B. Levrat

COMPARATIVE STUDY ON CRITERIA AND PROCEDURES FOR THE EVALUATION OF EDUCATIONAL SOFTWARE

A STUDY BY THE ICEM
INFORMATION TECHNOLOGY SUB-COMMITTEE
UNDER UNESCO CONTRACT
NO.106.534.6

Revised Edition September 1988

EDITED BY
RICHARD N. TUCKER

NEDERLANDS INSTITUUT VOOR
AUDIO-VISUELE MEDIA

FOREWORD

No account of what happens in the educational world in a number of differing countries can hope to give a completely objective view. This study is concerned with the criteria and procedures for the evaluation of educational software in a number of member countries of the International Council for Educational Media (ICEM). Whilst criteria may be set on paper and procedures drawn up by governments, the evaluation itself is a dynamic process involving a series of complex subjective judgements on the part of the evaluator. As will be seen from this report, whilst we attempt to define the evaluation of the product we have really to consider the evaluation of the process- the software in use.

On behalf of the ICEM Sub-Committee on Information Technology I would like to thank those correspondents who took such pains in the writing of their reports and the submission of sample forms. The report which follows is largely made up of the words of these correspondents. It was my intention in blending their words together to give a constant and consistant view of the way in which each country mirrored the questions. The resulting synthesis must therefore be speculative in trying to present a generalised view. If by taking any of the correspondents words out of their original context I have distorted their picture of what is happening in their own country then they have my unreserved apology. If the views that I eventually draw from their reports are not acceptable they should be attributed to me and not to the correspondents.

Richard N Tucker
Den Haag
January 1988.

CONTENTS.

Comparative Study on Criteria and Procedures for the Evaluation of Educational Software

Terms of Reference

Under the UNESCO program on enhancing the quality of educational software ICEM, the International Council for Educational Media, was asked to conduct a study based on reports from a selected number of the ICEM member countries. This selection was to represent different styles of educational organisation, from the centralised to the decentralised, and the different organisational methods for the production of educational software from governmental to commercial. The countries were also to be selected to represent different levels of integration of the computer into education. On this last point it was reasonable to select countries with a relatively high integration in order to find enough evidence of the evaluation of software

A summary of the national reports was then to be produced drawing out the main similarities and differences and commenting, in the light of existing reports, on software evaluation in general. In the conclusion special attention was to be given to recognizing those evaluation criteria, based on market factors, cost factors and social objectives which prevail in industrialised countries and also to pointing out evaluation criteria based on learning environments which are broadly applicable within the international community. The report was to end with a set of recommendations endorsed by the Information Technology Sub-Committee of ICEM, aiming at establishing the quality of software.

Basis of survey

The sub-committee considered the contract in great detail. It was clear that there were already many different systems for the evaluation of software. It was also clear that not all evaluators were evaluating the same things for the same reasons. The context of the evaluation was considered to be crucial to an understanding of the process.

It was therefore considered necessary to define the sorts of software to be studied and the types of evaluation which might be applied. Given that some countries produce the bulk of their software centrally and therefore build evaluation into the development process, whilst others have little or no institutional production and therefore place the emphasis on theevaluation and selection of commercial products, it was necessary to produce a structure of types of evaluation before proceeding to a questionnaire.

Types of program

a) Professional Programs. These are programs produced initially for the professional or business community and which may later be used for education, e.g. wordprocessors, databases, spreadsheets, graphics and CAD/CAM etc. The aim here was not to evaluate them against their own objective for this will have been proved in the open market but whether they meet the expectations of their educational users.

b) Open or Applications Programs. Designed for education these have no specific inbuilt content which makes them goal-specific. They are sometimes called "content-free" programs, although some are in fact built around a specific content which can be added to or altered by the user. Some are simplified versions of the commercial professional programs. The adaptation to the educational market often rests on the provision of different documentation and some in-built examples.

c) Didactic Programs. These are programs with specific content and usually with stated learning objectives or goals. Here can be found the bulk of the "educational" software including; drill-and-practice, tutorial, simulation etc.

Evaluation

Evaluation may take place during at least three phases. Although there can never be fixed boundaries between the categories these may be regarded as;

a) Development

The **Producers** of programs must have criteria on which they base the decision to produce a program, technical and methodological criteria for the development, and user criteria against which they test the product.

b)Selection for Translation and Conversion

The needs of the **Converters** are seen mainly in the smaller countries which, because of market sizes, are more concerned with the purchase and conversion of existing software than developing their own.

c) Selection by Authorities and Schools

the eventual **Users** of programs must have criteria by which they determine which programs are bought to meet their educational needs. This is perhaps the most widely known form of software evaluation, particularly because of the publicity which it has received from the USA and Canada.

The questionnaire, which was designed as the survey instrument, attempted to provide a matrix within which the commissioned authors would write the national studies. It was recognised that not all the elements in the matrix would be applicable in every country.

Software / Evaluator	Professional	Applications	Didactic
Developers			
Converters			
Users			

With this as a basis authors in several countries were approached. A questionnaire was devised, not with the intention that authors should slavishly follow the questions but in order to make it possible to construct a summary report.

Reports were submitted from the following;

Canada	Hans Kratz, Alberta Education
Japan	Haruo Nishinosono, Kyoto University of Education
Hungary	Ferenc Genzwein, OOK
Italy	Marcello Giacomantonio, TECOM
Netherland	Pieter Burghard; NIAM
Scotland	Alistiar Thompson, SCET
England	Richard Fothergil, CET
U.S.A.	Dr. Robin Taylor, University of Maryland

The Questions

CONTEXT

Please give a brief and clear description of the education system which provides the context for the rest of this study.

How centralised or decentralised is it? Who controls the curriculum?

Who makes the decisions about what is provided for the schools or bought by the schools? Whether it be the authorities or the school, how do these decisions relate to national, regional or local policies? How and by whom is the policy for computer hardware and software developed and implemented?

SOFTWARE DEVELOPMENT

Is the development wholly commercial, completely under Government control, through a central institution, through (a) project(s) or local education initiatives?

How does this software development relate to distribution and sales, and how is this done? From where do schools get their software?

Who decides what themes or subjects will be necessary for the schools/ or good for the market, and on the basis of what criteria? What systems are used to develop software in each of these instances? Where do the ideas originate? How much are the schools or practicing teachers involved in the development process?

What sort of field testing is there and are there stated criteria which are used during this development phase (by the developers themselves - not by the researchers)?

PRODUCTION

Are there criteria or standards by which production costs are estimated for items of software? For your country what is the number of copies which have to be produced to recover the investment? Where software is distributed, rather than sold, on what basis is this done - per school, per n pupils or per microcomputer?

On what basis are programs protected or not protected?

What computer languages are used now? What computer languages are envisaged for the near future?

CONVERSION

If software is bought in from other countries and converted for the local language and machines, how is this done and by whom?
Are there criteria by which this is done and if so how were they established? Are other, non-educational, criteria evident? Is there any evidence that these criteria are resulting in the best programs for local/national needs?

SELECTION

Who selects the software for the schools? Is this done nationally , is there national advice? Is it done by the local education authority, however that might be organised? Or is the selection left to the schools themselves or the individual teacher?

In each case, where applicable, are there predetermined criteria by which the software is evaluated? If so, is there a single instument which has to cover all the types of software or are there different sets of criteria/ variable criteria within the one list? what are the selection criteria for the different types of software?

USE

What criteria are used to evaluate the programs in the classroom? How were these established and by whom? Is there any measure of how much use a program gets in terms of time used, the number of users, and the frequency of use by the individual user?

The Report

Introduction

The distinctions drawn in the preceding questions have to be kept in mind in the following consideration of the reports from the eight contributers. In a country such as the United States of America evaluation of software calls to mind the reports from institutions such as the Educational Products Information Exchange (EPIE) and the many state or district organisations which have produced formulae for teachers to complete when using software, the summaries of which are regularly published. This results more often than not in a sort of descriptive cataloguing. Here it is clear that the emphasis is on evaluation for selection - a sort of educational"seal of good quality" which can be placed on a particular product. Naturally the program developers will have done their own technical evaluation and field testing, but the emphasis on the quality judgement of educational product arises from a market where there is an apparent abundance of good software and teachers without the time to look at everything feel the need for guidance.

As will be seen, the emphasis in other countries, especially those with a more centralised production of educational materials, may be placed on the evaluation of the product and its use during the *development* phases. Others may make the post production evaluation of the *use* of the program the more important aspect rather than an evaluation of how the *program* itself matches up to a set of criteria.

Therefore throughout this report, whilst the word **evaluation** will be used , the reader is asked to take some care to note the specific application. The writer will take equal care to try to make the differences clear. To start this process of clarification let us begin with a definition of evaluation. In the Oxford English Dictionary we find;

> Evaluate v. [f. Fr. *évalu-er* (see next *sb*) -ate^3] *trans.*
> a.*Math.* To work out the value of (a quanitative expressiom);
> to find a numerical expression for (any quantitative fact or
> relation). b. *gen.* To 'reckon up', ascertain the amount of;
> to express in terms of something already known.
>
> Evaluation. [a. Fr. *évaluation,* f. *évaluer,* f. *é* = *es*-(:-l. = *ex*)
> out + *value* Value]
> 1. The action of appraising or valuing (goods etc.): a calculation or
> statement of value
> 2. The action of evaluating or determining the value of (a mathematical
> expression, a physical quantity, etc), or of estimating the force of
> (probabilities, evidence etc).

The UNESCO "Glossary of educational technology terms expresses it thus:

> Evaluation: the process of delineating, obtaining and providing useful
> information for judging decision alternatives

In the Encyclopaedia of Educational Media Communications and Technology, where the relevant article is mainly concerned with curriculum evaluation we can read;

> Curriculum evaluation is the collection and provision of evidence,
> on the basis of which decisions can be taken about the feasibility,
> effectiveness and educational value of curriculua (Cooper 1976)

> Evaluation is the means whereby we systematically collect and analyse information about the results of a student's encounter with a learning experience (Rowntree, 1974)

> Evaluation consists of the collection and use of information concerning changes in pupil behaviour to make decisions about an educational programme (Wiley 1970)

> Evaluation is a tool in educational practice for ascertaining whether alternative procedures are equally effective or not in achieving a set of educational ends (Bloom *et al.*, 1971)

In dictionaries devoted to educational matters we find the following

> Evaluation is often used interchangeably with assessment.[ ] This is because there is a considerable overlap in their meanings. Both involve measurements designed to describe the amount of certain attributes. Both involve procedures for obtaining these measurements which can involve *tests* as well as less objective instruments such as rating *scales*. There is a tendency, however, for evaluation to be used in a more general way, involving a wider range of measures with a greater acceptance of subjective judgements. There is also a tendency foe evaluation to be used more when the subject of the evaluation is not a person (or group of persons) but the success of a course of teaching or the method of teaching. [...]
>
> (Hills P.J. *A Dictionary of Education 1982)*

> *Evaluation* (a) a *cognitive* process which involves making judgements about the value of ideas, works, solutions, methods, materials, etc for some specific purpose; the highest level (level 6) of Bloom's *cognitive domain;* [....]
>
> (Percival F., Ellington H, *A Handbook of Educational Technology* 1984)

A useful distinction to draw inevaluation studies is between *formative* evaluation and *summative* evaluation, and is concerned with their different purposes and scope.

The purpose of *formative* or on-going evaluation is to intervene at an early stage in the development of a software package or other learning material, so that suitable amendments and revisions to the package can be made before it reaches a finished state. Clearly formative evaluation is addressed to a local audience - the producers of the package, and its aim is to provide information to guide their decisions on any necessary or desirable changes to the package while it is still in a fluid state. It might well be carried out by a member of the package production team or by a critical friend, and it would typically consider individual software packages rather than a whole series of packages which might constitute a major curriculum innovation.

The purpose of *summative* evaluation is to give an objective, external critique to a finished software package and to make judgments as to its worth and possible usefulness to others. Summative evaluation is addressed to a wider audience - the learners; the development team; and academic and administrative staff of other institutions who might be interested in its use. It is more likely to be carried out by some independent external agent than by a member of the package development team; typically it would deal with a series of software packages, and attempt to evaluate a whole piece of curriculum innovation or development.

The UNESCO glossary further defines these concepts as follows:

> **formative evaluation:** 1. evaluation of instructional programmes whilst they are still in some stage of development; 2. evaluation which is intended to provide data for instruction product revision; 3. evaluation that occurs within and during the entire process of product, course, or programme design and production. Also referred to as "on-going evaluation".
> **summative evaluation:** 1. evaluation intended to provide data for product validation 2. evaluation oriented to consumer-administrator-teacher criteria and standards

From these definitions it can be seen that whilst the dictionary (at least in English) makes reference only to quantity and not to quality, the later definitions used in and developed from practice, are concerned with ideas such as effectiveness - but mainly with regard to the educational process itself or the changes which occur in the learner. It is only in these later definitions arising out of educational technology that one begins to see the identification of the subjective judgement of the learning materials themselves. What appears to have happened in the recent past is the shift from a complex, and to some extent subjective, process to one which tries hard to be objectively scientific. One has the feeling that much of this has arisen because what is now under consideration is thought to be capable of detailed evaluation because it is made up of lines of computer language. To go too far along this line is to ignore the fact that educational software, like all the other educational media, only has measureable value when it is being used .

It is clear from this study, and from the evidence of the last five to ten years, that when people talk about the evaluation of educational software they are also talking about a qualitative evaluation of the product itself. This has come to be conducted against sets of criteria which have more or less acceptance within the society or group which is conducting the evaluation. Thus we have to consider evaluation as a process which, in several manifestations, is applied to the development of programs, their selection and use, and their effectiveness in the learning process. Evaluation may be many things to many people and there has been much written about the use of evaluation within education as a political device. Perhaps we do not need to go as far as David Clark of London University who, when asked at a conference whether he had evaluated the interactive video program which he had been demonstrating, looked for a moment into the air and then said,

> "Evaluation, ah yes. That is what one pulls out of the cupboard when one wants to stop something happening."

SUMMARIES OF THE SITUATION IN EACH COUNTRY

Canada

The Canadian paper covers evaluation activities in the provinces of Alberta, Ontario and Quebec and although education is a provincial responsibility, it is noted that provincial Ministers of Education have all agreed to share their computer software evaluations using a data base established by the Council of Ministers.

The database provides a central source of descriptive and evaluative information on educational software and a standard data collection instrument and record format are used incorporating a standard set of evaluative criteria.

Although a common data system is used, provinces do not operate or fund computer program production, evaluation, distribution and utilization in the same way and the paper reports on activities in the three provinces which are most active in the field.

The most interesting aspect of Canadian practice in this field is the use of a standard data collection instrument and the exchange of information resulting from the use of a database by teachers and others in different provinces throughout Canada. The use of such a database must lead to considerable savings arising from the avoidance of duplication of evaluation effort and interesting exchanges of information and views by teachers on different software package.

Japan

The most significant development in Japan likely to lead to the enhancement of the quality of educational software was the publication in 1985 by the Sub-committee on Educational Media of the Social Education Council of a report entitled 'A Manual for Development of Educational Software'. The publication provided an outline of basic and desirable criteria for software development and it was published in the hope that it would serve as a manual and guide for educational software developers. It was was also hoped that it would be used as a guide to assist in the selection of appropriate high quality materials for the classroom. The report was primarily concerned with the development of instructional software for school use.

Another major development has been the foundation in 1986 of the Centre for Educational Computing which was jointly established by the Ministry of International Trade and Industry and the Ministry of Education, Science and Culture. Responsibilities of the Centre include research and development of basic software technologies in educational computing and dissemination of educational computer software.

Hungary

At present in Hungary, plans are being made to measure the efficiency of computer based educational activity. Such activity is largely limited to out of class settings (eg study circles, optional lessons) rather than to mainstream education but with software production and hardware availability increasing, the current use of computers in schools is likely to change. At this stage software production is seen largely as a central responsibility involving professional educational editors and teaching program makers.

It is believed that a number of investigations will be informative and these will be carried out in future. It is planned to carry out behaviour probing micro-tests, formal interviews, cognitive input-output analysis, operational input-output analysis, unrealised observations and information interviews.

The specific behaviour of pupils and their spontaneous questions concerning educational computing techniques can be investigated in these ways. The cognitive input-output tests are a traditional way of measuring the knowledge pupils had before the use of the computer and that which was gained by its use. The operational input-output analysis should provide answers as to how pupils approach a particular problem with a technique used in teaching. It is expected that the use of a computer program will not merely increase the knowledge of the children but will advantageously affect the quality and speed of their thinking. During the unrealised observations the behaviour of the children will be recorded so that information can be obtained about what the children say and do while they are using educational computing techniques.

Italy

In 1974 the Italian Ministry of Education abolished the national centres for audio-visual aids and set up new local centres and since that time there has been no central body with responsibility for planning, producing and distributing audio-visual and computer software for use in pedagogical situations. Software development in Italy is being conducted by private companies (sometimes in association with hardware manufacturers) or by individual teachers. It has been noted the 'pirating' of software (making illegal copies) is widespread and that difficulties in protecting software provide serious disincentives to publishers who might be contemplating the provision of substantial funds for the development of high quality software. One possible solution to this problem for publishers is for them to aim at a wide distribution of low cost software so that originals with manuals will be preferred to pirate copies which usually come without any support material.

A number of individuals and groups have drawn up educational software/evaluation forms and one of these is included in the Appendix.

Netherlands

Although theoretically free to buy learning materials of their choice in the new area of informatics, Dutch schools are guided by the government by means of various projects to buy selected and recommended hardware and software. Commercial publishers have been hesitant about entering the software develop ment market and the regular production of high quality educational computer software is being carried out largely by government institutions some of which are only partially subsidised. A number of individual teachers have become involved in local software development initiatives but it is not expected that many high quality educational software packages will emerge from such initiatives.

