Pupil Learning Assessment in the UNO/ESSP Curriculum

John Bowers

EDC Consultant (six weeks)

Final Report

(ESSP)

April 1993

Pupil Learning Assessment in the UNO/ESSP Curriculum

This report discusses how techniques for assessing pupil learning are determined by the structure of the UNO/ESSP curriculum model, the forms assessments take in the curriculum, and recommendations for assessment-related tasks that facilitate implementation of the curriculum.

Without going into an in-depth review of the UNO/ESSP curriculum model, it still is useful to touch briefly on its structure to show how the curriculum model determines the way pupil learning is assessed.

The '4-MAT' rubric underlying the design of a lesson unit (a set of eight sequenced lessons) is a key part of the larger system model. The whole model also encompasses supervision, management, teacher training, and community outreach. Implementation of the curriculum obviously depends on training teachers to use lesson units -- they must know what to teach, how to teach this, and how to assess the effects of their teaching, but success in implementation will also depend on training head teachers to become effective on-site supervisors to order to insure quality control over teaching.

Scope and sequence tables identify concepts and topics in all lesson units, show where these are covered at each grade level, and provide a basis for instructional texts. A teachers' guide contains all lesson units in a subject for a particular grade, and resource books serve as a common knowledge and methods base for teachers as well as for lesson unit writers.

A lesson unit is divided into eight lessons. Each lesson states a brief and an expanded objective, lists teaching activities and methods, identifies materials to be used, and indicates how to evaluate pupil learning. A standard format needs to be stipulated; this facilitates acceptance and use by teachers, presents a clearer template for lesson unit writers, and sets the limits for assessment task specification.

A concept is at the center of each lesson unit. The first three lessons are intended to connect pupils' past experiences with the concept to be learned; two middle lessons formally define the concept and develop mastery learning; and three final lessons enable pupils to apply the concept by creating small demonstration projects. These three components imply three matching assessments of pupil learning for each lesson unit. Documented assessment sends a signal to teachers that assessment is a significant process and that documented assessment evidence is open to pupils, parents, supervisors and others with a legitimate right to know how well children are doing in school.

Assessment in the UNO/ESSP model

UNO/ESSP assessment means continuous assessment, which is a feature of each lesson unit, and which shares similar activities with teaching. The implications of these two factors are several --

(a) The learning effects of each lesson unit can be an immediate basis for evaluating the teaching of the unit, the units themselves, and the assessment techniques. To put it another way, immediate small-scale formative evaluation can be carried out as soon as lesson units are produced and tried out in the classroom. It is not only possible to quick fix existing units, but also to adjust the production process itself when necessary.

(b) A major purpose of continuous assessment is to identify pupils who are experiencing learning difficulties so that remedial action can be taken before they fall too far behind. It is certain that a future testing program will be needed that examines all children at the end of each year or after some other time period. Blueprints for these "final" tests for a grade will be based on continuous assessment tasks that measure a concept domain and that lead to item specifications similar to those for developing mastery learning assessment.

(c) If the process is switched, and a "final" testing program is developed before continuous assessment procedures are set in place, the program will stifle the curriculum since these tests cover only a limited number of concepts in a curriculum, whereas continuous assessment at the unit level covers the entire domain of concepts. "Final" tests have appropriate purposes, so ultimately continuous assessment and periodic external examination will complement one another. In the beginning stages of curriculum development and implementation, the best rule is to start with continuous assessment.

(c) Since continuous assessment shares similar activities with teaching, and vice versa, assessment and teaching activities should be simultaneously planned and iteratively developed during the writing of lesson units -- we never want teachers or lesson writers to treat teaching and assessment as if they were independent.

(d) Teachers teach what they know is tested, so if testing reflects what teachers are to teach, then intended teaching can be driven by assessment as well as by prescribed teaching activities. A sort of "double whammy" effect operates. Assessment is an integral part of the curriculum which ideally can facilitate the implementation of the curriculum.

(e) A resource book of assessment techniques is as important (if teaching and assessment are two sides of the same coin) as the resource book of teaching methods. A resource book of assessment techniques is needed.

Documented Assessment

Documentation permits assessment scores to be reviewed by head teachers, pupils, parents, and supervisors. Teachers also learn that keeping pupil assessment records is an important duty and a supervised component of the curriculum.

Although assessments are made by teachers for each of the eight lessons in a lesson unit, three kinds of assessment scores are documented. These correspond to three broad learning goals of each lesson unit -- engagement, mastery, and application.

Assessment (Engagement)

Teaching is designed to insure that pupils connect prior learning (in or out of school) with the concept to be taught; assessment is designed to insure that pupils who fail to connect will not be overlooked. Teachers must identify pupils having difficulty, note the types of problems that occur, and record corrective actions taken. Three techniques can be used for this assessment.

- o A teacher's diary or log.
- o Behavior ratings and checklists.
- o Attitude measurement of positive response toward concept.

