample a much wider range of subjects in order to clarify their interests and abilities (and, perhaps, so that pupils and ex-pupils who have followed blind alleys can re-trace their steps and embark on a new course of career). Incidentally the way such systems of education work elsewhere (e.g., in Australia) should be thoroughly studied before their introduction here.
5. One could develop within schools a more extensive programme of individual guidance and placement for pupils, a programme that would involve considering much more carefully what is best for the personal development of each individual pupil. In due course this would involve recruitment of specialists and using specialist techniques, but in the meantime there is no reason why a move should not be made towards doing more effectively and explicitly that which many schools already try to do. In the long run this work would link in with the elective objective tests available from the Department and thus form the complement to the assessment system envisaged under 2. This elective system of papers would be designed to provide detailed information to individual teachers and pupils concerning their attainment of the objectives they set themselves.