With schools free to make their choices from the supply of available software, decisions are often taken by teachers playing the role of pioneers as regards the introduction of computers in the school and evaluation criteria. For some time no standard series of questions was used for evaluation purposes and no uniform criteria existed. However, from 1983 to 1985 as part of a project based at the Department of Education, University of Utrecht, an instrument was developed to serve as an aid for the evaluation of educational software. Apart from a description of the aims and target group for the software, the instrument distinguished the following categories of items to be used in the evaluation:- aspects of use and lay-out; educational effects; and goal effectiveness. It was found in practice that not all aspects of the software to be evaluated could be covered by the instrument and particular problems were encountered in attempts to evaluate open-ended or content-free software. As a result, attempts are now being made to construct a complete and 'objective' evaluation process consisting of field-testing (two teachers test the software independently of each other); a subject-oriented test (testing whether the software is correct and valuable as far as the subject is concerned); and a general educational test (based on the original instrument described above). The information which results from this evaluation process is published as a review in an educational computing magazine.

In the Netherlands, there is no structural evaluation of pupil improvement, improvement of learning overall or improvement of classroom practice in relation to the use of educational software.

Scotland

Since its foundation in 1980, the central government funded agency the Scottish Microelectronics Development Programme (SMDP) has evolved a sophisticated software development procedure which incorporates extensive formative evaluation. In its early days SMDP accepted contributions of software from teachers in project centres which had been provided by SMDP with equipment. The intention was that SMDP should enhance and distribute the software at material cost to schools throughout Scotland. In practice, it was usually found that the contributed software embodied sound educational principles but suffered from technical inadequacies and it was decided in the light of this experience to encourage teachers and curriculum working parties to channel their energies into the production of specifications for computer software which could then be taken up by professional computer programmers.

The process of software development used by SMDP is complex, with many educational and technical facets and main stages in the process are: pre-specification; specification and coding; documentation; field testing; and formative evaluation. A summative evaluation report form was supplied and has been incorporated in the Appendix .

Very brief details were supplied on the selection, evaluation and use of commercial software in Scottish schools.

England, Wales and Northern Ireland

Although individual schools are at liberty to purchase equipment and software as they wish, central and local government projects, schemes and services have had a major influence on individual teachers' and schools' practices in this field. In the period 1981-83 various government schemes led to the free or subsidised provision of certain makes of microcomputer. The wide availability in schools of microcomputers of a given type stimulated the software market and in the period 1981 to 1986, more than 2000 items were produced through the government-sponsored Microelectronics Education Programme (MEP). Commercial publishers became involved in duplicating and marketing software developed by MEP funded projects but after initial enthusiasm, many publishers dropped out of the educational software market as some of these felt that the limited size of the market and the high costs of software development would leave little or no profit for them. Individual teachers have also been involved in software development on a part-time basis and many business or professional programs are used.

In general there are no widely used criteria for judging software for use at individual school, local authority or national levels. Teachers are very conscious of reviews and many are available in educational publications. Many teachers make judgments based on their own practical experience of sampling the material and their views on how the material could be used in the curriculum and how it might fit their own teaching style. The view of trusted colleagues are considered valuable guides to quality and use of software. Only technical and accuracy criteria are likely totally to eliminate a program from further consideration.

Once software has been selected, its use in the classroom must be evaluated. In the early days, when drill and practice programs predominated, measured improvement in performance was a significant criterion. However, such improvement in performance is now seen as irrelevant and the general consensus seems to be to judge software on whether children are seen to undergo a useful experience that may assist them in developing useful skills. The 'illuminative' approach to evaluation is now widely used and case studies based on this approach have been published by MEP and have appeared in many educational journals.

USA

Executive control of all US Public Schools is decentralised and vested in the elected or appointed School Board of individual school districts. Although both federal and state Departments of Education recommend curriculum goals or changes, all policy decisions related to such recommendations are made by district Boards of Education. Individual schools may operate with considerable freedom within general curriculum guidelines adopted by the district Schools Boards. Each schools district must adhere to an approved budget and expenditure of funds is approved by the School Board - most instructional materials are purchased through a central district office.

Most software available for school use has been developed by commercial publishers often in collaboration with outside 'experts' including practising teachers.

There are many sources of information about instructional software and major evaluation projects provide their evaluations either through on-line databases or by mail for an annual fee. Lists of recommended software are published, publishers catalogues are widely and freely available and software reviews are published in teachers' magazines and periodicals.

Because of widespread dissatisfaction with courseware quality, several major course were evaluation projects have developed in USA including the Micro SIFT Project at University of Oregon and the Educational Products Information Exchange (EPIE). Several evaluation centres are operated by individual state Departments of Education and district School Boards and these centres have developed formalised criteria and formats to which all participating evaluators are expected to adhere. There is general agreement that any formal courseware evaluation must include testing the program with representative students. Apart form that, a wide range of particular criteria can be grouped into three categories: content characteristics (for example, accuracy); instructional design (for example, effective feedback); and technical factors (for example comprehensive user support materials). A single summative conclusion is also usually stated. Most of such formal evaluations are used with reference to a very specific form or checklist (samples are provided in the Appendix).

In spite of well-structured evaluation procedures, the conclusions reached by individual evaluators or by different evaluation projects have been found to differ significantly - in some cases one evaluation will highly recommend a program while another will conclude that it should not be used. While these forms or check lists provide useful guidance, it is clear that software cannot be reliably and objectively evaluated or selected on the basis of this guidance alone.

Chapter One: The Context

The differences between the educational systems in different countries are often greater than at first perceived and certainly greater than is allowed for in the writings of many scholars. At the end of the day the most important differences are to be found in the attitudes of the classroom teachers. It is a dangerous generalisation but there seems to be a standard national attitude to education in each country. This conditions not only how the teachers think about teaching but how they perceive themselves. One has to ask why teachers in neighbouring countries can come up with such different answers to the same question. Why, for example, do the teachers in the Federal Republic of Germany believe that there is no place for the computer in the school apart from the last years of secondary school; whilst the British primary schools appear to be making use of the computer for complex resource- and project-based education? Why is the teacher in the French school seen in such a majesterial role whilst a similar person in the USA regards his or her professional role to be alongside and supportive of the pupils rather than in front?

There is no question whether one system is *better* than the other. That is not the point under evaluation here. Each system produces the people that is needs. What is important is the *difference* in role-perception of the teachers and the methods which are dependent on that perceived role. This will be the major conditioning factor on what sort of evaluation is carried out, what sort of programs are sought and the criteria which are given priority. A major problem is that one can only with difficulty and with anecdotal evidence build up a picture of the teachers' perceptions. One can however give a more objective view of how the schooling system is organised since this will in itself have a major effect on the production, selection and use of software.

U.S.A.

If we first look at the American situation - since that is the context for much of the evaluation of educational software - we can provide a marker for the other seven lands under consideration. In the United States of America, education - like most aspects of society - is, in theory, controlled by the American people. However, various district, county, state, and federal agencies have been empowered to operate the system.

Executive control of all U.S. Public Schools is decentralized and vested in the elected or appointed School Board of individual school districts. (A "district" may be as large as a county or parish - the third level in the nation-state-county hierarchy - or as small as a village or rural township - the fourth level in this hierarchy.) Large cities usually consist of several administrative districts under a single governing Board of Education.

Most of the 50 states have county Superintendents of Schools. However, in only a few (Florida comes to mind) does this Superintendent run a school district complete with a Board of Education. Rather, his/her office is advisory to local school districts, with no real authority. Sometimes (for example, in Michigan), county offices contract with local districts to provide special services which the individual districts are too small to provide (for example, Special Education for the visually- or hearing-impaired). Administrative control is vested in a Superintendent, various Assistant Superintendents, Supervisors (by curriculum area), Principals (Headmasters), Assistant Principals, and so on.

At the state level, a Department of Education may establish state-wide performance standards, develop curricula, administer grants-in-aid for special student populations, and - in several states (among them, California, Florida, North Carolina, and Texas) - supervise a complex "adoption" process for instructional materials (chiefly basic textbooks).

The federal Department of Education establishes very broad goals and guidelines for U.S. education systems, but does not have any enforcement power. It also awards grants for major educational projects,

such as improving inner-city schools or developing an elementary science programme. Local School Board members are informed of federal and state recommendations by correspondence from agencies of those governments and in publications from a national association of school boards.

Curriculum Control. Although both federal and state Departments of Education recommend curriculum goals or changes, all policy decisions related to such recommendations are made by district Boards of Education, after recommendations from and consultation with appropriate administrative officials. Curriculum guides (containing level- by-level objectives, suggested instructional activities, and recommended materials) are then circulated to each school in the district for use by the teachers. Many individual schools may operate with considerable freedom within the general curriculum guidelines that are adopted by the district School Boards.

Purchasing Decisions. Most educational costs are funded through local taxes and special bond issues (many of which have been voted down in recent years). Additional funds are provided by many states at a fixed per-student rate. Some states also make "equalization" payments to poorer school districts.

Each school district must adhere to an approved budget, which includes monies for instructional materials of all types. In all cases, any expenditure of funds is approved by the School Board. Except for minor purchases, most instructional materials are purchased through a central district office - to enable the school district to take advantage of bulk discounts and to minimize confusion. In some districts, each school specifies which materials are to be purchased. (In the so-called "adoption" states, a mass purchase of approved texts is often made by the state. Individual schools then request the desired number of copies from a central book warehouse.)

Some specific projects - microcomputers for inner-city schools, Braille texts for visually-impaired students - are funded, in part, by state or federal grants (usually "seed money"). In those cases (and only in those cases) some control is exercised by the government, and then only in the acceptance or rejection of proposals and in audits of project expenditures.

Many school districts have an Assistant Superintendent or Supervisor of "Instructional Materials" (various titles are used) who is responsible for setting policy and for acquiring materials of an ancillary nature, including hardware and software. A few districts have a Director of Computing (also under various titles) who addresses policy and acquisition only of hardware and software. This person is often an "interested teacher" with part-time teaching responsibilities (except in very large districts). Or it may be a "computer specialist" whom the School Board originally employed to handle administrative computing, and who then "generalized" to become a specialist in educational computing! Money for the development of courseware, although rare, is also administered by the district School Board.

In this situation the administration of the schools is complex because of its decentralisation, but the school system itself is relatively simple. The K-12 grade school system is widely accepted.

The American system has been quoted at length because it contains in it many of the elements which are to be found in other educational systems where control and decision are distributed throughout the different administrative levels and structures. But if we compare this with other countries stark differences begin to emerge.

Netherlands

In addition to the school system there is also a system that deals with teacher training and with the provision of support and advice to all Dutch schools. Most of the time it is these institutions who initiate and implement educational policy. Whether the schools really apply the measures originating from this policy is another thing. As in most big organizations with a top-down approach to policy making, the "weak link" is found at the bottom, with the implementer of the policy, the teacher in the classroom.

In the Netherlands the situation may be described as follows: between the Ministry of Education, where overall policy is initiated, and the teacher in the classroom there operates a huge power zone of subsidized institutions that exist to support and advise schools and teachers. Together with the teacher training centres, they exercise a great influence on the way policy is applied in school practice.

To explain this, it is important to understand that since the beginning of this century the Netherlands have adopted a system in which institutions as well as schools founded on a religious or pedagogical basis co-exist with the official public institutions and schools (founded by local or regional authorities). As a result the organizations which founded these private schools enjoy considerable influence. In Dutch these umbrella organizations are called "koepels". They are organized in four associations.

1. public education (government initiated)
2. private education on a special pedagogical basis
3. private education on a Roman Catholic basis
4. private education on a Protestant basis

The three "private" associations receive for the funding of their schools exactly the same amount of money from the government as do the public schools. Of course they have, in addition to this subsidy, their own sources of income, which make their schools normally more well-to-do than public schools.

In view of this system, there is no real centralisation of the concrete decision-making about the contents of the Dutch education programme.While the structure of the curriculum is fixed by the central government in the so-called "leerplan", the filling in of these guide lines (number of hours per subject, choice of learning materials) is left to the schools themselves. Central government merely decides what subjects are to be taught in the different types of schools and provides a rough description of the kind of knowledge which has to be imparted.

The decision about what is provided or bought for the schools in the way of materials is generally left to the school. There are of course a few exceptions. As a part of the new law on "Voortgezet Basisonderwijs" (prolonged primary schooling), in which the bridge period in secondary education (one year) is extended and enhanced with more subjects, all of these secondary schools need technical equipment in order to run technical subjects. Central government has prescribed the exact amount and types of equipment needed. The budget provided by the government for this use has been tied to this description. This kind of centralised approach is not the usual one when the question of buying learning materials arises.

As for the buying of computer hard and software for use in schools, essentially the same rules apply as for acquisition of other learning materials. Schools are free to decide what to buy. However, certain factors are involved which do limit this freedom of decision.

In the Netherlands, as in most western countries, a National Plan has been developed for the introduction and use of computers in education. This plan is called the INSP (INformatica Stimulerings Plan). One project among others evolved from this plan is the NIVO (Nieuwe Informatietechnologie Voortgezet Onderwijs)project aimed particularly at secondary schools, of which there are about 2250. The plan is to provide each school with 11 MS-DOS Personal Computers, of which nine formed a network (eight slaves and one file-server) and two were subject-oriented colour display machines. The development of educational software would be stimulated by the government.One notable effect can be reported: a standard operating system for microcomputers in secondary schools has been provided. No such project has (yet) been started for schools in primary education.

Within the framework of the INSP, a number of Development Centres (for educational software) have been established. In these, small teams of educationalists on one hand encourage initiatives in the field, and on the other hand come up with ideas of their own for creating new or converting existing (sometimes foreign) educational software. The ministry of education has provided the money. In this way the supply of educational software for this particular area is influenced by the government.

For secondary education, an experiment is being started that is known as the "software-cheque" project. In September 1987, all secondary schools will receive a cheque worth two thousand guilders, which must be matched by the schools, to spend on selected educational software. The selection is made by a committee controlled by the government. The message is clear: you can only buy software backed by the government. In this way the government indirectly controls the purchase of learning materials.

All things considered, one can argue that although theoretically free to buy what learning materials they like in the new area of informatics, Dutch schools are guided by the government (by means of projects etc.) to buy selected and recommended hard and software. (This is, at least, very much the case for secondary education). Schools buying different hard- or software can reckon on little support from advisory institutions. Although this sounds negative, it is not in general seen that way. Suggestions have long been made that the government should set a standard for hardware in the educational field in order to resolve problems originating from the great diversity of machine types used in schools. The Netherlands is a small country; development of educational software is very expensive when seen in the context of the Dutch- speaking market, and the more so when you have to invest in converting your program into five or more different machine-versions. No wonder that commercial educational publishers did not leap blindly into this market.

Thus whereas there are similarities between the USA and the Netherlands in that there is a strong degree of local and school control of the practice of education the role of the central government and its appendant organisations is radically different.

Italy

Italy, which has a more unified school system, has a different organisational pattern. Here we need to distinguish between two sectors in state education: schools and vocational training. The former came directly under the Ministry of Education, while the latter comes under the Regions even though it has links with the Ministry of Labour.

In 1974, the Ministry of Education published new measures abolishing the national centres for audio-visual aids and set up new local centres, to which powers and responsibilities for educational technology were transferred.

The new bodies with decision-making powers were:
* IRRSAEs (Regional Institutes for Research, Experimentation and Training), one per Region;
* School district: a smaller territorial subdivision of the provinces of which Italy has about 400.

It is important to notice that these two types of institutions are not decentralised offices of the Ministry (these are still based in the provincial education offices) but democratic bodies regularly elected by the school population (teaching and non-teaching management bodies, students, parents).

For this reason, conflicts sometimes arise between them (they have adequate finance at their disposal) and the ministerial organisations sometimes charged with similar tasks. The new bodies operate with part-time or full-time personel temporarily seconded from school (generally managers or teachers, who have had to start their activities from scratch, facing all the organisational an logistical problems which that involved).

Under these circumstances many issues have been put aside, among them that of educational technology and teacher training in this area. There was no body or office to undertake the technical organisational aspects of planning, producing and distributing audiovisual and computer software for use in pedagogical situations.

It is, however, the Ministry of Education that draws up teaching programmes, even though individual teachers are relatively free to use whatever methodologies they think most useful for their implementation. The reforms of 1974 led to the abolition of the national centre which - as in other countries - coordinated production and activities in the field of educational technology.

The Ministry usually allocates funds to the different types of schools according to the projects presented. It is necessary, however, to distinguish within upper secondary schools between those with administratieve autonomy (such as technical schools) and those without administrative autonomy (such

as Lyceums).

The former obtain considerable finance for well equipped laboratories, while for the latter it is hard to purchase even basic equipment.A great many high-level training activities are being organised for teachers, thus making it possible to improve the poor situation in the schools.

Canada

The apparent lack of central control seems here to militate against overall production and provision of software to education, though as the USA shows there is nothing inherent in a federal or regional system which runs counter to this aspect. Canada has a federal system of government and is made up of 10 provinces and 2 territories. Provinces and territories have the responsibility for education according to Canada's constitution, the Canada Act,. Software approval therefore comes under the aegis of each provincial or territorial government.

The ministers of Education in Canada have a forum for discussion of common and related concerns. This forum is called the Council of Ministers of Education Canada. The Ministers of Education have all agreed to share their computer software evaluations using a data base established by the Council of Ministers.

The data base for educational software provides educators with a central source of both descriptive and evaluative information on educational software in both official languages (English and French).

All 10 provinces and the territories are participating in the project and are responsible for providing the original data for inclusion in the database. The data base design and a record format for the descriptive and evaluative information were established by provincial representatives on the Council of Ministers of Education Canada task force on Educational Software. A standard data collection instrument and record format is in place and incorporates a standard set of evaluative criteria used by all participants. To date the majority of items listed comes from Alberta who began the process of courseware evaluation.

However, having a common data system, and descriptive and evaluative criteria does not mean that provinces operate or fund computer program production, evaluation distribution and utilization in the same way.