Anecdotal records and notes are kept in a teacher diary in which problems in engagement are noted, as well as the names of pupils who experience these. Diaries are useful for recording unexpected critical incidents. Behavior ratings are made of specific kinds of desirable or undesirable behavior occurring over a fixed period of time (teachers can cooperate in doing this). Checklists are used to identify the presence or absence of specific listed behaviors. Group questioning and group response methods are used to elicit full attention and to identify which pupils have difficulties. Attitude measures are given to assess pupils' disposition to react positively to the meaning and the teaching of the concept throughout the rest of the lesson unit..

Assessment (Mastery Learning)

Assessment measures mastery learning and concept attainment -- the understanding of defined concepts and related information, links to other concepts, and concept attributes. Techniques used are common to classroom testing and the measurement of productive thinking. Context-dependent test items are used both for practice and assessment. The major objective is to assess the ability to attack and solve problems presented in new ways, as well as abilities that reflect flexible, original, and self-regulated thinking. These are what most people believe that school is all about -- learning the meaning of important ideas and measuring how well this meaning has been learned. Mastery assessment techniques also lay the foundation for future leaving and selection examination.

Assessment uses classroom testing techniques; several different types of item formats can be used depending on the nature of teaching activities. Oral quizzes are given in younger grade levels. Pupils are given short tests in fixed time periods. Items are of various types and similar to practice exercises. The kinds of items to use can include the following.

- Extended response items, when appropriate for a grade level, are good for testing depth of understanding, ability to organize thinking, supporting or justifying a position, stating a plan, and summarizing.

- Matching exercises are good for classifying entities and for identifying relationships linking classes. For example, matching pictures of common objects to their names is a matching exercise.

- Master list items are good for judgment discriminations.

- True-false questions efficiently increase content coverage and are useful for presenting generalizations, comparisons, causal propositions, instances and examples, procedural steps, and computations;

- Multiple choice items present new interpretive and context-dependent test material (similar to exercises), and can be used to measure higher-order thinking skills.

- Productive thinking is tested using measures of ideational fluency (production of symbols and ideas), flexibility (producing categories of ideas and symbols), and originality (producing novel or uncommon ideas); especially useful for metaphorical or analogical thinking, and for identifying relationships that elaborate meaning.

Assessment (Application)

Pupils develop demonstration projects that exemplify applications of concepts. Projects involve products created by pupils, either individually or in small teams, that demonstrate the meaning of concepts to parents, one another, or to others in the community. Teachers help pupils plan, organize, and produce at challenging but realistic levels. Assessment requires projects to be rated in terms of process (where taught procedures are important) or product (when the quality of the product is important). Assessment is documented by an overall merit rating made by the teacher of each lesson unit's project. Ratings must use simple process and product rating scales and rating criteria that are easy to understand and apply. Pupils may develop either individual or small group projects. Group projects foster cooperative effort. Ratings can be made by teachers, pupils, parents.

A combination of assessment methods is used to evaluate projects. Assessment ratings focus on the processes involved, on the products, or on a combination of process and product. Rating scales and rating criteria should be simple and easy to use so that there is not too much variation from one lesson unit to another. The following project ratings can be made by the teacher for each lesson unit project.

- Process ratings of the way a pupil plans the demonstration project; ratings can be made for clarity of plan, choice of methods, intended use of material, group cooperation.

- Process ratings of the steps used to develop a product; ratings can be made for conformity to plan, flexibility of development, wastage of material, use of time, and completion of all steps.

- Ratings of the qualities of products used for demonstration and the demonstration itself; ratings can be made for originality, relevance to the concept, social value, and organization of content.

Recommendations

Assessment-related recommendations are listed and discussed below. The aim of these recommendations is to facilitate implementation of the curriculum. They stem from the items discussed above and together represent a start in developing the assessment component of the overall curriculum. The recommendations address (a) sound linkage between teaching and assessment, (2) methods to document assessment results, (3) building assessment into supervision.

Linking Teaching and Assessment

The goal of the curriculum is train teachers what to teach, how to teach, and how to assess the effects of teaching. The first two items, the "what" and "how" questions, are treated in the two planned resource books -- the academic resource book and the pedagogical resource book. A third resource book is needed. This would be a catalog of assessment techniques with examples. Then three resource books would be consulted when planning a lesson unit. The books should be crossed referenced; this would reinforce the connection between content, teaching methods, and assessment techniques.

The assessment resource book will contain techniques applicable for each of the three facets of assessment: engagement, mastery learning, and application. Suggested techniques were listed for each in an earlier section. These would be fleshed out with full descriptions of procedures with examples. This is analogous to item

specification. Since item specifications have been developed at various levels of specificity, the project must set desired format. The format would presumably be consistent with the format of the other resource books.

> Recommendation 1: Design and develop an assessment resource book which (a) explains the rationale for three facets of pupil learning assessment: engagement, mastery learning, and application. This book is intended for use, together with the academic and teaching methods resource books, by unit writers when they produce lessons.