England

Though the schools in England are not federally organised, the chain of command through the local education authority and the school itself produceds, according to the English correspondent, a situation as complex as any to be found under other systems. For many, the education system in Britain is a perpetual puzzle, and that is particularly true in determining the responsibility for the curriculum. There is the mythical story of the visitor who asked a teacher if he was responsible for the curriculum. He replied that it was the headteacher, who said it was the local education authority, who said that it was the government and the examination boards. When these two bodies were asked, they said it was the teacher! So where does the responsibility lie?

At present, the answer is in all of those to some extent, depending on what one means by the curriculum. The teacher determines what happens in the class for which he is responsible, but if the children are to succeed in examinations, he is very mindful of the syllabus which the boards lay down. However, the balance between the subjects that the children take is determined by the Headteacher. He, however, is conscious of the policy of the local education authority (LEA) and responds to their requests while keeping the particular needs of the children in his school in the forefront of his decisions. The local education authority is conscious of the wishes of the government, and responds sympathetically to the mood of the times and the ideas put forward. However, these are very general, up to now, suggesting that a subject should be given greater weight, or that a particular approach to the teaching of it would be beneficial. This is more clearly stated if there has been a report from a curriculum development group which recommends a certain direction.

However, even if the government makes such a recommendation, the teacher may find it difficult to implement, and in the end, it is what happens in the classroom that counts. Thus, in the end, there is no check or authority over the individual teacher, and he holds the ultimate responsibility for what is presented to the child. Such arrangements as these demonstrate a more decentralised than centralised approach to the curriculum, though the politics are concerned with one part of the cycle taking more authority at one time, another at a later time.

Just as judgements on the curriculum are distributed around a cycle, so this is reflected in the purchase of hardware and software. If the teacher does not support any purchase, then it will not be used in his classroom. However, if the teacher wishes for an item and there is no money to buy it, then he will not get it. Thus, a compromise between the two has to be reached. While individual schools are at liberty to purchase equipment and software as they wish, and many of them do, some LEAs offer schools support schemes of maintenance, training and special prices for items of their own choice, as recommended by their professional advisory staff. Such benefits are considerable for the schools under their direction and are thus usually accepted by the teachers.

Central purchases by the government have also led to equipment for schools. In the early schemes in 1981/3, the computers were selected by the government, after consultation with professional advisors, but in future schemes, it is likely that guidelines on configuration will be given although names of models will not be cited. Software schemes offered by government have not identified items, but have offered a financial subsidy for purchase up to a specific total. The actual items were then selected by the schools or the LEA advisors. In all these government schemes, it has been usual to require the schools to make some financial contribution towards the purchase, although in some instances this has not been necessary.

Japan

Japan, where the schools show many similarities with the American and British structures, has not advance so rapidly into the integration of computers into schools and therefore has not yet had to deal with some of the issues. However it is necessary to understand the approaches being made by the Ministries in order to make comparisons. The Japanese Ministry of Education, Science and Culture is responsible for the integrated planning and coordination of the development and improvement of education at different levels and in different areas. In Japan there are 47 prefectures, each of which is further divided into a number of municipalities. In every prefecture there is a prefectural board of education which is responsible for the administration and management of services relating to education, science and culture.The Ministry of Education has the power to provide guidance, advice and financial assistance to local education authorities, and to require them to submit such reports as deemed necessary relating to educational activities under their jurisdiction. The Minister may also give local education authorities orders to correct or improve their policies or measures, when he finds that these violate the provisions of laws and regulations, or impede the primary objectives of education.

The responsibility for financial support of public education is shared by the national, prefectural and municipal authorities. Each provides support for its own educational activities by means of funds derived from its own taxes and other income.The policy of introducing computers into education is carried out along with that of national and local organization of educational administration and fits within the patterns which have already been established for textbooks. Educational aims and instructional objectives are also controled by these organizations through the national "courses of study" and textbooks. All elementary and secondary schools in Japan are required to use textbooks in the classroom teaching of each subject. In principle, textbooks used in schools must be either compiled or authorized by the Minister of Education, Science and Culture. Most of the textbooks currently used in schools are published by commercial publishers and merely authorized by the Minister. Textbooks compiled by the Ministry itself are limited to those for certain vocational subjects in upper secondary schools and those for special schools for the handicapped, both of which, due to a very limited market, no commercial publisher

would be willing to publish.

The decisions on textbooks for use in local public schools rests with the local board of education administering these schools. On the other hand, textbooks for use in national or private schools must be adopted by the principal of each school.

There is as yet no nationally authorized system for the development and production of educational software, but the procedure for textbook production seems to be influencing the national system for software production.

Thus whilst the American and Canadian systems leave a great deal to local administration and decision making, and the English seem to concentrate much of the control at the local authority level, decision making in Japan seems by contrast to be much more in the hands of central government through the provincial authorities. As will be seen this tends to create a specific focus for software evaluation - as with textbooks more at the formative stage rather than teacher evaluation of commercial products.

Scotland

Scotland, though having a school system which is simple in comparison with many of its European neighbours (primary and secondary comprehensive with some of the schools being Roman Catholic but within the same form - there are also a very few fee paying schools), has a much more centralised system than England. This has had a considerable effect on the way that software development, and therefore evaluation, has found form.The Scottish Education system is divided into three main sectors. The primary sector deals with compulsory education for children aged 5 to 12 years, over a seven-year period, and the schools are designed and equipped for this age range. Classes are mixed in sex and ability, and sometimes also in age, depending on local staffing and class size. Pupils normally work with the same teacher all day every day for a year and there is little movement between rooms. The curriculum is broad-based and non-compartmentalised, and the children normally work in small groups. At the end of the seven-year period, pupils transfer to the Secondary school for which their Primary school is a "feeder".

The Secondary sector covers the remaining years of compulsory education to the age of sixteen, and post-compulsory education, which normally lasts for a further one or two years. The secondary school is divided into departments and as many as fourteen departments (e.g. Mathematics, Home Economics, Art) can exist in one school. Compulsory schooling ends at the age of sixteen. Pupils then face a range of choices. They can remain at school and continue studying for certificates at a higher level, either in order to improve their employment prospects or to meet entry requirements for Further and Higher education courses. Alternatively they may leave school and study immediately at a Further Education (PE) college for certain other certificates, also awarded by the Scottish Examination Board. Finally, a new choice available is to remain at school but to study there, just as if they had left school and enrolled in a FE College, for an alternative set of awards new offered by the Scottish Vocational Education Council (SCOTVEC). This new system is intended to promote a close association between the schools and the FE colleges in terms of the post-compulsory curriculum.

Schools and FE colleges are the responsibility of the Education authorities, of which there are twelve in Scotland. While the Authority creates overall policy, the actual details of curriculum and methodology are left to the institutions themselves. For schools, the Consultative Committee on the Curriculum (CCC), a national advisory body, is responsible for advising the Government on the curriculum; this influences national examination syllabuses, which in turn influence the Secondary schools. The CCC also produces papers and materials to assist schools in their choices, although the schools themselves have freedom of choice of curriculum, method and materials selection and purchase.

At regional level, advisors appointed by the Authorities assist in the implementation of Regional and national policy, and provide advice on the selection of hardware and software. The Scottish Council for

Educational Technology is involved here also, in some cases by directly advising schools in some cases, but more usually through links with the Authorities in support of their policy decision-making.

The degree of central organisation within the Scottish system will become more apparent in the later consideration of software development. One might expect the educational system of Hungary to be even more centralised. The following description from the Hungarian correspondent appears at first to back this up.

Hungary
In the development of the Hungarian school system, emphasis has been placed on content and methods rather than on alterations in structure. The system of education can be outlined as follows:

The age-group 3 to 5 may attend nursery school. Participation in this form of education is not compulsory but according toestimates is becoming quite general, so that even today, 90% of the age-group participates in this form of education.

The basis of our school system is the 8 years of compulsory education, followed by pupils aged 6 to 14 years. This is followed by secondary education, in the form of:

apprentice school - 3 years
professional training school - 4 (or 5) years
grammar school

Training and grammar schools award the so-called final examination certificate which is necessary for further education. Professional training schools also award a certificate in the chosen field. 92-93 per cent of the relevant age-group complete elementary schooling and about 90% of them enter one of the secondary schools. This means that about 80% of the relevant age-group undertake some study at secondary level.

The Hungarian system of education is on one hand highly centralised and on the other decentralised. This means that teaching and lesson schedules, books and other media are fairly uniform in anyone particular kind of school. These are prepared and approved by various institutions of the Ministry of Education (National Pedagogical Institute, National Centre for Educational Technology, Textbook Publishing Company, Company for Production of Teaching Media, Institute for Scientific Organisation and Information Science).

Ministry of Education control is exercised partly through its institutions and partly through the county councils. Pedagogical institutes have been established in each of the 19 counties in the country as well as in the capital city. These institutes are independent and control and assist the actual work of the schools through a network of advisers and consultants.

Schools are maintained by the local (municipal or county) councils.Occasionally, the Ministry of Education also provides financial support for schools. This happens e.g. when personal computers are purchased by the schools. This made it possible for the number of computers in the 4.5 thousand schools to rise by 10 thousand machines in 1986.

In completing this section on the educational context of the study it is worth noting the opening remarks from the Hungarian paper.

> In Hungary today it is considered out-of-date to carry out sporadic fundamental changes in order to adjust the function of the schools to the social and economic needs of society. Development is determined in practice by the directions forecast by our national social development programs. The difficulty that results is that one is constantly having to think over

a perspective of at least 25-30 years. Beyond outlining main features, it is not possible to determine details because that would risk pedagogical developments implying methodological innovations. Our development strategies aim at "optimal function" in theoretical questions, but to leave the door open for the elaboration of details or for the design of particular variants within the main framework later.

> The functions of our school system are determined by the cultural leadership of the country in line with our education laws. An important function is the continuous development of cultural education. A modern cultural education is considered as a combination of motivation, systems of knowledge, customs, beliefs, needs and activities. At our present stage of scientific and technical development, our schools can meet social requirements only if efforts are made to teach basic knowledge, to stimulate demands for further education and to prepare children for continuous self-development. Information science (including computer technique) is today considered an integral part of basic knowledge.
>
> It should be recognised that in the two or three decades ahead less well-known or competely new technologies are likely to spread. Scientific, technological, economic and social circumstances will harmonize more fully.
>
> These trends presuppose a rapid development of scientific and cultural knowledge in the whole population. Schools can help in this development by providing high-level and broad-based cultural education, including as a matter of course up-to-date computer science and its applications. We believe that the content of general and technical knowledge should be such as to enable pupils to progress to further education and retrain for other qualifications in later life.

The Hungarian situation highlights the dilemna which is facing many other countries but may not be articulated in quite the same way. What can be seen in a community which is rich in commercail products and has the luxury of a large language market may not be applicable in a country where the national policy strives to reach its goals through other means. No substantial tradition or good policy for software development has yet been established because the whole school computer system in Hungary was only started in 1983.

Software production in Hungary is increasing, but the use of these programs is limited to out of class settings e.g. study circles, optional lessons. The reasons are manifold. Teaching schedules, books, other media and traditional methods of teaching do not facilitate any integrated use of available software in the learning-teaching process. The available software sometimes preserves out-of-date teaching methods showing that views and policy for software production need to be changed. Software prepared for the teaching of a particular subject may attain its optimal effect only when achieves "optimal mass" in terms of quantity, and covers the whole topic of the subject unit or concentrates on a particularly difficult point.

Experience corroborates the hypothesis that computers are most successfully applied in those schools where educational technology is at a reasonably high level anyway. It has also turned out, that teachers with poor backgrounds in educational technology are enthusiastic about using these "magic' tools, but unfortunately do so with only a low level of efficiency. This obviously means that continuous help should be given to the teachers for the development of their knowledge of educational technology. This can usually best be done through participation in organised courses.

Computer programs are evaluated in Hungary in line with the above-mentioned principles. The main point is how a program developed for a particular teaching purpose helps in the achievement of the educational goal.

Professional programs (originally made for technical and business groups) are hardly ever used in the schools. Their use is, however, widespread in the professional training schools for which their special

content and background assumptions are relevant to the teaching programmes. The efficiency of the so-called open programs is much higher. Their main advantage is that the teacher may extend or change their content as the local situation requires. Didactical programs - due to their special educational content - seem to be most suitable for use in the general grammar and in the professional training schools.

.In the Hungarian opinion computers in schools provide educational profit only if the software conveying the content is an organic and integral part of thecurricular plan. It is thought that the problem can be solved - at least in those countries where the educational system is similar to the Hungarian - only if a "critical mass" of computer programs are made centrally and subjected to uniform principles. This does not mean that teachers should be deprived of the opportunity for computer programming. Teachers should indeed be involved in this activity, because otherwise they would be unable to deal with the need for local innovation. Experience shows that programs made by non-professionals are rarely suitable for nation-wide dissemination, but are suitable for use locally or regionally.

In comparison when a country, such as Italy, has no central institution, market forces come into play in a situation where the market itself is very fragmented. But most software in circulation is however pirate copies. Because it is difficult to protect the software each programme is immediately spread through the fastest and most mysterious channels.

This raises serious problems for the production of good-quality software and for the cost of the "general purpose"programs, the astronomical price of which is thought of as morally excusing the school in pirating what they will never be able to afford to buy. In this vicious circle it is those who would like to produce good-quality software for sale at a reasonable price who lose out.

There is also a big growth in software produced by teachers themselves or by specialized people who develope it out of personal interest or within training and experimentation projects for which they are already fully remunerated. In this situation, software is being produced by public institutions and above all by private individuals according to thier own various needs, and not according to any general development plan.

Chapter Two; Development of Software

Is the development Commercial?

It was felt important to begin from a question about the commercial nature of the predominent software in the schools. If commercial pressures were the driving force we would have expected different sorts of evaluation to be in operation than if the development process was in the hands of the educational establishments.

Italy demonstrates clearly the two strands which have emerged elsewhere. The lack of overriding control has stimulated a lot of private enterprise which is producing the majority of the software. There is also a big growth in software produced by teachers themselves or by specialized people who develop it out of personal interest or within training and experimentation projects for which they are already fully remunerated. North America and to some extent Britain have developed in a similar way, as will be seen, but it is noteworthy that the Japanese have consciously begun in a more controlled way.On December 11, 1985, the Subcommittee on Educational Media of the Social Education Council published a 42-page report entitled "A manual for Development of Educational Software", including evaluation items. This manual describes its aims as follows:

> "The development of high quality software cannot be realized through the developers' efforts alone; a great deal depends upon smooth channels of distribution and appropriate utilization. However, even though educational software (and, in particular, instructional software for school use) is now becoming more and more readily available, a review of the situation concerning microcomputers in education shows that not nearly enough is being accomplished.

Software for the educational use of computers is developed by commercial enterprises, local schools, and other educational institutions. At the moment, there is no official establishment for the registration of programs which have been developed nor any comprehensive catalogue of educational software. However, in April 1986, 958 commercial programs were counted and classified according to school levels and subjects.

The British developments with MEP and SMDP are already widely known. Here a government initiative was taken to stimulate the sales of computers and their integration into schools. Naturally these programs were also concerned with the production of software. In England the initial impetus to development in 1981 to 1986 was given by the government-sponsored Microelectronics Education Programme (MEP), which was responsible for over 2000 items. Some of these were small programs, used to stimulate development, and now superceded. However, the majority are still in use. At the same time, many small companies were set up to prepare and market software,. some of them very successfully, others producing programs of total triviality. Several of these companies no longer exist.

One tactic for the distribution of MEP programs was to involve educational publishers in duplicating and marketing them. This was a successful enterprise in that nearly 20 companies became involved, and subsequently began to develop their own software. After initial enthusiasm, however, they dropped out rapidly, as the size of the market, the pricing structure and the costs of development made it clear to them that there was no profit in this area. Now there are only five major commercial publishers in the educational software field. The proportion of educational software developed in this way was only about 40% of the total market.

Teachers themselves were not slow to see that there was potential income for them in the development of programs, and a thriving 'part-time' industry started and has continued. Some developed their materials in association with LEA groups and local curriculum development projects, some of which were located in LEA centres, and these were found to be relevant and useful and were widely used in their

localities. Together, these constitute some 20% of all the software used in schools.

Another source of software has been the materials that are used in commerce and industry, the Professional Programs. As they were not made for or directed at education, much doubt has been expressed about their applicability, but some schools have found them useful. Usually, however, they have been used in special areas like business studies, and it is rare to find a piece of software from this source used elsewhere in the curriculum. Thus only a very small proportion of software in schools originates here.

A major priority from the beginning of this technology has been to ensure that it is being used, and preferably helping children to learn successfully. To achieve this, considerable emphasis has been placed on teacher training, and a wide range of approaches have been taken. Some have involved the distribution of packs of learning materials to the teachers, either for distance learning or at courses. Many of these packs contain substantial numbers of programs which are available for use in the classroom and are therefore given free to the teachers involved. This approach is valuable as teachers not only receive the software at no real cost but also get training in the most appropriate use of the material.

Another almost free supply of software initiated by MEP, and much used by staff involved in helping children with special learning needs, has been christened the 'blue file' system. This supplies master copies of the programs to regional or local teachers centres, where copies can be made by teachers for use in their schools. For educational purposes, all copyright restrictions are waived.

Finally, programs have been available from telesoftware sources, either broadcast by the BBC as part of the teletext service or obtainable on-line from the viewdata service or the National Educational Resource Information Service. While the viewdata service can levy a charge for the software, the others are available at no cost.

The programs made available through the teacher training packages have usually been especially developed or obtained directly from educational development units financed through government organisations. The free materials and those on the viewdata service may have come from these organisations as well, but many are developed by individual teachers or LEA development projects. Commercial sources are obviously not appropriate, but teachers are always on the look-out for a bargain if they can find one. A current focus of interest is the 'shareware' movement, where costs for quite sophisticated products are low. Many are also studying the bulletin boards to see if there are free materials there, and The Times Network for Schools electronic mail service sometimes has free software available to subscribers as well.

Thus the variety of sources of software is extensive, but the quality of what is obtained from them varies a great deal. Both application and didactic programs are obtainable from these sources, though many of the computers sold for education are purchased with a number of application programs as part of the package. Word processors, spreadsheets, Logo, and databases are commonly supplied in this way. However, many of the commercial and educational developers have produced and marketed their own application programs, which they consider better in an educational sense then those supplied. These usually contain support materials and examples which are more helpful to children.

Here is it clear that a pragmatic approach was taken with all the advantages and disadvantages of a mixed economy. With the end of the MEP project software production has consolidated into the strong, high quality producers many of whom are small companies.

Scotland adopted a more centralised approach. Naturally commercial products were on sale and were bought and it must be noted that many schools bought the products of the MEP work south of the border. The Scottish Microelectronics Development Programme was set up within the Scottish Council for Educational Technology with the development of software as one of its main tasks.Software

development takes place within SCET and within individual Education Authorities. Within SCET, a Software Development Unit produces packages to meet national needs, and some individual Authorities have set up small development units modelled on the national unit at SCET, to support particular Regional priorities.

Over the past seven years, SMDP has been involved in the production of computer software for distribution to Scottish Educational Establishments through Regional Distribution Centres. This software has been developed in a variety of ways. These will be considered later when software production is considered, but the central role of SMDP must be recognised since it has shaped much of the use and evaluation of software in Scottish education.

The Dutch on the other hand have relied in the first instance on that peculiar mixture of commercially available materials (usually from the USA and often in another language) and teacher-produced programs which do not get effective distribution. Though the commercial programs were soon brought out in Dutch there was little which related to the educational market and commercial publishers have been rather hesitant about entering the software market. A number of partially or wholly subsidized educational institutions have on the other hand gone quite deeply into this field. In addition to the Development Centres (tied to a specific school-type), the Centre for Education and Information-technology (COI) has also evolved from the INSP. This institution possesses, among other things, a fully equipped department for software development, which lays emphasis on research and the development of prototypes.

SLO (Institution for the development of the curriculum) is another institution that from a researchers point of view incidentally develops software for educational aims. The few, though well worked-out, programs emanating from this institution are distributed at relatively low prices.University Education Departments are also active in this area.

Another institution that must be mentioned is NIAM (Nederlands Instituut voor Audio-visuele Media) in The Hague. Traditionally, the role of NIAM has been the production, conversion and distribution of audio-visual material for Dutch schools. A few years ago, the same activities were extended into the area of computer software. In addition to distribution (for SLO, OPSET, CDO and others), NIAM is busy converting and developing educational computer software.

From this list, it will be clear that the quality and regular production of educational computer software is, first of all, a matter for the (sometimes only partially subsidized) government institutions. Commercial publishers seem reluctant to commit themselves deeply to this market, partly because of experience they have observed in other countries (the US and UK), partly because of their own experiences. There is at present no real expectation of profits from this area.

The *project* approach, as seen for example in the INSP, still seems to be the most effective way of producing adequate educational software. Examples of this approach are the "Voorbeeldreeks" (Exemplary package) produced by the Development Centre for Primary and Special Education (OPSET); a series of programs developed in a project led by SCO (in combination with OMO, an organization of schools working together, and IBM); and the NIVO-project, where a so-called Starter-package of software is with government money provided for secondary schools. This package includes business-oriented software like *PC-Type* and *PC-File* in addition to more dedicated educational software like Dutch versions of well-known British educational programs such as *Developing Tray* and *Dynamic Modelling System* . In this last project, the government negotiated, through the official "Rijksinkoopbureau" (State purchasing department) with the suppliers of software in order to get a reasonable rebate and to demand special adaptations (where necessary) to meet the particular educational needs. This may prove to be the long-term solution to providing the schools with good software.

Local initiatives have always existed and will always continue, so long as there are teachers who are willing

(and often eager!) to produce software to meet their precise educational needs. However, their inability to cope with the more difficult computer systems and languages hampers the production of original software and their breaking through the drill-and-practice barrier, so that their products are unlikely to find their way through the official distribution channels, as against the restricted channels of the informal circuit. No substantial contribution, in quality or quantity, to the solution of the problem of the shortage of acceptable educational software should be expected from this quarter.

Thus the Dutch software production can be seen as partially commercial with a large input to education from the project-based and government funded activities. To date, whilst development has taken place in educational establishments, distribution has only been successful on a sound commercial basis.

It is when one turns to the decentralised lands that one finds the strongest influence from the commercial sector. The United States situation is well known and in Canada the major software development trends are from the commercial sector, particularly the United States. The example of Alberta where the review body is based shows clearly just how little of the software originates from their own educational institutions.

Software development does not relate directly to the needs of Alberta schools. The Clearinghouse reviews thousands of software items annually, but less than five percent are recommended for purchase by Alberta schools. Most software purchased by Alberta schools originates in the United States of America (perhaps over 90%). The Clearinghouse is contacted frequently by commercial developers to review curriculum needs. Alberta schools or practicing teachers are almost never involved in the development process. Since Alberta Education itself does not to date develop programs, teachers do so and market them themselves. Themes or subjects are decided by the commercial producers of software. The marketplace decides which programs survive.

Ontario province with its provincial program based on its custom made computer has a somewaht more centralised approach. Educational software is developed by private companies or educational institutions under contract to the Ontario Ministry of Education. Copyright remains with the developer, with the Ministry retaining a licence for use in all Ontario schools and educational institutions.

In Quebec the Ministry of Education does, however, to a very limited extent support development of some software. In 1986-87, it supported the development of 10 pieces of software for the educational market.

From this it can be seen that in those countries with a mainly commercial production the emphasis in education has to be on criteria for the selection of software. The british models show a mixture of project dvelopment and commercial production, and thus one expects to find a more even balance between selection and development criteria. The Dutch and Hungarian models show more clearly the need for evaluation during development, however because the government lead development in the netherlands is also matched by a large measure of freedom of choice in the schools one might expect criteria for selection to play a major role.

SOFTWARE DEVELOPMENT

Having gained an overview of the nature of the provision the next step was to look at the development processes themselves. Staying with the Canadian examples the major trend in software development is to follow the lead of the American counterparts. Much of the French material available for the Quebec market is translated from English versions produced elsewhere. Recently there has been some original French software produced for the school milieu, and this has been very refreshing and encouraging. Ideas for software can and do come from a variety of people. Most of the new software produced locally comes from a small group of software development firms working in isolation, trying to meet the demands of their French clientèle. If Quebec could arrive at an industrial

standard machine for the school environment, software development might proceed more rapidly. The ministry of Education in its funding of software development requires that all software be produced by a Quebec based software company. This requirement has effectively ruled out school boards and teachers from taking an active part in the development process. The importance of this should not be underestimated.

On the francophone side there is very little in the way of software support. Any material which does come out in French for Apple computers is usually an IBM translation of something produced locally. It is virtually impossible to get the rights to translate any of the American produced English software without a tremendous cash outlay.

South of the border the American position is that over half of all commercially published didactic courseware is aimed at the elementary school market (kindergarten through grade 8), with courseware for secondary schools (grades 9-12) comprising most of the balance. Although drill or tutorial programs still account for a large proportion of new courseware, the number of programs in problem-solving or simulation mode is on the increase, especially among publishers of excellent courseware (Taylor, 1986). Also, the quality of many new releases is much better than was the case previously, possibly because publishers are responding to detailed post-production evaluations carried out on earlier efforts (Wighton, 1984).

If what has been shown here relates to the financial/commercial pressures then the normal distribution channels should also have an influence on the development of software. Quebec is typical. Software sales in the educational market are handled through software firms that promote their products through direct marketing and catalogue sales. This form of marketing does not give people much in the way of interaction with developers. In nearby Ontario the schools receive copies of software through TVOntario, which acts as a distribution agent for the Ministry. This service is termed the Ontario Educational Software Service (OESS). Software developers are encouraged to market their products outside of Ontario, and assistance in this regard is provided from time to time by the Ontario International Corporation of the Ontario Ministry of Industry, Trade and Technology. In this instance the connection between commisioners, developers and users is much closer. This is a good example of a government lead project. One wonders under this structure how many wheels get invented all over again, but perhaps that is also a necessary criterion for the bonding together of the people involved in the process.

Hungary, being more centralised has more control over the process.Works may be submitted by professional teachers or non-teachers. All are evaluated and assessed by a jury (consisting of people of various qualifications, including teachers) and the best programs are copied and can be purchased by the schools. Booklets have also been issued by the National Centre for Educational Technology, and made available free of charge to the secondary schools. Programs are also available commercially. The main problem with this is that they can only be used with difficulty in schools.

The degree of central control appears therefore to act against the development of effective software by commercial houses, just as the plethora of commercial product in America and parts of Canada seems to reduce the need for educational institution based production. The vigorously mixed Dutch system ought to find a balanced position somewhere in the middle. Educational software is distributed through a variety of channels. The commercial publishers make use of the same channels which they use for distributing traditional learning materials such as books, namely specialized bookshops and direct distribution from publisher to school (by means of mailing, catalogues etc.). No standard channel exists for the distribution of products resulting from government projects or activities. While distribution of non-commercial audio-visual learning materials (including computer software) has traditionally been organised by NIAM, alternative distribution channels emerge from time to time.

Where, as in Japan, the numbers of commercially available programs is still limited the need for evaluation on the American model appears less necessary. Japanese teachers tend to prefer their own hand-made programs to those in the market. Fig 2 shows the tendency in software development at the moment.

	Elementary	Secondary	
		Lower	Upper
Number of schools responding	410	1,040	2.945
Personal development	198	548	2.115
Collaborative development within a school	52	129	845
Collaborative development within an area	32	60	53
Purchase of commercial programs	147	339	867
Supply from other institutions	82	189	446
Others	78	204	302

Fig 2. Schools' policy for software development (as of Dec. 1986)

Publishers and private learning institutions other than schools are major suppliers of instructional programs. Some tutorial type programs are widely purchased by families for their children as remedial material in relation to classroom work, or to enhance their academic achievement. The governmental budget for instructional software is limited and teachers are expected to develop their own programs. This is therefore an apparent shift of responsibility from the government and its central institutions to the teachers themselves.

At this point it is worth noting a comment from the Canadian writer which undercuts much of what has been written about the development and evaluation process. "Most schools purchase software that they have seen demonstrated, but which has not been field-tested in the schools by the teachers themselves. This course of action leads to inappropriate software."

SYSTEMS OF SOFTWARE DEVELOPMENT

Where there is such a spectrum of software provision and sources of development a consideration of the systems of development ought to elicit some useful information regarding the criteria for evaluation during the development process.

The dutch situation reveals little. There are no fixed systems for the development of educational software. Apart from some research institutes such as COI and the Education Departments of the Universities, most software developers proceed rather pragmatically. This is caused to some extent, by the terms of their subsidies (often short-term and product-aimed), and by the complex use situation (because of the wide variety of computer systems used in education, a great deal of the energy is put into conversion).While the INSP promulgates the idea of a uniform approach to educational software development, the circumstances indicate that no concrete results are to be expected.

Scotland on the other hand has used the central production unit to generate well defined criteria for the development stages. The procedure for the development of high-quality educational computer software has now matured, and the National Plan for Microcomputers in Scottish Schools (Microcomputers in Scottish Schools: A National Plan, ISBN 0-86011-102-4 (£5), available from SMDP) describes how this procedure can be applied to meet Regional and National objectives.

Since the process of software development in Scotland is complex, with many educational and technical facets, it will be worthwhile to describe the main stages in the process.

Pre-specification

The most important considerations in the development of educational software are those which address

educational issues: objectives, subject matter, methodology, target groups, and so on. An initial brief statement involving these issues can and should be made before consideration is given to technical implementation (or even technical feasibility). This statement, which is in fact a pre-specification, has several uses: preparing it helps to crystallise the thoughts of the specifier; it provides a solid educational foundation on which to build the package and its documentation; and it assists in making decisions about which specifications are to be converted into programs.

Note that at this pre-specification stage, there is no requirement for discussions with the staff who may later be involved in writing the program. It is, however, essential that the pre-specification author has knowledge of the facilities offered by computers, or can at least discuss computer facilities with a knowledgeable colleague. This is likely to increase the chances of the pre-specification going further.

Specification and Coding

Once it has been decided that a program will be developed, the job of specification can be undertaken. It is at this point that a programming team becomes actively involved.

As the specification develops, the programmer will probably start working on some of the routines or on some key frames or graphic animation, and these will be viewed by the teachers before the specification is complete. This previewing of modules of a program, coupled with discussions between educational and computing staff, is bound to have an effect on the development of the specification itself. In fact, it is difficult to separate the two steps of specification and program writing. They are two aspects of an interactive process in which, at the start, most effort goes into specification activities, and nearer the end most effort goes into programming.

Documentation

An educational computer program can be a very complex resource, requiring much in the way of support material of different kinds. It also requires an intimacy with the provisions of the package, which often bely its surface appearance to the users. Some support must therefore be provided in the form of teachers' notes, enabling teachers to makenthe best use of the package within an appropriate context. In addition, it will often be necessary to provide materials, in the form of notes, worksheets, etc. for pupils.

Guidance in the writing of educational documentation to accompany computer software is available in a SMDP publication (SMDP Presentation and Documentation Standards (available from SMDP)) and this forms a nationally-adopted standard which is now followed by all specifiction groups.

Such a detailed model is not to be found in any of the countries reported here. The Ontario project has, nonetheless, many of the steps which are to be found in the Scottish model. Software developers submit proposals to the Ministry following a "Call for Proposals" in a given area, although some proposals are funded on an unsolicited basis. Consultation with practicing teachers is a requirement in these proposals. Development is undertaken in three stages: Design, Prototype Development, and Field-Testing in classrooms. The field-tests must be supervised by a professional, external to both the developer and the Ministry. In addition, there is a continual needs assessment process involving the cooperation of teachers, curriculum developers and University Faculties of Education.

Selection of appropriate subsections of subject areas as the content of programs has a crucial role to play as seen in the MEP developments. The steps taken are however, typical of the (anglo-saxon) pragmatic approach which characterises the British and North American systems rather than the more 'scientific' approaches required as one moves eastwards into continental europe.Choosing the area in which to develop materials seems to depend more on the needs identified by the developers than on any central direction. Ideas for application programs come readily from commercial and industrial uses, and from general concepts in education of an open environment in which children can work. With an increasing emphasis on exploratory learning, such content-free programs are attractive to teachers, and new environments or supporting programs like spell-checkers are being conceived by teachers and advisers.

Ideas for didactic programs are less obvious, and arise for a number of reasons. The topic may clearly lend itself to treatment by computer. Another reason may be that the concept is difficult for children to understand, and presentation with the aid of a computer may help to clarify it. More importantly perhaps, the topic may benefit from children exploring it and identifying the results of different scenarios. Other reasons, like personal interest, the topic operating within a clearly defined system, or responding to interesting graphics, also apply.

There are many sources for these ideas, but the most valuable are teachers, LEA advisers and curriculum development groups. Other educationalists in universities and colleges contribute, but whereas the former group seem to be collectively more interested in solving practical issues, these are more interested in ideas that develop extensions of the use of the computer. From the commercial publisher side, there is much sampling of teachers' opinions, and some reliance on the views of respected authors. Attempts have also been made to devise software to accompany successful standard text books, but such programs are not notable for their value in the classroom.

Such pragmatism, an approach which at least pretends to base itself on the needs of the learner, is resistant to the distillation of definitive criteria for evaluation. From all the reports it does seem to be a fair reflection of what happens in most circumstances (except in this case, Scotland). The Dutch comment comes nearer the mark. "Ideas originate either from the field (minority) or from the staff of the project groups and institutes involved.

Ideas for software can emerge from the analysis of an educational need or from already existing software applications. Teachers are often involved in projects, but their contribution to the final product is relatively small. Because of their authority in the domain of information technology, it is generally the project leaders (educationalists, psychologists, didacticians, in short the researchers from the universities or from the advisory units) who leave their mark upon the final programs."

The Scottish contributer is aware of the need to protect against this influence. For those teachers unable or unwilling to program for themselves, the provision of a system which nevertheless allows them to be involved in the creation of computer-assisted material, is to be welcomed. The existence of professional programmers and analysts committed to supporting teacher-specifiers has put an end to the view that computers in education are the province of an élite minority who can speak the appropriate computer language, and this is to be applauded. However, it is not simply a matter of a teacher providing a programmer with a written description of a program for him to "get on with", while the teacher starts on his next idea. Professional programmers committed to education are still a rare species, and programming is still an extremely labour-intensive activity. It is therefore essential that very close attention is given to the educational considerations accompanying a specification, and also to the balance of educational and technical input as the specification and the eventual program develop. This approach is forced on the creators of software by the current shortage of experienced full-time programmers in education, but is, of course, a welcome subjugation as it can but increase the quality and value of the software eventually produced.

In the early days, SMDP were the recipients of software produced by teachers, generally through the sixty-eight project centres. This software usually embodied sound educational principles, but it often suffered from technical inadequacies, and it was one of the tasks of SMDP to make good these deficiencies prior to its distribution. As awareness increased, standards and expectations rose and it became obvious that it was generally inappropriate for the already hard pressed teachers, to contemplate writing software.

The Dutch and the Scottish writers are each arguing their own particular case. One man's helpful specialist is another man's interfering power grabber. One suspects on the basis of this study and much reading that the Dutch situation is more common. We should not be surprised if there are really just a few

people who effectively control what goes into educational software. This has long been the case with other learning materials and is now the case with commercially produced software. Even where there is central control the text book is more often than not the work of one or two people rather than a committee and certainly not the mass of the teaching profession. We seem to have few fixed criteria for textbooks but want them for computer programs, perhaps because the program itself appears to be exact and scientific.

The Hungarian correspondent presents the alternative case. Although it is not an aim, it is considered important that most teachers should be able to write computer programs, because only then can they be expected to take part in the preparation of programs with their experience and knowledge which is one of the conditions of efficient school work. Experiences at home and abroad proves that it is not, for instance, a necessary condition for writing a good book that the writer should be an expert in the field. A textbook needs to meet special requirements in terms of its language, logic, illustrations and printing. Standards for computer programs are not exactly the same but are very similar to those for books.