Documenting Assessment Results

Assessment results need to be documented for several reasons. A major one is that continuous assessment in the UNO/ESSP curriculum is criterion-referenced. This means that results are interpretable in reference to a well-defined domain of tasks.

Criterion referenced tasks are typically expressed as performance objectives, and these of course are integral to the lesson units in this curriculum. In this curriculum, the domain of concepts to be learned is the underlying basis for lesson unit objectives, pedagogy, content, and assessment techniques. Since the curriculum is organized into lesson units, assessment can not be anything else but criterion-referenced.

The great advantage of criterion-referenced assessment is that pupil performance can be reported in simple and meaningful ways to a variety of persons with a stake in education. These persons are teachers, pupils, parents, curriculum developers, teacher trainers, supervisors, educational planners. With a common documented basis of assessment results, a variety of reporting formats can be developed to serve the needs of these multiple audiences. For each of these, results can be interpreted simply as the number of concepts that a pupil has connected with, mastered, or successfully applied. This legitimizes percentage scores, which always appears to be the measure preferred by persons at all levels.

There are two factors to consider when documenting assessments. The first concerns procedures developed for teachers to record and maintain a roster of assessment results for lesson units covered in class. The second is concerned with gathering this information for central computer processing and analysis by the research and planning department. Design features are important in both instances. High quality on-site record keeping is necessary to insure the authenticity of results both for those directly examining the records and for researchers who transfer the recorded information for analysis. This seemingly simple notion of record management by teachers is not a trivial matter. Poor record keeping decreases the reliability of assessment measures.

Recommendation 2: Design and develop assessment record keeping procedures that enables teachers to maintain accurate continuous assessment results for pupils.

Recommendation 3: Design and develop procedures to computer process and analyze school site assessment result. (Recommend the Data Entry Management (DEM) software developed by Andreas Schleicher of the International Coordinating Center of the University of Hamburg, Germany. The software has data compatibility with dBase, SPSS and SAS.)

Assessment Supervision

The UNO/ESSP model addresses the need to supervise instruction in ways that support pupil learning. Assessment is designed to inform people what students know and can do; assessment is intimately tied to teaching, so the supervision of teaching implies supervision of assessment procedures as well. Supervisors must be trained how to evaluate the ways teachers use lesson unit assessment methods and how they also maintaining records of assessment results.

The development of a supervisor training program has always been part of the UNO/ESSP curriculum development and implementation plan, but details can be worked out as lesson units are written and teachers are trained to use them. Since no supervisor corps currently exists, head teachers must be trained for on-site supervision. Formative evaluation of early implementation trials will shed light on what tasks head teachers do and what tasks they will be required to do for successful implementation of the curriculum. The discrepancy between what is and what needs to be done will lay the groundwork for supervision training.

Recommendation 4: Design and develop a training program for head teachers and (ultimately) supervisors that among other things focusses on the rationale for assessment supervision, on assessment procedures, and on the functions that assessment serves -- valid measurement of learning outcomes, criteria of school effectiveness, and the promotion of positive attitudes towards education.

I want to express my appreciation for the opportunity to have worked with Dr. Donald Schutte and all of the Specialist Facilitators who generously allowed me to participate in daily seminars throughout this consultancy. They are an extraordinary cadre of professionals and it was a privilege to have worked closely with them.

References

- ESSP Assessment of Pupil Learning: Pupil Assessment Tasks, Briefing Report, March 1993.

ESSP Assessment of Pupil Learning: Pupil Assessment Tasks

John Bowers

EDC Consultant

Briefing Report

(ESSP)

March 1993

Table of Contents

ESSP Assessment of Pupil Learning: Assessment Tasks

Assessment of pupil learning in the ESSP curriculum here denotes continuous assessment. Developing continuous assessment tasks is an immediate priority with curriculum materials being prepared for use and evaluation. The need for conventional summative testing will no doubt arise in future, and that kind of testing program can be developed rather easily from a base of continuous assessment tasks. Before discussing pupil assessment, however, it may be best to review the materials being developed, their purposes, and how training is expected to influence curriculum implementation.

The ESSP/UNO Curriculum

The ESSP/UNO curriculum model designed by Donald Schutte combines instructional content, teacher training, and pupil assessment in a single total educational systems development program. Specialist facilitators first prepare academic and pedagogical resource materials and initial classroom lesson units for teachers and pupils. They then train teacher facilitators who in turn train head teachers and teachers. Teacher facilitators prepare more lesson units as part of their training. This procedure is followed when teachers and head teachers are trained. In this way, all teachers and trainers are brought into curriculum development as they learn about the model and how to implement it.

Scope and sequence tables in four subjects (social studies, science language arts, mathematics), prepared in three languages (English, Dari, and Pashto), identify subject concepts and topics, show where these are covered in grades one, two, or three, and provide the basis for four instructional texts. These are a Teacher's Academic Resource Book, a Teacher's Pedagogical Resource Book, a Teacher's Lesson Guide, and a Student Text.