The USA under the pressure of a huge commercial market finds that very little courseware has been developed in any organized fashion by school systems. Certainly, some individual teachers' efforts - first just locally distributed - have eventually achieved commercial success. (The best-known examples are the simulations, *Odell Lake* and *Oregon Trail*, which are now published by the Minnesota Educational Computer Corporation.) Teacher-developers usually program in a dialect of BASIC and do not use sophisticated copy-protection mechanisms.

DEVELOPMENT CRITERIA

The USA does, for this very reason have a lot of experience of the commercial development of programs for sale into education. This production has been studied byDr Taylor who explains the situation as follows. Rather than being the result of a needs assessment or market analysis, a publisher's decision to develop a piece of didactic software is usually based on "matching" an already-existing product (such as a textbook); on "an intuitive sense" of what the market wants; or on the personal interest and expertise of the developer. Then, too, when one program is widely accepted, related or similar programs tend to follow. (Cases in point are the popular *Terrapin Logo* and *Bank Street Writer*, for each of which several supporting programs and related materials have been issued.) The concern of commercial software publishers, then, is not "what is good for the schools," but "what the schools will buy."

There is some evidence (Taylor, 1986) of a positive correlation between the production of excellent courseware and a publisher's collaboration with outside "experts," who may very well include practicing teachers. Although it seems as though little money or time is being spent on formative evaluation (Sturdivant, 1984), according to the responses received in separate studies by Truett (1984) and Taylor (1986), some 60-70% of courseware publishers claim to conduct formative evaluations and to revise materials on that basis. In addition, about 65% of the publishers in Taylor's survey indicated that they validate objectives with teachers, and about 70% said that they also validate instructional sequences with teachers. [*Caution:* Both studies involved small-sample surveys.]

When locally developed (either by individual teachers or by a district), courseware programs are usually created to "do" something (for example, factor large numbers or solve quadratic equations), rather than to "teach" something. Sometimes, these efforts are "published" as program listings in journals or distributed at-cost to other schools in the district. There is essentially no testing of such software beyond debugging, and scrutiny of the program listings reveals that even debugging is limited.

The Japanese approach to commercial programs is more circumspect and is based on the policy outlined in " A Manual for the Development of Educational Software". Commercial programs are evaluated by a study group of the Japan Association for the Promotion of Educational Technology. This group classifies the types of software into six categories. The first category consists of programs suitable for the

instruction of textbook-type material including tutorial and drill mode. The second category is the software used by teachers in conventional classroom teaching. The third category is the information retrieval type of software used for guidance of learning and consultation on self-learning, and other database type programs. The fourth type is for the creative and problem-solving mode of instruction. The fifth category is used for experimental and practical learning by using computers as a tool for learning. And the sixth type consists of gaming and simulation programs. Programs other than of the tutorial type are being gradually introduced as shown in Fig 3.

Types of Software	Number of schools
word processor	15
spread sheet	4
design and graphics	8
OS and programming language	9
Database	1
Tutorial programs for subjects	19

Fig 3. Schools using general purpose software (as of Dec., 1986)

The British report introduces the notion of formative evaluation during the development phase, although this also has a major role in the process outlined by the Scottish report.

Developing software is a time-consuming business, and accordingly relatively expensive. Because the investment is likely to be high, formative evaluation is considered important, and has to be built into the production schedules. In some instances, such evaluation is assigned several iterations to ensure that the programs are widely applicable. One unit, which is closely involved in research ideas, studies the behaviour of children as they use the software, and rewrites the programs to get closer to their expectations. A special system, called SCAN, has been developed to help with this, details of which are obtainable from ITMA, College of St Mark & St John, Plymouth.

The more common formative evaluation seeks general answers from teachers using the prototype material with children. These include comments on questions seeking technical and presentational faults or ambiguities, but stress is laid particularly on the clarity of the information and graphics being used. Thus, questions like 'do changes in the display happen too quickly or too slowly' and 'how can the display of information be improved' predominate.

All programs are accompanied by a certain amount of supporting information. Sometimes this is brief notes on the operation of the software, but other packages may include extensive worksheets for the children and training materials for the teachers. Again, formative evaluation is undertaken on this material, and questions cover a range of points. On the students' materials, the emphasis is on the links between them and the program, the level of appropriate prerequisite information, the level of language and the clarity of illustrations. For the teacher, while there is some interest in the technical issues of operational instructions, the predominant themes for the questions are the clarity of the objectives and rationale, the background knowledge required, references and hints on appropriate use and ideas for extensions and the need for further flexibility. The results from these surveys are used to rewrite the materials and develop them to a higher standard, and in most cases they only go through one iteration of evaluation before publication.

While it would be true to say that the careful analysis of the results of field trials is undertaken only by the major educational development agencies, whether government-backed or commercial, nevertheless the smaller groups, particularly those sponsored by the LEAs, do undertake testing of their prototypes as well. However, the results of classroom and teacher use are gathered in a less structured manner, talked through at personal meetings and treated less thoroughly. The concern is often about the same areas:

presentation, accuracy and technical competence, ease of use and the quality of the support materials. The context of valuable use of the software is also of concern, and is usually highlighted in accompanying literature when the program is finally distributed. Thus the direction of the formative evaluation of the materials is the same, but the degree of care and accuracy varies.

Few companies or distribution agencies will undertake the duplication and marketing of software without some evidence that formative evaluation has taken place. At the minimum, they wish to be assured that enough has been done to ensure that the presence of technical bugs has been checked and that they are as far as possible eliminated. However, most also wish to know that the software has been used by teachers in practical circumstances and proved at least helpful. There is no figure for a minimum number of teachers in the trial, as this will vary according to the professional stature of the people involved and the subject being covered. If the software is to be marketed as having multi-disciplinary value, then evaluative evidence of that will be sought, but if it is accidentally found useful in subjects for which it was not orignally intended, no effort will be made to prove this.

This is reflected in the Dutch experience but in a less structured way. Whether field-testing is done depends entirely on the design and structure of the project. Experience suggests that field-testing seldom occurs. The teachers involved are usually the only touchstone. Themes and subjects are seldom indicated by government. In the acceptance of project demands however an equal distribution between subject areas is taken into account.

Scotland has perhaps the most formal field testing procedure.Once a program has been completed, and its documentation written, educational field-testing is undertaken in the classrooms both by those teachers involved in its creation in the first instance, and also by teachers who are seeing the software for the first time. In this way, the software and supporting materials can be tested against the stated objectives of the package, and any necessary modifications can be made, before the program is put on general release.

This pre-release evaluation is the joint responsibility of the specifying group and SMDP. The members of the specifying group test the package themselves, possibly with the help of colleagues in their own schools, and results of these trials are discussed by the group before the package is finally declared ready for release. It is, on the other hand, the responsibility of SMDP to arrange for field trials beyond the specifying group's locus, in order to obtain disinterested comment, which is also to be taken into account before the software is released.

A final comment in this sector from Hungary sounds a cautionary note.

It is worth noting that among the makers of the educational programs (the universities, high schools and other schools) there are differences in basic principle and opinions. The main source of these differences lies in the educational background of the two groups as well as in differing interpretation of requirements. One group would rather concentrate on improving research work regardless of whether or not it serves any practical purpose. The other group neglects pedagogical principles and considers only those solutions which may result in a better achievement of the teaching goals. This generates ill feeling between the two groups, leading to a deteriorations in results.

chapter three - production

Since production is a major stage in the development of software, and since there are at least two major forms of production (institutional and commercial) it was deemed important to enquire whether there were discernible evaluative criteria on which production decisions were made and which governed the production processes. One of the first constraints must be cost. The size of a potential market must have a considerable effect on the investment decisions. These in turn will have effects on the amount of time and money allowed to the (educational) developers of the programs. At least one would have thought that this would be the case. The evidence to prove this is hard to find. There doesn't appear to be that much difference between the development time in a small country such as Scotland (which has access to the whole English speaking market) or the Netherlands (which doesn't) and those programs produced in the USA or Canada.

General USA courseware production costs are difficult to determine, since much has been provided as "add-ons" with hardware or with textbooks, and the cost built in to the cost of the basic texts. (*Note:* The "textbook case" has now almost disappeared, as have the electronics divisions of traditional textbook houses.) However, the related factors of "cost" and "time" are the two principal reasons publishers of less-than-excellent courseware give for cutting back on analysis, validation, and testing (Taylor, 1986).

Smaller countries with nascent industries (as far as educational software is concerned) reveal little. Italy faces all the problems of pricing commercial and non-commercial production in a context where pirating is (or is thought to be) rife and the Japanese are at such an early stage of investment in educational software production that no cost-effectiveness study has yet been set up. Thus little is known about how the necessary decisions are taken.

The British have had longer experience of institutional production and to some degree reflect in a different context to those views expressed by the American correspondent. Preparing educational software for the market is an expensive operation, as the package, unlike a book, is multi-facetted, and handling this is labour-intensive. Being certain that sales will result is therefore important. MEP subsidised the commercial sales of its products by using its government funding to pay for development and design. Thus the commercial agences were required to finance only the presentation and duplication, as well as advertising and distribution. Costs were therefore reduced, and the risk capital invested by the agencies was much less than it would have been if they had attempted to cover the design and development as well. A major reason for this approach was to give these publishers and agencies an opportunity to try out the marketing of software without too much capital commitment on their part. That some of them still continue in the field, and are now financing their own programs, is a measure of the success of the operation.

Costs for distribution by telesoftware, whether broadcast or on-line, are difficult to calculate. The same MEP approach was used where these methods of distribution were concerned, and again the emphasis was on seeing how successful such an approach would be. In practice again, there has been steady (or continuing?) use of the systems and the testing by MEP has proved valuable.

In a way, the MEP era in the production of software was artificial . Returns from sales were not high, and partly because schools took some time to acquire sufficient computers to make use of the programs, especially the didactic material, partly because the teachers' approach to the curriculum had not changed sufficiently to fit the materials into the teaching and learning strategies that were being used. However, most agencies involved decided that they would find it difficult to develop appropriate materials and market them at a reasonable price without making substantial losses of money.

The picture becomes more depressing as the complexity of software increases. Costs for development are usually calculated on the basis of the man-hours involved in producing the software after the idea has been conceived. The complexity of the specification increases cost levels by a considerable factor, and current trends are aggravating the picture. In the early days, the simple programs took only a month or two from idea to completion, and even with the small market that then existed, it was possible to recover costs and even to make a small profit. This was especially true for programs for the primary sector. Current trends, however, make it more difficult to make a profit from programs that may take twelve to eighteen months to complete.

For didactic software aimed at the secondary sector (about 6000 schools) companies would expect a sale of between 500 and 1500 copies, and initial production runs would be unlikely to exceed 500 copies. In the primary sector (about 24,000 schools) sales of 2-3000 would be expected, but the intial production run would probably be the same figure. Applications programs would seek a larger market, about double that for didactic software at least, but most companies would be likely to produce in batches of no more than 500 copies. This contrasts sharply with smaller countries which may represent a complete language market, in which sales of only a few hundred (or even less) may exhaust potential sales.

To make the picture more complex and difficult, the variety of machines for which the materials would be produced includes four different models, plus networked versions of each, all within the totals mentioned above. It is little wonder from these figures that government subsidy is thought to be essential if a profitable market is to exist! It is also apparent that applications software is more likely to be profitable than didactic materials, as the spread across schools is greater. (What the British see as a major barrier might be viewed as a situation to be desired in other countries. In 1987 a report from the Dutch Schools Inspectorate noted that in 1700 schools survey there were 83 different types of computer!)

Another way to increase profitability is to ensure that schools buy more than one copy of a piece of software, and that networking is seen as a 'multi-copy' use. Schools are naturally unhappy about this but, with the authority of LEA advisers encouraging obedience to the Copyright Act and the growth of central purchasing through them for the schools in an LEA, there is an increase in the purchase of more than one copy per establishment. In the case of applications programs that reside in ROM, this is inevitable, but those on disc are also being bought in this way too. The market, however, is still not a profitable one in the didactic software sector.

The emergence of these moves towards portability, together with the increasing interest in joint development of software between countries, may lead to costing and profitability figures coming out on the black rather than the red side of accounts. This is particularly important if the relatively small markets for didactic software are to be increased. It is evident that the computer has much to offer for learning, and it would be unfortunate if this was not to be exploited through subject specific materials, and only applications materials were to be used.

This contrasts sharply with the Dutch situation where a sale of one hundred copies might fully exploit a market yet in no way cover costs. Even where a program is funded by the Ministry of Education and distributed to all secondary schools the maximum market is only 2250 copies. There are no standard criteria for the calculation of production costs for software development. Obviously there are differences in approach between commercial publishers and partially or wholly subsidized government institutions because they have different objectives. It is hard to fix the break-even point, on one hand because the production costs for various programs can differ considerably, and on the other because versioning for the different machines brings extra costs. Another problem is that the selling price should not exceed a certain maximum because schools, with their tight budgets, may not buy high priced software.

Where the government project is more integrated, such as the Dutch NIVO project it is possible to calculate the unit cost of delivered software and base production decisions on this criterion. The Ontario project ties the hardware and software provision closely together. The installed base of approved

microcomputers now exceeds 15,000. The average cost of development of a new product is now $ 5.00 to $ 10.00 per installed unit. Development funding is a "once only" payment, and hence unit cost drops as the installed base rises.

In Quebec, where the construction is not so rigidly set up the pricing of locally-produced software is a very complex matter. While the cost of its production must be recovered, the reality of what the marketplace can bear is a bigger factor; therefore, many of the graphic enhancements possible are left out to keep costs within the market range.

In order to make their products more attractive, Quebec software manufacturers are developing licencing agreements with school boards for the use of their software. The government has also entered this area with its limited mixed licence agreement. Here, school boards are able to purchase the software themselves but the government picks up the cost of the licencing agreement. To be able to participate in this programme a software producer must meet a rigid set of guidelines.

In the USA where there are effectively no government projects the commercial picture is clearer. Most single didactic programs sell for about $50 U.S. At this rate, some 3000-5000 copies would have to be sold to cover a developer- publisher's development, production, marketing, and distribution costs. It is perhaps worth commenting that whereas these are high figures and the marketing costs are extensive, the breakpoint can be reached with only a small percentage of the potential market. In many other countries or language areas the fixed costs of some types of program development are greater than the potential returns from the total market. This leads countries towards government supported production, which can lose some of its competitive edge, and to the conversion of existing software which has proved its worth in another context.

Related to the production costs are criteria about free distribution. This is dependent on covering the production costs through subsidy and significantly alters the evaluation of programs considered for production and distribution. The Dutch example of free distribution of software to the schools is the already mentioned Starter-package from the NIVO-project. Schools receive one copy of each program with the explicit permission to use it in a network configuration. In Scotland all software produced by SMDP is distributed free of charge to Regional and Island Authority-nominated distribution centres, which then distribute it to schools on a non-profit basis, in a variety of ways appropriate to each Authority's mechanisms.

Two other factors have to be considered in evaluating software production; copy protection and the language used. Protection of discs is in itself expensive but could make a product commercially viable. The choice of language is becoming more important as the chance of sales to other countries increases and barriers to conversion have to be lowered. The limitations of one language, even when set against development time, may be crucial.

The Italian experience is that the problems of protection are not easy to solve, because it is possible to find programmes that copy the diskette bit by bit without checking. Protection will therefore have to be based on a different approach: one is to aim at a wide distribution of low cost software so that originals with manuals will be preferred to pirate copies which usually come without any support material.
The issue of copyright and program protection in Japan is being discussed in a committee of CEC.

In the early days, British teachers were frequent offenders against copyright by making copies available to their colleagues. Such activity was met by a surge in copy protection being applied to programs, and clever children spent many hours proving that all the techniques could be broken! The major disadvantage, which cost companies money and time, was that this copy-protecting frequently prevented use of the programs on upgraded equipment subsequently purchased by the schools. There has also been an increase in honesty and the acceptance of multiple purchasing, so the trend now is to avoid copy protection systems wherever possible. Several companies still retain the procedure but

they are becoming the minority. As portablity of software between machines is now of considerable importance, the added inhibition of protection seems to be of less concern.

The Dutch approach is similar. Generally, programs aimed at the educational market are not copy-protected. Copying can be done without any great difficulty. However, there is a tendency towards formal registration of the rightful owner (for example the program presents the name of the owner on the screen each time it is started up). This "moral protection" has been borrowed from the American commercial world.

Educational software is currently not produced in Alberta by the Alberta Department of Education. However, a fair amount of educational software is licensed in Alberta, meaning that copies are made available to schools at a nominal cost. The originator of the program usually includes a copy protected system.Ontario products are licenced for the entire Province and hence freely reproducible by school boards for educational purposes.

In the USA where commercial criteria rule, the approach is much harder. "Working copies" may be made of most applications software, but professional and didactic software is usually copy protected - and often drastically so! To offset this, special site-licence and networked versions are available from a few publishers for some of their packages. The number of these is increasing, as school systems' purchasing becomes more organized and the demand for such services increases concomitantly.

COMPUTER LANGUAGES

When it comes to the consideration of computer languages it is encouraging to see more an more countries arriving at the same (albeit temporary) solution. Much of the Canadian material is developed in C or NATAL. In the USA , for sophisticated software, the language of choice is Pascal - and its use is on the increase. However, a BASIC dialect, often supplemented by machine- or assembly-language routines, is still the most commonly used programming language.

Traditionally in the Netherlands, if one can speak of traditions in so young an industry, BASIC has often been used as a programming language for developing educational software. The use of (especially Turbo) Pascal is encouraged by way of subsidized projects and by the influence of the institutes. In some cases C is also used. For the time being Pascal is likely to remain the most popular programming language, not least because of the growing procedure libraries which can be used time and again.

In the English project transfering programs for use on different machines, and in some cases different models, has been a major problem in the early stages of software distribution. Early programs were written in Basic, and the different dialects that were available componded the difficulties. Much time was wasted, and money expended, in rewriting programs so that they could be used on different machines. Varying graphic capablities also raised difficulties, and for programs written in assembler major rewriting was needed. Some were and are written in Logo, and again some translation is required. The position has eased a little with the growing use of Pascal and C, but some rewriting is still necessary. However, the present interest in virtual machines and the dominance of C as the language through which these virtual environments will be accessed, together raise the prospect of much easier transfer between nations. Hopefully, this will also lead to improved exchanges of software between countries, for up to now there has been little importation except in the professional program area, mainly because of the issues surrounding versioning for different machines.

Japan appears to be moving in the same direction for whilst the most widely used computer language is BASIC, C is also widely used among the professionals in program development. LOGO is used in elementary and lower secondary schools, not for programme development, but for education in computer literacy.

BASIC is one of the languages used by the Italians to develop the software; PASCAL is used for scientific programmes and general purpose programmes, due to the wide availability of library routines in this language. For widely distributed programs C is used owing to its portability which makes it possible to run programmes under MS-DOS and UNIX.With the increase in languages connected with artificial intelligence, Prolog has become more widely used especially since compilers in this language became available. In the future, programs are expected to be written in several languages, one for each routine, thus making for easier data management, inference processes, graphic display, etc.

Chapter 4 Conversion of materials

Correspondents were asked to consider the question of conversion of materials from another language. As it turns out few countries have really entered on this task. What has been demonstrated is that programs have to go through three levels of conversion:- language, computer language/system, and cultural conversion. This poses difficulties when operating criteria for selection since one is selecting something which might look quite different by the time that it has gone through the process. What has been learned in the Netherlands is that it is not only possible to build on the evaluations in another country and assess these against ones own national needs but that the conversion of programs can be considerably cheaper than original production. However these conditions are not repeated elsewhere.

The lack of conversion is not surprising in large language areas where competition ensures that there is a sufficiency of programs. It would however seem only sensible in human and economic terms, that smaller countries begin to share their products and buy from the larger groups, just as they already do with audio-visual media and books. Because in many respects the production of software is still a "cottage industry" the *not-invented-here* syndrome still has force. Programs are only felt to have value if they have been developed locally. M.J. van Dalen speaking at the EURIT conference in 1986, on behalf of the educational publishers argued forcefully that there was no way that commercial publishers in a small language area could publish software economically without Government support. The figures look convincing but one must note that publishing is geared to traditional book publishing and marketing and thus each product bears large overheads. Other institutions and some very small companies are however proving that production or conversion of programs can be economic, even in a small market, through a structure with lower overheads, free-lance workers on contract, and a different approach to marketing. At that same conference David Walker of SCET argued that conversion was as expensive as original production. Given that there are many different ways of presentiung accounts, the Dutch experience shows that the use of professional but freelance workers, standardised proceedures, and a planned building of this special expertise can not only considerably reduce costs but that the time taken to produce a product which meets the standards of production evaluation can be shortened.

General criteria do not exist for translating and/or adapting software for the Italian user. It often happens that subject-oriented programs (physics, chemistry) coming from other countries (the USA) have different ways of presenting materials from these needed by our schools and are not suited to the local curricula. Nevertheless, this software provides a good standard of comparison for local production, which has to face the difficulties of a rather restricted Italian market.

Conversion of instructional software from foreign languages into Japanese is regarded as inefficient except in the case of general purpose programs, because of the difference in language systems and cultural background. At the moment, there is no attempt to systematically convert foreign instructional software for Japanese use.

So far only a few programs have been acquired, by the Hungarians, from other countries, but in the future more are planned. This has both theoretical and practical significance. In this non-traditional area much experience can be gained by learning from the principles of teachers in other countries. This approach may be especially important for a small country like Hungary, which could produce the above-mentioned necessary "critical mass" of programs domestically only at the expense of a considerable intellectual and financial effort. Experience shows so far, however, that only a very limited number of programs can be taken over without changes and that this channel may work only in the field of mathematics and other natural sciences. Among the criteria for re-writing, the most important aspect is that programs should accord with the requirements of our teaching plan and with those media familiar in Hungarian schools.

Conversion in Quebec could be a big industry. However it is not. Simply put, the cost of translating existing software from English to French plus the royalty costs involved discourages many local software

development firms. Any software firm that does undertake conversion, usually has to take into account the whole French-speaking world market. With impediments like these, French educational courseware will be very slow in coming.

Very little instructional software is imported into the U.S. To a great extent, this situation exists because of the incompatibility of various brands of microcomputers, because of wide variation in curricula among countries, and because so very much "home- grown" software is already available here. A little instructional software is imported "as is" from other English-speaking nations, such as Canada, Great Britain, and Australia. An even smaller amount is imported and adapted by the U.S. representative of a foreign publishing house.

In the context of this report it is important to note that the British correspondents indicated that the matter had not received consideration to any significant degree.

Chapter 5. Selection of Software.

Correspondents were asked to report on the criteria for the selection of software. It may be significant that those who wrote extensively on this subject did so in descriptive terms of the institutes and instruments and provided examples of the resulting forms. Some of these forms can be found in the appendix to this report. There was no specific consideration of the actual questions (criteria) used. This tends to suggest, since the authors are all familiar with this field, that they share a view that the questions are fairly obvious and that it is the descriptive reporting which is more important.

In considering the selection of software for purchase by schools we are touching on the central issue in this study. When discussion about criteria for the selection of educatioanl software takes place in almost any international forum the model which is quoted is that of the systematised evaluations published in the United States. Some of these are widely known and almost as widely read. Their effective use can be considered as less than the total readership. Whatever the evaluation instrument, and some are to be found in appendix, there is always the question to be asked as how flexible they are to take account of the changing nature of software. It is possible to find review sheets which would mark a program as poor because it did not give cumulative scores. When applied to a graphics program this can be seen to be an inappropriate criterion. Yet there appear to be many who would like to have an all purpose instrument which would give some sort of objective measure of each piece of educational software. Because computer software is made up of digital information recorded on a magnetic medium and is therefore in some strange way "scientific" there seems to be the need for an equaly "scientific" evaluation. We want a universal applicability of the medium to education which we have long since abandoned in our selection of books and have not as yet achieved in regard to audio-visual media. Nevertheless there is obviously value in the evaluations which are carried out.

The comments which follow will indicate not only the wide range of evaluation proceedures and practices but also the frequent gaps between aspiration and reality.

The Italian report shows no formal evaluation in practice with teachers using their own judgement. In Japan the Manual issued by the Subcommittee of the Social Education Council is expected to be used by educators for the rational selection of software, but in fact is not widely read. There is no monitoring system on software-quality and only poor information regarding the quality of productions. Even in the report from england there was little evidence of any formal evaluation for selection.

Dutch schools are free to make their choice from the supply of software available. However, besides the direct influence of the government via the exemplar series called "Voorbeeldreeks" and the list of programs for the "Software coupon", an indirect influence on selection is also exercised by the institutes belonging to the training and advisory structure.

Nevertheless, schools do finally make their own choices. Decisions are often taken by teacher(s) playing the role of pioneers as regards the introduction of computers in the school. It is self-evident that the subject taught by the teacher(s) in question is likely to be favoured above other subjects when deciding on software purchases.

Hungary takes a much more cautious stance stating that the efficiency of computer-based educational activity can be measured only be in the future; all one can do today is talk about the kinds of investigations that may be informative and which we may be able to carry out in the future. Hungary plans to carry out behaviour probing micro- tests, formal interviews, cognitive input - output analysis,

operational input - output analysis, unrealised observations and informal interviews. These are all suitable for probing the effect of the technique on the pupils.

All software purchases by an Albertan school jurisdiction are the responsibility of that school jurisdiction. The Alberta Department of Education provides advice through the Clearinghouse publications. The school is where most software purchase decisions are made, so they may reflect the interests of an individual teacher or of a group of teachers. There are no predetermined selection criteria, unless developed at the school jurisdiction or school level.

Selection of software for the Quebec schools has recently taken place on two levels. Firstly, the government imposed its selection on the schools by providing free French Software for the Comtrem Max Computers. This approach has met with very stiff opposition. The quality and content of the "valise" supplied left much to be desired.

Learning from the errors of the past, the government has now moved from the distribution of totally free software to approving software, after field-testing for a "mixed licence". The school board will pay for the cost of the original software and the government will pay for the user-licence so that the software may be used throughout the school board. The government publishes a magazine called Bip-Bip that highlights technological developments and lists software produced by teachers in Quebec. This magazine is available to all Quebec schools.

At the school board level, most software purchases are made on a limited consultative basis: - computer and subject consultants, teacher and the school principal. When this process of consultation becomes a little more thorough, more people will be aware of what the board/school has in its software bank. Most allocate very little to software acquisition, so it is obviously not a priority item.

Crossing the border into the USA we find the most developed and extensive use of criteria for the selection of educational software.There are many sources of information about instructional software. The major evaluation projects cited below provide their evaluations either through on-line databases (as in the case of MicroSIFT) or by mail for an annual fee (as with EPIE). Several states operate "clearinghouses" that disseminate information about software (including evaluations) to local schools.

General software availability is identified annually in *TESS: The educational software selector* (EPIE Institute, 1986), which also includes useful information about hardware requirements, cost, and preview or return options. A more limited selection of highly recommended software is identified and briefly described in the annual publication *Only the best: The discriminating software guide for preschool-grade 12* (Mattas, 1986). And there is a publication just for parents - *The yellow book: A parent's guide to educationally sound courseware* (NEA Educational Computer Service, 1985). Publishers also distribute their catalogues freely to potential buyers.

Some instructional software is supplied "free" as ancillary materials to accompany basic textbooks. The remainder is marketed competitively by commercial distributors, and may be bought "in bulk" by School Boards. As site licensing and networked versions become more readily available, district Boards of Education are becoming more deeply involved in the acquisition of instructional software and its distribution to local schools.

The School Board usually approves all expenditure on software. Recommendations to the Board come from the Assistant Superintendent (or Director) of Instructional Materials and/or from the Director of Computing (if such exists). When "petty cash" budgets or teachers' "discretionary funds" are used for software purchases (seldom), decisions are made on a school-by-school or teacher-by-teacher basis and may still have to be requested through the School Board. (*Note:* In the "early" days, parent groups often raised money and purchased both hardware and software - a source of much long-range dissatisfaction with computers in the schools.)

Thus whilst there are more published sets of reviews (of a greater or lesser scientific nature) there is still no compulsion on the schools to follow one or other of these guides. The very structure of the school system militates against the development or acceptance of a common evaluation instrument. A number of American evaluation sheets will be availbale in the appendix and the reader should be ready to invest some little work in looking through these to judge which criteria suit their particular needs. It would be tempting in a survey like this to start laying down a series of ' laws' establishing what had to appear as criteria for evaluation but, as has already been shown, there are such significant differences in the school systems and the approaches to the use of the computer that it would be unwise to assume that forms of evaluation can be transfered from one country to another without modification. Equally one has to be aware that many of the American evaluation forms were evolved to compare didactic programs of the drill and practice or tutorial sort.

England has fewer regularly published reviews but the provision of guidance to teachers on what to use in schools was a major feature of the MEP project.Cost of production, and thus price to schools, is one reason why the amount of software in use has not been as great as had been hoped. Other factors, like the density of machines and the level of training of teachers, have also played significant roles, but it is only in the last year that it can be assumed that about 10% of teachers are using the computer in their classes at least once a month. This is still not a high level of use, but then this is an innovation which has the potential of radically altering the curriculum and the teachers' role in presenting it. Such innovations take time to be accepted and then absorbed, and even among those who are active users, the subjective evidence is that most are twisting the presentation of the new materials to fit the teachers' old styles.

Another well-known excuse for under-use is that the teachers did not know that the software existed. Many methods have been used to keep them informed. Newsletters with lists have been circulated to schools from LEAs and the old MEP regional centres. The commercial companies have placed advertisements, sent circulars to schools, visited teachers' centres and demonstrated the programs and offered 'on approval' purchases. MEP took space in the main educational newspaper to list and describe items of interest, and similar articles were used in specialist magazines and publications produced by subject associations. So, from the point of view of publicity, much has been published. However, actually handling the software is important, and so travelling exhibitions have been arranged at which teachers could use the programs. Teacher training courses have had software available for trying and examining, and there have been sample extracts distributed to give teachers an opportunity to get the flavour of the material. All of these help to develop an understanding of the value and relevance of the programs on offer.

With the help of the broadcasting authorities, a number of radio and television programs have also carried descriptions and demonstrations of software. Some of this has been the material that these authorities have sponsored to accompany their school broadcasts, but all have contributed to generating an awareness and interest in the use of software in education, which in turn encourages the acquisition of items that have been publicised through these media.

It is noteworthy that the emphasis in the English and the American reports is on the review nature of the evaluations. These reviews can overcome the shortcomings of a rigidly applied set of criteria. There is a definite role for the reviewer to express that personal feeling about the suitability of a program which matches the individuality of approach which is demonstrated by each individual user of the program. But by the same token there are drawbacks to relying too much on a review written by someone with whom the reader has no agreement about the meaning and value of the terms being used. Much American courseware is selected solely on the basis of teacher preference which, in turn, is usually determined by a combination of catalogue descriptions and reviews in teacher magazines such as *Electronic Learning*, *Arithmetic Teacher*, or *Science Teacher*. This is unfortunate, because catalogue descriptions are often misleading, especially with regard to the mode of didactic software (for example, a "tutorial" may simply be "practice with feedback" and incorporate no instruction at all!) and the management features

included (for example, "student records" may hold only the result of the most recent program run for one student!). Rarely, a preview copy of the instructional package is obtained and "screened" before the decision to purchase is made.

Because of the wide-spread dissatisfaction with courseware quality, several major courseware evaluation projects have developed in the U.S. Chief among these are the MicroSIFT Project, which is a non-profit, government-sponsored undertaking that operates out of the Northwest Regional Educational Laboratory at the University of Oregon, and Educational Products Information Exchange (EPIE), which is a private, self-supporting endeavour that grew out of an earlier commitment to evaluate "programmed instruction" materials. Also, several evaluation centres are operated by individual state Departments of Education and district School Boards (for example, the States of Minnesota, North Carolina, and Texas, and the School Boards of Baltimore in Maryland, Iowa City in Iowa, and San Mateo County in California). Most such systems have created formalized criteria and formats to which all participating evaluators are supposed to adhere. The Alberta Clearinghouse has developed evaluative criteria, which are now used by all provinces and territories in Canada. But here too there are no guarantees that the reviewers adhere to them.

Under the Scottish project the fact that there was a central production unit, which had adopted certain standards for development, production and presentation, tended to produce work which was within a certain stylistic band. It was therefore somewhat easier for SMDP to construct evaluation procedures for teachers. At least they had the knowledge that they applied to the SMDP programs. They also had and still have a strongly founded belief that these criteria were more universally applicable.

Guidance on the issues to be addressed during formative evaluation, is given in the SMDP publication "Validation of Educational Software". This document teases out the issues surrounding the place of the package in the curriculum; its aims and objectives; the value of its planned content, the types of learning embodied; the reasons for using a computer rather than other media; implications for off-computer activities; interactive features; appropriate use of facilities offered by the computer with special reference to the locus of control of learning; and a series of subservient issues involving layout and presentation, ergonomics, robustness, documentation, and testing and trials. The document includes a suggested summative evaluation form which represents a necessary compromise between what information could be obtained in the light of a full evaluation of a package, and what it is reasonable to expect from busy teachers who have other things to do apart from evaluating software. A copy of this form is included here in Appendix.

The Scottish correspondent goes on to reason that these completed forms ought to help others in the selection of items for particular purposes, but there are two factors which militate against this. Firstly there is as yet no formal mechanism for distributing evaluation information (and for what to do if two evaluations of the same package are totally contradictory). Secondly, no evaluation of educational material is portable: that is, there are in the final analysis no good packages and no bad packages, but there are good and bad ways of using any package. This implies that a package is educationally neutral until applied in the learning situation. All an evaluation report can hope to do is to indicate versatility of application, or in the absence of this versatility, to show the effectiveness of the package in one (possibly very restricted) set of circumstances.

As the software is in general and extended use, comments will come in from a variety of sources. This post-release evaluation is likely to be very different from the pre-release evaluation. It tends to center on how the package can be used, based on successful experience, rather than simply how the package fulfils its original aims. Arising out of this may come innovative applications of a package which had not originally been envisaged by its authors, as well as implicit or explicit suggestions for extension and improvement.

It is recognised that the availability of a computer to support the learning process provides the teacher and learner with a range of possibilities, in terms both of methodology and content, which have previously been unavailable. One such example is the use of data bases, where it is now possible (and indeed becoming common) for children to be involved in constructing hypotheses, and testing these by use of an interactive data base containing relevant information. It is therefore frequently desirable for training materials to be available for teachers, over and above the documentation provided by the package authors, to help introduce unfamiliar aspects of the package in terms of its methodology or content.

These are crucial comments in the context of this survey. Despite the fact that SMDP is operating in very favourable circumstances and with a far greater degree of centralisation than is to be found in most countries their system is still set up to recognise that evaluators (who are really reviewers) are going to provide a response from a particular context and that this response will be to a large measure subjective. Even though the SMDP document provides fairly strict guidelines for the judgement of programs which, if they are from SMDP, already conform to these guidelines, there is a recognition that the resulting reports will be rooted in local conditions. As such they can be illuminating and provide valuable guidance. In Scotland the reader has a good knowledge of the general circumstances and prevailing views. Because of the size of the educational community the reader may also have a good knowledge of the specific context from which the reviewer is writing. How much more difficult it is to obtain the same sort of enlightenment in a market the size of the USA.

Dutch software is usually obtained from the institutions publishing the materials, i.e. the commercial publishers, or the the educational institutes that produce software. There are a few examples of shops specializing in learning materials and offering educational software, but they are not widespread.

Information about software is distributed in many different ways. Besides magazines and advertising there are also two magazines especially about computers in education, "Computers op school" and "Scenschrift", which are government subsidized. The advisory units regularly organize seminars and courses. Finally, databases exist in several locations containing information about educational software. These databases are sometimes accessible through Viewdata systems.