The two resource books comprise a common knowledge and methods base for teachers. The Teacher's Lesson Guide (with the Student Text) includes all lesson unit content for each grade level in a subject.

o Teacher's Academic Resource Book: This book for each subject contains information about concepts, topics, and subtopics that pupils are to learn. Content covers grades one through three. Knowledge is propositional (knowing about), procedural (knowing methods), and conditional (knowing when to apply knowledge). This book is used by teachers as a reference and by facilitators and trainee/curriculum developers who select from it relevant content related to concepts, topics, and subtopics for each grade.

o Teacher's Pedagogical Resource Book: This second resource book consists of general pedagogical prescriptions. These include, for example, teaching strategies related to "concept attainment" and a teaching method that "compares and contrasts."

o The Teacher's Lesson Guide: There is a Teacher's Lesson Guide for each subject at each grade level. It consists of lesson units, each following a "4-MAT" framework (McCarthy, 1980), with expanded objectives, teaching and learning activities, materials to be used, and ways to assess pupil learning. These components are organized to facilitate the learning of selected concepts and topics developed in the Teacher's Academic Resource Book, as well as the application of appropriate teaching activities described in the Teacher's Pedagogical Resource Book.

o Student Text: The Student Text presents content for pupils that corresponds to that for teachers in the Teacher's Lesson Guide. The format of Student Texts provides pupils with material to stimulate their awareness and response to concepts in the lesson units.

The two resource books are now developed in each subject to a stage where they are ready to use for training, with some expansion needed as further lesson units are produced. One lesson unit has been produced for grade one in each subject. Production now will center on the completion of Teacher's Lesson Guides and Student Texts as specialist facilitators begin the training of teacher facilitators (who become producers), who then train head teachers and teachers (a third tier of producers).

The ESSP/UNO Lesson Unit

A lesson unit is divided into eight sequential segments that focus on four types of learners (imaginative, analytic, common-sense, and dynamic) and that represent different stages in teaching the unit. The heart of the lesson unit is a concept to be taught. The three first segments engage the pupils' prior personal experience with the concept and enable them to analyze and reflect on this, two middle segments formally define and develop the concept and provide practice exercises, and three final segments enable pupils to extend and apply the concept in his or her daily life through individual projects. Each segment contains an expanded objective, lists teaching activities and methods, identifies aids and materials used, and suggests ways to evaluate pupil learning (these point to the design of more detailed assessment tasks).

Developing the "formal" segment is done first; the lesson unit's selected concept is the basis for preparing an expanded objective around which to organize information about the concept for pupils. Teaching methods, strategies, and techniques are identified; teacher aids and materials are indicated; and ways to evaluate pupil learning are specified. Instruction does not start with this first-developed segment. This segment is taught after pupils complete the first three segments.

The eight segments are described briefly below.

Segment A: The objective is to create learning experiences that enable pupils to connect personal experience to the concept. The teacher is actively in charge, and attempts to provide conditions for pupils to discover the concrete meaning of a concept. The assessment of learning at this stage is very informal and is actually hard to distinguish from teaching. Teachers must observe and note whether all students are becoming involved, must identify those pupils having difficulty, and must be confident that all pupils are engaged for further learning. Teachers should begin a teacher log or diary at this time.

Segment B: The objective is to examine the "whole" experience of Segment A. Pupils now begin to analyze the concept, more as a concept outside themselves which has external meaning. They can be asked to provide additional concrete examples of items in the real world that may or may not meet the attributes of the concept. All students may not yet grasp the concept in a precise way, but the segment's activities enable them to do much trial-and-error learning. Though evaluation is still informal, use of checklists and ratings by teachers is indicated.

Segment C: This is a bridging segment. The objective is to help students clarify understanding of the concept apart from personal experience alone. Teachers should document assessment at this stage to insure that all pupils have indeed connected their concrete experience with the concept being taught and have met all of the preconditions for the formal teaching in Segment D. Assessment at this time should also involve attitude measurement to determine whether pupils may be expected to respond positively to further teaching of the concept.

Segment D: This is the segment that is developed first. It treats formal definition of the concept, gives information to pupils about the concept, and provides learning experiences that enable pupils to develop semantic understanding of the concept. In one sense, this is what most persons think school is about, and in most instances this is where ordinary lesson units start. In the ESSP/UNO curriculum, however, the very important need is to first connect the pupils' personal experiences (and attitudes and motivations) to the concept to be taught.

Segment E: This segment emphasizes practice that enables pupils to gain independence in manipulating and generalizing the defined concept. Pupils begin to become self-regulated learners and gain confidence through exercises involving the concept. Assessment at this stage uses formal methods and techniques, but does not have to be dull. Learning acquired at the end of Segment D and Segment E (formal teaching plus practice) should be regarded as mastery learning. Pupils, parents, teachers, persons in the community, supervisors, educational authorities, curriculum developers, and teacher trainers expect mastery learning assessment to be documented, and this is the place to do it. This assessment should not, however, be regarded as superior or more important than informal assessment carried out during the first three segments.

Segment F: Learners begin concept applications in this segment. Pupil independence begins with opportunities to use the concept by adding something of their own. They do this by starting development of a project by the end of this segment that exemplifies their personal view of the concept. The project might be a stand-alone product or a home or community demonstration activity. The project will carry the pupil through to the end of the lesson unit. Projects involve trial and planning, setting goals, organizing, creative thinking, and other combinations of higher-order abilities and skills. The teacher is a facilitator, and fills this role more and more as pupils gain confidence in their independence and their success with projects. Each pupil's projects will be placed in a portfolio. Evaluation of the portfolio requires rating criteria and methods for rating. Since there is probably more than one teacher in a school, teachers should share rating tasks with one another, and there are good reasons to allow parents and pupils themselves to rate projects. In fact, when a project involves a home or community demonstration activity, those persons to whom the demonstration is made should be asked to provide judgements of the pupil's performance and the perceived social value of the project.

Segment G: The pupil continues to integrate and develop project applications of the concept. Pupils are given opportunities for self-discovery by being allowed to actively exercise their inventiveness. Projects involving pairs or small groups of pupils might be designed to allow pupils to teach concept applications to one another. The teacher's role is that of constructive evaluator and remediator, helping pupils to exemplify their best work.

Segment H: Projects are completed and demonstrated as planned at the end of this segment. Ratings and judgements are made and documented. However, as anyone who has dealt with portfolios knows, portfolio products become sources of pride and discussion in the community. Portfolios that contain serious work with social merit can create immense good will for schools because they symbolize what children can do far more dramatically than rosters of examination scores. When an open-house is held at a school, or when a supervisor visits a school, attention is conventionally directed to formal assessments and grades (like those from Segment E), but it is almost a certainty that parents, supervisors, and others who visit the school will remember and talk about folio products.

Implementation Linked to Training

Successful adoption and implementation of the ESSP/UNO curriculum depends on the expected involvement of head teachers and teachers in the preparation of lesson units. Producers become consumers; teacher participation in curriculum development leads to their acceptance of the model and their commitment to its effective implementation. Teachers-as-developers learn lesson unit objectives first hand and then create appropriate teaching activities, methods and assessment tasks that measure the attainment of objectives.

Behavioral Objectives and Assessment Tasks

Objectives of course are the basis for assessment tasks. Objectives serve many purposes. They direct instruction, provide framework for analyzing what is taught, give teachers a basis for evaluation, tell pupils what they are expected to learn, and communicate the intent of instruction to parents, other teachers, administrators and the public (Nitko, 1983). These multiple purposes carry over to assessment tasks -- so we must always remind ourselves that these tasks and their measures carry meaning to a variety of audiences such as students, teachers, headteachers, educational authorities, parents, and community figures. In order to reach these various persons, the development of assessment tasks must be accompanied by the development of methods to store, manage and report information about learning effects measured by assessment tasks.

Objectives have to strike a balance between being too vague or too specific. They are described best when teachers can identify what students must do so that intended learning can be validly inferred. Assessment tasks specify these intentions in greater detail. Several kinds of assessment tasks may be developed for any one objective. Assessment task specification must clearly describe methods to evaluate what pupils learn. An earlier overview listed the elements of objectives and the need to cross-classify these into categories representing content and learning dimensions. That information is presented in a note titled *ESSP Assessment of Pupil Learning: Cognitive and Content Taxonomies,* which contains a brief statement on objectives and different ways to classify cognitive, affective, and psychomotor taxonomies. We need now to develop a framework for assessment consistent with the Schutte model. An earlier note listed some guiding principles:

(1) Assessment is part of each stage in each lesson unit.
(2) Assessment purposes and criteria must be open and clear.
(3) ESSP assessment is criterion-referenced to well-defined tasks.
(4) Continuous assessment has priority over certification testing.
(5) Affective as well as cognitive outcomes are assessed.
(7) Assessment uses different formats.
(6) An objective may be assessed in a variety of ways.
(7) Assessment features individual and group activities.

Assessment is always part of all eight segments of a lesson unit. However, it was proposed above that assessment be **documented** in three ways for each unit. There is **Assessment 1** done during the first three segments of a lesson unit (A, B, C) which relates to pupils' connecting their experiences to a concept; **Assessment 2** done at the end of the next two segments (D and E) which is like conventional classroom examination, and **Assessment 3** based on projects developed in the last three segments (F, G, H).

The 4-MAT model is designed to insure that all types of learners receive equitable conditions of instruction. When 4-MAT based teaching is fair to all learners, then assessment that is tightly linked to teaching will also be fair to all.