Evaluation criteria applied in the selection process may differ widely. Until recently, there was no question of a standard series of questions used for evaluation purposes. Apart from the many questionnaires (often subject-oriented), composed by the teachers themselves, no uniform criteria existed.

Under the project name "Voorwaarden voor coursewaregebruik" (conditions for the use of course ware), an instrument was developed between 1983 and 1985 to serve as an aid for the evaluation of educational software. This instrument, later known as "Courseware nader bekeken" (A closer look at course ware) was developed by order of the SVO (a large educational research institute) in the Department of Education of the University of Utrecht and completed in 1985. The instrument (covering only general educational and no subject specific criteria) was meant originally to serve in the INKON-project (part of the National Plan INSP), which was later on rechristened SCEN (Software and Courseware Evaluation Centre Netherlands). The specific task of this project was to evaluate educational software and course-ware in order to provide the necessary information about these new learning materials for the schools.

The instrument developed for the Dutch evaluations will be looked at later together with some comments from the USA. But before that it is useful to consider the pragmatic approach adopted by the MEP project which attempted to recognise the subjective elements in the selection process:

Knowing of the existence of software does not mean that it is purchased, even if the money is available. In the early days, almost any program found a market, but teachers and advisers have become more

selective over time. However, the mechanisms which they use to choose are still more subjective than objective, and much dependent on how the individual sees the program fitting into the teaching scheme rather than on its inherent merit. Teachers are very conscious of reviews, and there are many of these in educational publications now. *Following the views of a trusted reviewer is a well known way of selecting purchases. Hunch and whim are still significant levers in deciding to purchase, but more valuable is the practical experience and sampling of the material, which for books used to be known as the inspection copy service that publishers offered. Word of mouth is also a key factor in deciding to acquire a piece of software, the views of trusted colleagues being considered valuable guides to quality and use.* This is particularly useful when linked to a description of how the program was successfully used, the way in which the children were prepared and also what happened afterwards.

At the LEA level, there are no authorities with a fixed set of criteria for selection. All will wish to be satisfied that particular items are free of operational bugs, and that they meet a need that the adviser has identified or noted being expressed by one of his or her professional colleagues. This means that software houses that have a tradition of providing good materials are supported, whereas those with some failed items are not so uncritically used. Several advisers use panels of teachers in different subjects to vet the software on their behalf, and purchases for all schools are based on their views. The option sought is very much a subjective one concerning relevance, accuracy, presentational quality and good communication.

Thus, on the whole, there are no widely used criteria for judging software according to objective qualifications. In the end, the teacher is concerned to see whether the materials support learning for the children, and this means that he will be watching their reactions carefully, as well as the work they do before, during and after, and will note whether or not they have gained from the experience, again a subjective judgement. If the program's graphics and presentation are motivating, and the support materials develop particular skills and insights, these may be important values for the teacher, but he is likely to buy a program in the absence of these because the program illustrates a difficult topic. Only technical and accuracy criteria are likely totally to eliminate from further considering a program. Another subjective key will be whether the teacher feels that he is covering the same topic as well using other means like a film or even blackboard and talk. As the teacher views the computer as a supporting resource, he will not use it purely for its own sake.

Applications programs respond to the same questions. However, the LEA view of what will be used becomes more important with these programs. Where judgement is made, the criteria are mainly user friendliness, the range of facilities and interconnections with other software, and general presentation. With schools interested in sharing materials, particularly worksheets and similar materials, the use of the same applications programs is a clear advantage.

These programs are now so well tested and trialled that increasingly the judgement is on comparing facilities. Certainly accuracy is not an important factor, but it is essential in deciding about the didactic materials. Early drill and practice programs suffered a great deal from criticisms that they were providing or testing wrong information. The 'teaching how to tell the time' programs were infamous in this respect. If the hour hand was shown against the two when the time was 2.30, it was wrong. However, if it was shown accurately half-way between the two and three, young children found great difficulty in interpreting it. Thus, some programs were issued in two versions, one of each type!

The demise of drill and practice programs as serious contributors to education has seen the end of this as an issue, but accuracy remains an important feature for selection other forms of didactic programs. Any teaching material may fall into traps associated with the compromises inherent in simplification. But because of the power of the concept of computing, more is expected of the programs than would be of books on the same subjects, and in particular expectations of accuracy are raised. Some tutorial programs are criticised for this, but these are not as important educationally as those simulations, models, adventures and investigative programs that stimulate exploration. Here, the problem of simplication

leading to inaccuracy often leads to criticism of the programs and their elimination from consideration.

Developers are aware of these difficulties and try to deal with them. Reviews and comments from teachers are taken seriously by software designers as guides towards future practice. When one developer uses a new presentation method that finds favour with teacher and child, most of the others

are quick to examine whether a similar approach would be suitable for their material as well. Thus, the feedback from approved techniques into styles of development is quite rapid and helpful in improving standards.

If England has arrived at a rather flexible solution and the American at something which suits the publishers and gives them the apparent respectability of producing objective evaluations (when it is mostly the reviews which influence peoples decisions, then an examination of what the Dutch are producing as an evaluation instrument will serve as an example of a broad "middle of the road" approach. This not selected because it is any better or worse than any of the others, but simply that the selcted parts illustrate some of the issues which have been discussed here.

Let us first consider the instrument itself. In addition to the description of aims and target groups, the following categories of items by which the evaluation should take place can be distinguished:

A. Aspects of use and lay-out ("gebruiks en vormgevingsaspecten")

B. Educational effects ("Onderwijsgevende werking")

C. Goal effectiveness ("Doelrealisatie")

Copies of sample pages taken from these categories are shown in the following figures:

3.3.2. Screen Presentation

35	The form of presentation is functional	disagree	1	2	3	4	agree	
36	Information is presented in complete units	no	1			4		ˇ n/a
37	The introduction of new information is clearly signalled	bad	1	2	3	4	good	
38	Shifts from one screen to another takes place "per page"	no	1			4	yes	
39	The screens are quickly enough composed	disagree	1	2	3	4	agree	
40	The program never gives the impression of having stopped	no	1			4	yes	
41	Text on the screen is easy to read	disagree	1	2	3	4	agree	ˇ n/a
42	Graphic presentations are clear	bad	1	2	3	4	good	ˇ n/a
43	Animations are of good quality	bad	1	2	3	4	good	ˇ n/a
44	The lay-out of the screen is well organised	bad	1	2	3	4	good	

Fig.4. Sample pages from Evaluation Form on Educational Effects

4.2 Presentation of Lesson content

DESCRIPTIONS

11.	What structure does the package use?	ˇ strictly linear			ˇ linear with menus			
		ˇ branching structure			ˇ networking structure			
12.	How is the material ordered?	ˇ ascending level		ˇ constant level			ˇ descending level	
13	In determining the size of steps, is any account taken of the differences of level within the target group?	not at all	1	2	3	4	very much	
14	To what extent is text used?	not at all	1	2	3	4	very much	
15	To what extent are drawings used?	not at all	1	2	3	4	very much	
16	To what extent is animation used?	not at all	1	2	3	4	very much	

EVALUATIONS

17	The sequence of sections in the package is reasonably logical and well-structured	disagree	1	2	3	4	agree
18	The material is well constructed	disagree	1	2	3	4	agree
19	The lay-out and presentation are good	disagree	1	2	3	4	agree

Fig. 5. Sample page from Evaluation form on Goal Effectiveness

Any other points or special characteristics which have not been coverd in the questionnaire?

2.3 Evaluation of Goal-effectiveness

In view of the evaluation of the package as a learning aid and medium of tuition, to what extent is it likely to achieve its stated objectives?

not likely 1 2 3 4 5 6 very likely

Qualifications on this evaluation:

The forms shown above are the scoring-forms. These pages are accompanied by extensive documentation with extra information on each of the questions.

INKON worked in a variety of evaluation situations with this instrument or with a subset of it. It emerged that the criteria often failed to cover all aspects of the software to be evaluated. In particular new developments such as open-ended or content-free software could not adequately be evaluated by the use of this instrument alone. SCEN (the former INKON) is currently engaged in constructing a complete and "objective" evaluation process, consisting of field-testing (two teachers test the software independently of each other); a subject-oriented test (testing whether the software is correct and valuable as far as the subject is concerned); and a general educational test (based on the "A closer look at courseware" instrument). Apparently distinctions were drawn according to user-group and not according to type of software.

Sample pages from the three questionnaires:

Fig. 6. Sample page from Teachers questionnaire

6.0 Learning results

6.1	This question refers to the fact that the program may have other effects on the behavious of pupils in addition to those intended; e.g. pupils may learn from a program to solve sums according to one particular method, and afterwards apply this strategy to types of sums to which it does not apply. We need to be alert to spot any results, either good or bad, occuring with some consistency. If this is the case here, please describe these effects briefly and concisely.	Did you observe any other learning results or side-effects, good or bad, connected with the use of the program? If so, what were they?
6.2	A program may sometimes have more uses in education than the designer foresees. If you think this is the case with this program, please give a short, concise description.	Are there potential uses other than those suggested in the package description?
6.3	Not all individual pupils possess the same previous knowledge, skills or learning style. It is therefore possible that not all pupils will learn as much or as quickly from the program. For instance, pupils keen to solve problems on their own may gain more from a program caaling for this skill than will pupils who are unused to solving problems for thenselves. Indicate here what pupil characteristics contribute to their gaining to a greater or lesse extent from this program.	Are there certain types of pupils who will gain more from the program than others?
6.4	It is equally possible that the program will not be equally suitable for use in every teaching situation. For instance, a program that performs excellently as classical demonstration material will be less suitable for use in a situation where no classical instruction is given. Indicate the teaching situations for which you consider this program especially suitable, and those for which you consider it not suitable.	Are there certain types of teaching situation in which the program proved to be particularly ineffective in use?
6.5	Indicate briefly the general reactions of the pupils to the program. Did they find it, e.g. easy, difficult, boring or interesting?	How would you describe the reaction of the pupils to the program?

Fig. 7. Sample page from Subject Specialist's questionnaire

16.0 Subject Content Evaluation

16.1 Is the subject content of the teaching material correct?

16.2 How well does the material fit in with the usual teaching material for the relevant target group?

16.3 Do the contents of the package match the entry level of the assumed target group? (Is the target group correctly chosed?)

16.4 Does the package match any particular method used to teach this subject?

* no * yes

If you answer is yes, please expand on it:

Fig. 8. Sample page from General Educationalist's questionnaire

The Evaluation of use and lay-out aspects

4.0 Use of package by teachers

4.1 Is it possible to adjust the progam as regards programming level?

* no * yew

If your answer is yes, how?

4.2 Is it possible to adapt the program in advance (time, duration, level, optional choice sections, etc.)?

* no * yes

4.3 Does the program offer opportunities for differentiation?

4.4 Do you expect the program to require much preparation, supervision or explanation from the teacher?

* no * yes

If your answer is yes, on what basis?

The information that results from this process is written up in the form of a summary describing with all the strengths and weaknesses points of the particular software package. This, illustrated if possible with screendumps of the program, is then published in a magazine, called "SCENschrift", to which schools can subscribe, and which carries in each issue a series of these reviews.

There is no evidence as to whether the use of criteria results in better software. No research has yet been done on this. Of course, not all educational software is subject to review by this organization anyway; publishers are free to cooperate or not. Neither has there been any research as to the possible effect of the reviews on commercial advertising.

Finally to round off this section with some observations from the American report, there is a suggestion that the same doubts exist there and that the reality of the selection of software has more to do with subjective and factors which arise from that complex combination of program sort, target group application and context. So far, very little evaluation of professional and administrative software for schools has been formalized. Most U.S. evaluation efforts have focussed on didactic and applications software. There is general agreement that any formal courseware evaluation must include testing the program with representative students (Fisher, 1983; Taylor, 1985). Apart from that, a wide range of particular criteria are usually grouped into three categories: content characteristics (for example, accuracy); instructional design (for example, effective feedback); and technical factors (for example, comprehensive user support materials). These criteria are critical, because a publisher's claims may not be supported by sensitive testing. A single, summative conclusion is also stated (for example, recommend highly, recommend with reservations, or do not recommend).

Most of such formal evaluations are made with reference to a very specific form or checklist. Sometimes, slight variations in the format are developed for certain modes or content, as MicroSIFT has done for science courseware and EPIE has done for programming languages. In spite of well-structured evaluation procedures, the conclusions reached by individual evaluators or by different evaluation projects may differ - sometimes to the extent that one evaluation will highly recommend a program, while another concludes that it should not be used (Taylor, 1986). [Sample evaluation forms are provided in the Appendix.] Evaluations that are prepared for publication in a magazine are often more descriptive than critical, and may not provide answers to detailed questions.

As far as system-sponsored evaluations are concerned, some school districts or counties (for example, San Mateo County Office of Education in California) have developed formal evaluation criteria and have actually used these. However, this is not the typical case, for "intuitive" evaluations based on sales promotions still predominate.

chapter six : use in schools

The Italian report, whilst noting that there is no organisation with a central role in evaluationg the use of software in schools, makes the cogent point that we must not forget when evaluating the product, even within the perspective of an educational application, that this is quite different from evaluating "the process" triggered by inserting a give programme into the curriculum.

The collection of the views of teachers, which has proved so useful under the MEP project, help in the identification of the software that they will use, but the drive behind the acquisition is its use with children. Because of the very private way in which teachers work, there are again no fixed or agreed views on how to measure its value. In the early days, particularly when drill and practice programs predominated, measured improvement in performance was a significant criterion, but this was rapidly seen as irrelevant. Instead, the general consensus seems to be to judge software on whether the children are undergoing a useful experience that my assist them in developing useful skills. The fact that they 'talk maths' after using Logo is considered beneficial, and the exploration that they undertake in developing the pictures via turtle graphics helps to expand their problem-solving skills.

This means that the only evaluation technique that counts is illuminative. Several case studies based on this approach in periodicals have been published, and MEP produced a short series known as 'Reports from the Classroom' in which teacher, observer and children commented on their experience. These reports did not evaluate the software itself, but discussed ways and means of making appropriate and educational valuable use of it. They highlighted the skills and openings for learning that the children used, and left it to the reader to judge whether this was applicable elsewhere, or even valuable as an activity.

Once again the illuminative character of evaluation is highlit by the English report. The Italian comments cited above seem to point in the same direction. If actual classroom practice is the eventual test of the value of a piece of software then one has to ask how far the results of one classroom can be transported to another. The Italians appear to want a more scientific instrument, rather than relying on the accumulation of illuminative reports. The Dutch correspondent notes that although SCEN provides teachers with an evaluation form to use with the software to be evaluated, no special items are included regarding the time the program is used in the classroom situation, neither are there any questions which ask for the number of pupils working with the program, or for any information about the frequency of the program's use in the classroom. In brief, there is in the Netherlands no structural evaluation of:

- pupil improvement
- improvement of learning overall
- improvement of classroom practice

in relation to the use of educational software. This leads that correspondent to conclude that evaluation activities in the Netherlands may, until now, be characterized as being summative product evaluation.

Whilst emphasizing the flexibility of the evaluative criteria developed by the Alberta Clearinghouse, which are now used by all provinces and territories in Canada, the correspondent notes the lack of effective evaluation in schools. A single instrument is available and used by the Clearinghouse. This is modified for different software. These criteria are used by the Clearinghouse to reject or approve software. In practice, over 90% of software screened is rejected. Many commercial developers are very interested in the Clearinghouse results and some changes have been made to software products. Schools consult the Clearinghouse published reports and use the Clearinghouse published reports and use the Clearinghouse evaluation instrument. In our experience commercial advertising is not very reliable. In general, the Clearinghouse has encouraged the trained field evaluators to use the software with students under realistic conditions. This is rarely possible, therefore the Clearinghouse reports are

based on teacher perceptions and not necessarily student experience. No criteria for classroom evaluation have been developed.

Once in the Quebec schools, software is left to the teacher to use as he/she sees fit. It is therefore rare to come across any formal evaluation of its educational value beyond verbal reactions expressed by teachers or students.

Because of the decentralised nature of the United States situation it is necessary to look at the background of the use of computers in schools, in order to make sense of what little evaluation has been done. After 15 years, there are still very few microcomputers in U.S. schools - in fact, the overall median ratio of students to computers in Spring 1985 was 42:1 (Becker, 1986a; p.2). Of the six possible school configurations in the country, the "best" student- computer ratio - 25:1 - was found in junior/senior high schools (grades 7-12), and the "worst" - 60:1 - was found in K-6 elementary schools. And the computers are not used as much as they might be during the 30-35 hours per week most public schools are open. The median use of microcomputers in elementary schools (kindergarten through grade 6) is 15 hours per week, in junior high schools (grades 7-9) is 20 hours per week, and in senior high schools (grades 10-12) is 23 hours per week (Becker, 1986a; p.6).

Not surprisingly, then, students still receive very little exposure to computers. Becker's survey (1986a; p.5) established a median of 0.6 hours of computer use per student per week in elementary schools, 1.1 hours per student per week in junior high schools, and 1.8 hours per student per week in senior high schools.

It is also important to note that despite all that is written about the educational software and is place within the system, most of the courseware now in schools consists of professional programs - spreadsheets, word-processing packages, database managers - that are commercially produced. The most popular professional packages are *Lotus 1-2-3*, *WordPerfect*, and *dBase III* (all in IBM-compatible versions) and *AppleWorks* (for the Apple II series). Any available applications programs are usually limited-function spreadsheet, database, and programming-language packages that accompany "computer literacy" or programming textbooks, or introductory-level word-processing software.

As far as didactic software is concerned, drill programs (e.g., math "facts" or capital cities) are by far the most readily available in schools. In fact, over 30% of all school microcomputer time is devoted to drill (Becker, 1986a; p.5). Drill packages were the earliest ones produced and, therefore, the first purchased. (The initial flood of insipid software in drill mode is probably responsible for wide-spread teacher dissatisfaction with computer-assisted instruction.) Fortunately, there is evidence of increasing emphasis on simulation and problem-solving/logic software, both in purchasing and in development.