Assessment 1

Purpose: The teacher assesses in order to verify that pupils connect their prior experience to the concept being taught. This early connection is an essential part of the ESSP/UNO curriculum and is to happen for all types of learners. Teaching is designed to insure that different types of learners are successful in achieving this connection; assessment is done to insure that no pupils are overlooked.

Teachers must identify pupils having problems with the lesson, note the nature of these problems, and record what corrective action was taken. Pupil involvement will to a large degree be assessed using group response techniques (e.g., hand card response and thumbs-up-and-down). The teacher must also learn whether all pupils are motivated and willing to receive further teaching of the concept.

Learning Levels: Knowledge of specifics from prior learning that are connected to the concept taught; comprehension of the concept in both concrete and abstract terms; application of inductive reasoning to identify examples and non-examples of the concept and identification of rules that define the concept; analysis of entities to identify attributes corresponding to the concept; synthesis/evaluation that integrates concept relationships and applies criteria to determine appropriate concept meanings.

Methods: It is proposed that three techniques be used for this assessment: (1) a teacher's diary or log in which successes and problems are noted (and individual students who experience these), (2) behavior ratings and checklists that indicate pupil response and involvement, (3) attitude measurement to learn whether pupils show positive response toward the concept.

Anecdotal records and notes are kept in a diary or log by the teacher. These are useful for recording unexpected incidents that involve only one or a few pupils; these can be critical incidents. Behavior tallies are made of the number of times a specific kind of behavior occurs over a fixed period of time. Checklists are used to identify the presence or absence of specific listed behaviors, one student at a time. Group questioning and group response methods are used to elicit full attention and to determine which pupils may be having difficulties. Attitude measures are given to assess pupils' disposition to react positively to the meaning and the teaching of the concept.

Activities: Pair-sharing; teacher questioning with pupil group response; self-report by pupils; teacher observation and recording.

Assessment 2

Purpose: Assessment here measures mastery learning and higher-level thinking skills. The focus is on the understanding of defined concepts and information related to them, links to other concepts, attributes of the concept, and the ability to use higher-level thinking in new and novel problem-solving situations. Developing these kinds of formal assessment tasks will lay the foundation for constructing conventional certification examinations when these may be needed in the future. The main things to assess at this stage are concept and principle learning. Techniques used are familiar ones common to classroom testing as well as unfamiliar ones used to measure productive thinking (creativity). Much use should be made of context-dependent testing techniques both for practice and assessment. The goal is to assess abilities to attack and solve problems presented in new ways, and abilities that reflect flexible, original, and self-regulated thinking.

Learning Levels: Knowledge at the lowest level is indicated by recognizing and recalling essential concepts in the subject's body of knowledge, the ability to quickly recall many examples and non-examples of the concept, the ability to identify or state rules that define the concept. Comprehension is shown by finding similarities between the concrete and abstract forms of the concept and by comparing the concept's tangible attributes and its defined characteristics. Concept learning is measured in three ways: by producing new (untaught) examples of the concept when asked, by idetifying or defining the concept when new (untaught) examples are given, or by paraphrasing the meaning of the concept. Application of rules requires relating concepts (e.g., if-then, cause-and effect, before-after) to new situations and by combining previously learned rules in new ways to solve a problem. Knowing when the concept applies or not is application. Lower level deductive reasoning is shown by developing conclusions and generalizations based on the concept. Reasoning involving analysis and synthesis is shown by analyzing concept attributes to form generalizations. Going beyond simple analysis to complex comparisons that classifies entities on several attributes reflects a combination of synthesis and application.

Methods: There is use of several different types of examination item formats. These include essay and short-answer or completion items when these are grade-level appropriate; matching exercises (good for classifying things and for identifying relationships), master list items that are similar to matching exercises and good for making judgement discriminations, true-false questions that efficiently increase content coverage and are useful for presenting generalizations, comparisons, causal propositions, instances and examples, steps in procedures, and computations; multiple choice items which present new or novel interpretive and context-dependent test material (similar to exercises).

Creative thinking is tested using measures of ideational fluency (production of symbols and ideas), flexibility (producing categories of ideas and symbols), and originality (producing novel or uncommon ideas).

Activities: Pupils are given short tests in fixed time periods. Items are of various types and should be similar to practice exercises (and the other way around). The teacher should discuss the tests immediately after they are completed and the papers collected. Pupils should be encouraged to explain and even defend their "wrong" answers, and asked to share their creative thinking responses with the class.

Assessment 3

Purpose: Assessment is of pupils' projects in which they develop tangible applications of concepts. It is expected that projects will typically involve demonstration activities by pupils either individually or in small groups (where teamwork is learned). Teachers play significant roles by assisting pupils to plan, organize, and produce at challenging but realistic levels.

Projects will have tangible products. Assessment will consist of ratings in terms of process (when procedures are taught, departures from steps can be identified, and the way things are done is important) or product (when performance is shown mainly in the product, process is indeterminate, or steps may not be specifically taught). It may be more important to assess process rather than product in earlier lesson units until pupils acquire confidence.

Assessment criteria must be defined and ratings on these criteria must not impose undue time demands. Criteria should have social relevance to pupils, parents, educational authorities, and persons in the community. Ratings of projects could, for instance, be made in terms of connection to the lesson unit concept, organization, quality, and social merit. Ratings must be simple enough so that persons other than teachers can apply them. Even though there will be issues regarding rater reliability, development of assessment procedures should begin quickly without spending too much time now on psychometric issues. Rating reliability will develop over time. The immediate first need is to establish a framework for project creation, rating, and the roles and activities of teachers and pupils.

Learning Levels: Projects do not easily lend themselves to analysis in terms of inferred levels of learning, even though we might reasonably believe that all combinations of learning levels are reflected in the processes pupils use to create their projects. Pupils appear to apply the core thinking skills of Marzano et al. when they are focussing, goal-setting, planning, assembling needed materials, attempting various trial efforts, and integrating their activities into a completed project. In the Bloom taxonomy,

knowledge and comprehension are shown through the project's relevance to the concept; application by applying rules relating to concepts and by appropriate production procedures; analysis and synthesis are continuously blended during the creation of the project. Productive thinking skills such as fluency, flexibility, and originality are certainly called often into play.

Methods: Probably a combination of methods should be tried out. This is certainly a matter to be addressed when evaluating the curriculum's implementation. Factors to consider are the role of the teacher, whether process or product assessment is emphasized, how to rate process (e. g., procedures are planful, clarity of goal, choice of materials) or product (e.g., conformity to the plan, suitability to the intention, and social value of the product in the minds of parents and community), and whether others besides the teacher are to rate products.

Activities: Pupils may develop individual or small group projects. Projects are completed within the lesson unit's time frame so that incomplete projects do not pile up. Ratings are made by the teacher and perhaps by other teachers, by pupils themselves, or by parents.

Databases for ESSP Curriculum Evaluation

ESSP continuous assessment produces a large amount of data, so means of efficient storage and quick access for statistical aggregation and analysis need to be developed. If a data file contains three assessment measures for each lesson unit, and if there are 30 pupils participating in 25 lesson units in a year per subject, the assessment database contains 30x25x4x3=9000 measures, or 2250 measures per subject each year per grade level. One can easily get carried away by asking for too many measures to be stored at the individual level. That is one reason to limit the storage of assessment measures to three per lesson unit.

Teachers will have access to the most detailed evidence of pupil performance, but the database mentioned here should provide good outcome information for purposes of pupil grading, supervision, curriculum evaluation, teacher training, and so on. Curriculum developers need to know whether some curriculum areas are learned better than others; trainers and supervisors need information on pupil learning for preservice and inservice; policy makers will depend on the assessment database when determining whether there are substantial school differences in achievement and to what these may be attributed; parents and the community will look to student learning measures as direct evidence of the effectiveness of schools.

We should bear in mind that data other than assessment measures must be collected in order to evaluate the ESSP curriculum. We need to create computer data files to document various training activities, and to have easily accessible information about

pupils, teachers, and schools. One source of information will be the IEES database. IEES work is now in the planning stage; the intention is to collect information in special studies. ESSP/UNO curriculum evaluation will be one of these special studies conducted when the curriculum is implemented. It is important that specialist facilitators determine the kinds of information needed for evaluation so that cooperative work can be organized with the research and planning department.