In the early grades (kindergarten through grade 3), instructional use of computers is most prevalent in mathematics (about 42%) and English (about 40%). In the middle grades (grades 4-8), "computer subjects" (computer literacy, programming, problem-solving) claim the greatest proportion of computer time (about 34%, compared with 28% for mathematics and 25% for English). Finally, in high school (grades 9-12), business subjects account for much of the computer use (about 18%), although computer subjects still draw the greatest time (about 50%; Becker, 1986b; p.6).

Quite a lot of administrative and management software is in use by School Board administrators (for inventory control, purchasing, payroll, and accounts payable). However, very few packages have penetrated individual schools or classrooms. Major exceptions include attendance reporting packages, sports statistical programs, and guidance counselling software. Rare is the classroom teacher who uses software to manage his/her instructional programs and record-keeping.

Some in-class evaluation is done in conjunction with post-production evaluations, using formalized criteria. More and more courseware publishers are encouraging teachers to take advantage of a 30-day preview period. This gives teachers (or district evaluators) an opportunity to evaluate the software's

usefulness to them and to their students before purchasing it. Other publishers allow schools to return software they don't like for a full refund (within a similar time frame). These two policies are enabling more in-class evaluation of didactic programs than formerly. Because so many teachers are unhappy with earlier purchases, the trend toward pre-purchase, in-class evaluation will probably increase.

Little data is available about the instructional effectiveness of commercially published software, because most publishers neither conduct effectiveness studies nor publish the results (Truett, 1984). However, scholars are now beginning to conduct their own effectiveness research using commercial software. For example, Perkins and Bass (1984) tested the effectiveness - for improving four thinking skills - of seven commercial problem-solving packages over a nine-week period. They found measurable gains in favor of the software (compared with traditional instruction) for verbal analogies and inductive/deductive reasoning, but not for logical reasoning or word problem analysis.

Becker's periodic surveys, *Instructional Uses of School Computers* (1986a; 1986b; 1986c), also attempt to investigate effectiveness issues, for example, the impact of computer use on student motivation, confidence, and self-discipline (1986b), and various organizational effects within schools such as teachers working together (1986c).

Thus one can see that even in a land rich in software and richer than most in academic study there is a dearth of systematic evaluation in schools. The work of Becker and others cannot be overestimated. Few people have done serious evaluation of the effectiveness of programs in schools.

Within the British projects there has been very little scientific evaluation but a great deal of reported experience from schools. The English correspondent argues the case well for the value of such illustrative reports

Selecting appropriate skills to mention in reports like these has been relatively easy, as the need for such reports has featured in a number of papers and publications from Her Majesty's Inspectors and curriculum bodies. These have become notable parts of the preparation of the curriculum that teachers use. It is interesting to hear discussions with major developers identifying the integration of such skills in the materials they are making. The more experienced developers are also eager readers of case studies on the use of their software, and the ideas and problems that are described are taken into account in future designs. Unfortunately, the less experienced are usually convinced that their approach is the right one anyway.

Because this approach dominates the current attitude to evaluation, there is no published evidence of pupil improvement in a statistical sense. Improvement in general is a very subjective concept, which is probably best measured through the illuminative techniques that are being used. The arrival of software has opened up a new range of learning opportunities, and these must alter the content of learning and the approach that children take to it. Comparing performance after software with that produced after other techniques have been used is to achieve a non-sequitur. The use of software should lead to different results, the level of benefit of which can be measured by the teachers' view of that is a good educational outcome. Thus the problem faced by educationalists today is to reassess what is important in education, appreciating what the tools now available offer different avenues to understanding and to the expansion of abilities from their predecessors. If the old approaches and the old techniques were successful, the new technologies are not important, for they are not just reinforcing agents. Instead they lead to new pathways which need to be assessed and evaluated against new criteria. In Britain, such criteria have not been determined, and until they are, judgement rests with analyses of what teachers and observers see. At the moment, the new experiences of children appear to be beneficial, as viewed through the eyes of the professional educator.

Summary

1. The Contexts

Though evaluations, one might even say the evaluation industry, to be found in the USA are frequently cited in reports on the use of microcomputers, a number of specific aspects are clear: many people (outside the USA) seem to feel that they should also have institutions such as EPIE; it also seems that very few people outside the target area of the USA do anything about these American reviews, in that they buy and convert very little foreign material; and the contexts in which these distant viewers stand is often radically different from that of the USA.

America with its system of devolved school control, operates with a large body of commercial production originating outside the education structure. Thus the market place imposes a demand on education to produce a "best buy" list. Canada, also a federation, presents a wide range of provincial organisation. Though there is a national software review system based in Alberta not all provinces base their software selection so strongly on reviews of commercial products. Ontario chooses projects for development based on submitted proposals and then ensures widespread distribution of the finished products.

When one considers Europe, other patterns emerge. The UK development projects have ensured a high degree of hardware penetration into the schools. The software development is not so consistent. In Scotland it is largely centralised whereas in the rest of the UK the products of the projects are sold off to commercial publishers. Another factor which must be taken into account when considering the UK schools is that not only is there a largely unified school system but also that these schools ahve been working in a context of resource-based, project-based, and even inter-disciplinary education for the past fifteen years at least. This has undoubted advantages for the integration of computers into education. The Dutch, Italian and Hungarian examples run the gamut of a highly complex system of schools in the Netherlands with an inbuilt freedom of choice, through the Italian situation where the school system is not complex but there is apparently a lack of central guidance for the provision of hardware or software, to the Hungarian state system where the decision making is perhaps the most centralised. Japan despite it thrusting image in the world of computers is currently taking the most cautious stance of all the countries reported on here. The penetration of information technology into the Japanese schools is not very far advanced. The traditional nature of many of the attitudes in Japanese education provides a resistance to innovation which is also to be found in many other countries.

What can be drawn from the descriptions is that it is not easy to separate software and the criteria for its evaluation from the context and history from which it has grown.

2. Development of Software

Wherever software is being developed it is with an eye on the market. Apart from a small amount of development for personal satisfaction or to solve a very local problem, software developers want others to use their programs. That market may range from a small group of teachers to whom that program is freely given and who pay with their gratitude and accolades, to a potential pool of many millions of people - e.g. the major language groupings of the world. What can be deduced from these reports and others which were previously available is that criteria for the evaluation of programs during design and development are most likely to be found where educational institutions play a central role. The Scottish SMDP and the French software development linked to "L'Informatique pous Tous" bot show clearly the way in which proposals from teachers can be formulated within guidelines and against sets of standards. This ensures at least some of the aspects of the quality of the program before it is finally produced. Toronto, for those programs which fall within the provincial program, has equally stringent criteria.

Because the criteria and procedures for these projects have been widely published we have become

sensitised to the fact that we frequently know very little of how other programs are produced or evaluated. One has reason to believe that much of it, even the very best programs, is written by the sort of educationalist who previously wrote those excellent text books - out of personal and professional experience but without much field-testing and rarely against lists of educational criteria.

3 Production

The same sort of range of experience can be seen in the production stages of a program, from countries which are mainly commercially based to those in which central government or large institutions have a major role to play. In both cases economic factors have a higher priority than educational criteria. One has to assume that these educational criteria had full operation in the early stages of planning and development. In one sense no sensible producer would bother bringing a program onto the market if they didn't believe that it had a chance. This usually means that they have to be able to demonstrate the educational worth of a program as part of the marketing. Where national projects such as those to be found in the Netherlands or Hungary arise, decisions can be taken to make large scale provision of particular programs, which negates the normal market criteria. One has also to assume that these programs have been selected for sound educational reasons.

When producing for an open market the normal criteria for deciding on any product must operate. Is the market big enough? Will somebody (and not necessarilly the customer) pay for the product? and there enough people willing to pay enough money to make the production viable. The payer may be the state or an individual. Large language markets such as the English speaking world provide a good test of the quality of a product. For many other countries such criteria cannot be applied. If the maximum potential market in a particular language area is insufficient to cover the development and production costs then either those costs have to be lowered, the selling price increased, or a cost- covering subsidy has to be found from somewhere.

One of the results of these pressures - in an attempt to cut costs - has been the selection of cheaper systems which later fail to match up to the needs of the users, the rejection of many of the facilities offered by the computer because of the costs of more sophisticated programming, and the marketting of programs which may create negative reactions to the computer over a period. One can cite here the underuse bordering on rejection of the computer which has been reported from the USA and attributed by some writers to the plethora of dull boring (but cheap!) drill-and-practice programs. Short term economies in order to turn a profit may lead to longer term damage to the market. In this sort of situation the publishers, with their large overhead costs may be reluctant to enter the market and the future may lie with the small unit which can manage to produce high quality software.

A further factor affecting production is the illegal copying of programs which seems to common practice the world over. Where programs are freely available under a government scheme there is no incentive to make copies. Where higher costs are being charged piracy can destroy a market in days. Some countries are aligning software to published courses or textbooks in order to ensure some returns and to reduce the effect of the black market. Copy protection of programs has been rejected by most countries. Such protection is a challenge to every juvenile hacker and good protection is very expensive. Some American companies are using quite drastic protection on some of their products. One method may prove more succesful - and cheaper. That is the moral protection of having the purchasers name built into the title page of a program. Any copies will show who has permitted the illegal action.

The international flow of programs across political and language boundaries could be a stimulous to (better) production. Thus the gradual move towards Pascal and C as languages for development of educational programs bodes well for the desired transportability. Meanwhile Basic, with all its different forms and dialects, continues to show that a common language is anything but basic.

4 Conversion

Of the all the contributing countries only the Netherlands has made any consistant effort in converting software originally produced in other countries. It is therefore not justifiable to attempt to draw conclusions about criteria in this field. Whilst not underestimating the problems of converting programs from one computer system to another, from one language to another and from one cultural context to another, the Dutch have shown that there can be considerable benefits and cost savings if programs are selected carefully. The process is similar to the normal selection of software with the added dimensions of being able to determine if the way in which a program has been constructed makes is relatively easy to convert to another system. NIAM has discovered the value of having a few people who can build up their experience with specific systems. A further dimension is the ability to recognise which parts of a given program will need to be rewritten or replaced entirely with new examples in order to make the program acceptable culturally to the intended users. Underlying these questions are the essential elements to be found in the criteria for selection.

Other than the Dutch experience English speaking countries reported virtually no conversion of software. Given that they are working with such a large pool of available programs the problem does not appear to be relevant. However one must ask how much good material is not finding its way into other countries because it is written for one system which is not predominent in the other country. Equally one has to raise questions about instances where software is bought, say from the USA into the UK. How much consideration is given in that selection to items which are specifically American (context, terms of reference, even differences between the two types of English) which really ought to be converted if the program is to be used effectively. One suspects on the basis of these reports that the problem is not recognised. Conversion is something that is necessary for "foreign" programs. It would appear that smaller countries could do well to consider the criteria and processes of conversion. Despite all the problems posed one can ride on the back of development work in other countries. One can buy in excellence. One might even be able to sell ones own excellence back into the market. The wheel is perhaps being invented more times and in more shapes around the world than is wholely necessary. To give them their special due, the Japanese may have a particular problem

5. Selection

Criteria for the selection of existing materials are perhaps the best known and most developed. They are most readily to be found where there are large amounts of software and the problems for the purchasers are those of knowledge of what is on the market and discrimination between the programs offered. It is not surprising therefore to find that in the USA there are a number of evaluation services. At a national level these may be published in journals, available to subscribers by post or even on-line. Some states have their own evaluation services and there are many more local or district services.

The selection of software is made in some cases by state bodies who then bulk-purchase on behalf of the schools, and in the majority of situation the selection is made by the individual teacher. In all cases it is important to the software producers that their products get evaluated. Without that sort of publicity (even with the risk of a bad review) making such a large audience aware of your products can be very expensive. The provision of information about what is on the market is obviously of major importance in a market such as the USA. Without this knowledge teachers will be hampered in their choice. Therefore many people demand, subscribe to and actively support evaluation services such as EPIE, MicroSIFT.

However some critical points emerge from the report. It seems that despite all these reviews the selection of media is more llikely to result from having seen a demonstration than from reading a review and that even where a review is read, the general description carries more weight than the apparently more scientific evaluation table. Despite their inaccuracies and variable terminology publishers catalogues still function as one of the main sources of guidance for the selection of software.

There are also some problems with the instruments which have been devised for the evaluation services. They purport to be scientific, and present their evidence as though it were objective. Yet most of these evaluation forms were evolved for the recording of both fact and personal judgement about drill and practice programs. It is not now certain that they function as well when applied to the more sophisticated applications programs. If for the sake of argument considerable importance is given to the criterion which checks whether a program gives a cumulative score for the user, then this will have little relevance to a program which is built like a data-base. A largely graphic program may not match the criteria which were established when most programs were textual.

It is not simply a question of having a more flexible set of criteria. One has also to recognise that these reviews are applying, of necessity generalised criteria to specific programs which may be used in educational contexts in the same country which are very different from each other. Where these reports do seem to score highly is in their informative function and in the general description of the programs. Research has shown that it is the descriptive element of these reviews and journal articles which have a major influence on the way in which teacher select software.

The canadian reviews are also based on what look like objective reports. Yet when one considers the forms it is easy to see that apart from details like to size of the disc, the type of computer system, the language and the memory requirement, most of the answers required are a matter of judgement on the part of the reviewer. Most of the instruments check whether the program is technically good and whether it has any factual errors. There are also general checks as to whether it matches the curriculum, but against which curriculum is the program to be matched when there are so many to be found in a country as diverse as the USA or Canada. Local or state reviews may have less of a problem in this regard, where curricular decisions may make it easier to match programs to lesson requirements.

What one ends up with are generalised criteria which require subjective judgements (often on a 1 - 5 scale) from the reviewer. These are then presented in tabular form as though they are definite values. Such presentations have to be suspect. However both the US and Canadian experiences seem to point to a heavy reliance on the part of readers on the descriptive reviews. For these to have currency there has to be an agreed basis for comparison. Here the generalised criteria have a major role to play. With these as a background the reviewer can then cope more easily with the varieties of programs and possible applications.

The Dutch review instrument relies heavily on subjective responses from the reviewers but recognises that these should be based on professional experience of the educational level and subject in question. It also provides quite clear guidance in its questioning. The resulting reports in SCEN are written like literary reviews. Readers can begin to get a feel for the way in which a particular journal reviews, just as they have their personal preferences for reviewers of football or theatre.

The strength and value of the descriptive review, particularly reviews of the application of a program - how it functions in a classroom - is strongly reinforced by the English report. Here there pointers towards a more fruitful use of software reviews. Those countries with less of a commercial market, either because of a dearth of software or because the choice of software is in the hands of national projects, offer little evidence of there being any criteria for the selection of software. It seems that where there is an opportunity for teacher to choose software then this is done on the basis of catalogues, magazine articles and demonstrations. It could even be that once political decisions have been taken to produce particular items of software, that evaluation may expose the project to the risk of bad reviews. Policy may on occasion take precedence over pedagogical judgement.

6 Research

Given that so much attention is being paid to evaluation around the world.It seemed only right to ask

whether it was of any value. Does it improve the choice of programs? Is program which receives good reviews going to make a greater impact in the classroom? Given the vaste range of questions that are asked, if one bothered to put all the evaluation forms together, one would have expected to have found some research to back up all this effort. Sadly, but perhaps not surprisingly there is very little evidence of research. Some academic researches are reported from the USA. What is of interest an value is the comment which is made, explicitly or implicitly, by the American, Canadian, Dutch and English correspondents that out of all this reporting and evaluation what has proved to be the most valuable has been the illuminative research and evaluation. This is by its very nature more an evaluation of the software in use in a learning situation than an evaluation of the software as a "thing". The most prized evaluations are of the *process* rather than of the *product.*

CONCLUSIONS

1 Development: There is considerable value in having guidelines and principles for the development of software. These should be evolved with a strict relation to the existing curriculum and methods of learning. At the same time they should be designed in such a way as to guide program developers towards those methods and subjects which are desired for the near future. Where programs are developed within the educational structures, such guidelines should be applied as widely as the organisation of education permits. The only proviso is that those administering these standards should recognise that they are guidelines and not apply them so rigidly as to crush innovation. Where development takes place outside the educational system one can do little more than offer the guidelines. In smaller countries and language areas where government support may have to be greater. Here the golden rule applies. He who has the gold makes the rules. The authorities have the opportunity to influence development to a greater degree than has perhaps been recognised.

2. Production: After years of program production of very variable quality it ought to be possible within education to limit some of the idiosycracies. It ought to be possible to agree to a limited number of languages (if not a single language). It ought to be possible to agree to a small number of processes such as always keeping text in separate files so that translations or versions of a program can be made more easily. It ought to be possible to operate with a limited number of common commands. It ought to be possible to agree within any one system to a range of basic standards for screen presentation, e.g. number of lines of text on a screen, the design and spacing of questions, the use of colour and colour contrasts, and so forth. This should be possible to make it easier for the users to work with new software. There could even be agreement about the basic principles of putting menu bars and pull-down menus into programs. It ought to be possible.

3. Selection: The conclusion has to be drawn from these reports that where the schools have a choice of software there is an important role for what are now termed evaluations. It also has to be recognised that the majority of what are put forward as "evaluation forms" are very flawed and when applied rigorously across the range of software, largely inappropriate. Where evaluation forms do score highly is when they give the reviewer guidance on the aspects to look for and how to make judgements in the broad sense. They also act as a check list to reduce the risk that a reviewer might forget aspects which may be crucial to other potential buyers. If we recognise that the apparent objectivity of most of the published evaluations is an unsupportable pretence and that they are subjective but written on the basis of agreed standards, then we can give that subjectivity its true value and start treating them as reviews of products. What we then need to do is extend this practice of reviewing to equally "guided" reviews of the programs in used - the process for which the software was originally written. Let us recognise that this illuminative evaluation of the process is the most valuable guidance that we can provide for education.

4. Conversion: Without labouring the argument that there are good reasons for converting software, criteria can be established by which the value of a program for possible conversion can be