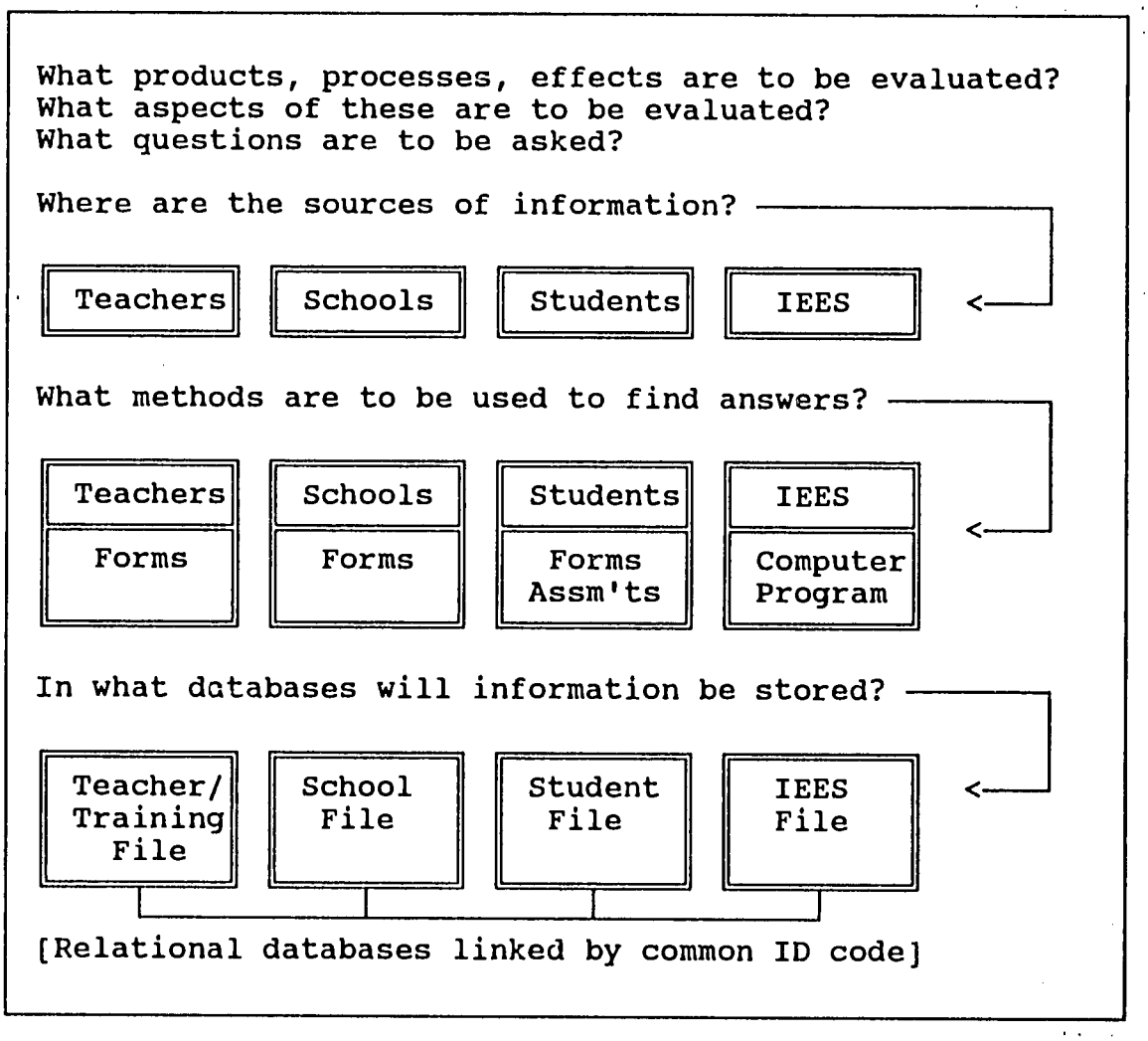
Information for the separate databases will largely be collected using questionnaires or special forms (interviewing should not be ruled out). Information about teachers' qualifications, age, sex, experience, subjects and levels taught, and so on, is collected with a Teacher Questionnaire. Information about schools such as number of teachers, water and latrine facilities, available materials will be collected by a School Questionnaire filled out by headteachers or principals. Pupil information will be obtained from assessment rosters maintained by teachers and from data collected with a Student Questionnaire. A special form is needed to detail information about training. The IEES master database will be accessed by a computer program in order to select necessary data. Use of any of these data collection methods should be piloted and instruments should be precoded as much as possible.

Each piece of information in a database (e.g., a questionnaire item) should be described in terms of its instrument reference, the meaning of coded values, the range of acceptable values, the type of variable the information represents, how missing values and "not applicable" responses are treated, and whatever other characteristics would enable a stranger to gain a clear picture of the information contained in a database.

An excellent program is available to ESSP/UNO for database entry, data cleaning, and documentation. It is DataManager, written by Andreas Schleicher in the International Coordinating Center at the University of Hamburg. It was used in over thirty countries in a recent International Reading/Literacy Study to create databases for teachers, students, and schools. DataManager requires the user to develop a well-defined data structure file for variables. Once this is done, the menu-driven program acts like dBASE but has added features such as entry checks, data cleaning rules, import/export options, and creates SAS and SPSS+ input statements and a codebook. Dr Schutte has this program and so does the Research and Planning Department for examination.

Planning ESSP curriculum evaluation starts by asking the critical questions: What curriculum products, processes, and effects are to be evaluated? What particular aspects of these components are of main interest? What questions are to be asked? Where are the sources of information to answer these questions?

The major sources of information of course are pupils, teachers, headteachers, and the IEES database. ESSP curriculum evaluators will develop questionnaires and forms that correspond both to specific sources and to databases. For example, ESSP evaluators will need to develop a School Database, a Teacher/Training Database, and a Pupil Database. The IEES File will be accessed by computer request to extract relevant information. These are relational data files, which simply means that variables across forms can be collated using key variables common to all, such as school identification code. IEES must set a standard code for school identification. The figure below illustrates the scheme.



Readings

McCarthy, B. (1980) The 4-MAT System. Oak Harbor, Illinois.: Excel,Inc.

Nitko, A.J. (1983) Educational Tests and Measurement: An Introduction. New York: Harcourt Brace Jovanovich, Inc.