DOCUMENT RESUME

ED 076 365 SE 015 683

AUTHOR	Tisher, R. P., Ed.
TITLE	Research 1972.
INSTITUTION	Australian Science Education Research Association.
PUB DATE	72
NOTE	131p.
AVAILABLE FROM	University of Queensland, Department of Education, Brisbane, Q 4067, Australia (no price quoted)
EDRS PRICE	MF-$0.65 HC-$6.58
DESCRIPTORS	Curriculum; *Educational Research; Educational Researchers; Instructional Materials; International Education; *Learning Theories; *Measurement Instruments; *Research Reviews (Publications); Science Course Improvement Project; *Science Education
IDENTIFIERS	Australia

ABSTRACT

Presented in this monograph are 10 articles read at the Third Annual Conference of the Australian Science Education Research Association which was held in Melbourne in May, 1972. The articles are given under the following headings: A Model for Curriculum Evaluation; The Australian Science Education Project as a Stimulus for Research--A Progress Report of a Study of Teaching Strategies Used with Australian Science Education Project (ASEP) Materials; Some Suggestions for Research Related to ASEP; A Checklist for Analyzing the "Style" of Instructional Materials--A Report Based on Work by S.R. Shepherd; Classroom Interaction: The New Religion; Prior Knowledge--A Source of Negative Factors for Subsequent Learning; Difficulties with Non-technical Vocabulary amongst Junior Secondary School Students: The Words in Science Project; The Development of a Creativity Test; Outline of a Learning Hierarchy Investigation; Basic Skills of Graphical Interpretation; and A Survey of Evaluation Instruments. The first four articles are associated with ASEP materials, the fifth is a review of research, and the remaining deal with antecedent variables, learning theory, graphical methods, and measurements. A biographical sketch is given for each of the authors at the end of the monograph. (A related document is ED 065 325.) (CC)

THE AUSTRALIAN SCIENCE EDUCATION RESEARCH ASSOCIATION

RESEARCH 1972

Edited by

R. P. TISHER

A publication containing papers presented
at the Association's Third Annual Conference
held in Melbourne, in May, 1972.

Executive Officers: *Dr. C. N. Power*
Dr. R. P. Tisher

Address: Faculty of Education,
University of Queensland,
BRISBANE, Qld. 4067.

PREFACE

The first publication of the Australian Science Education Research Association. *Research 1971*. was a modest one for a group with limited funds and clientele. However. the encouragement. co-operation and assistance of many persons and organisations guaranteed the success of the publication. Consequently. the association. with greater confidence, is producing its second modest, yet significant, monograph which contains the papers read and tabled at the third annual conference in Melbourne. May 1972.

The conference was held at the headquarters of the Australian Science Education Project. The venue was significant and appropriate on several counts: first. A.S.E.P. had assisted greatly in the production of *Research 1971*. and second. a major theme of the conference was "research that can be based on A.S.E.P. materials". The first four papers which appear are associated with the major theme. The remaining six deal with a variety of topics but appeared to belong to certain clusters depending on themes which appeared in them. The responsibility for the grouping rests entirely with the editor. but the categorization does highlight some of the areas in which research is proceeding in Australia. It is appropriate to point out that the views. opinions. interpretations and implications expressed are those of the individual authors and. not necessarily those of the Australian Science Education Research Association or the Editor. Furthermore. rigorous cutting and editing of manuscripts have not occurred because the Editor believed the publication should present an accurate record of the proceedings of the annual conference. The reports evidence a wide range of interests and styles in research. This is an encouraging feature and an acknowledgement that a range and variety of studies are essential, and appropriate, if research in science education is to have an impact in the real world.

Research 1972 is further encouraging evidence of a continued interest and activity in research in science education in Australia. Hopefully, other research workers will now build onto rather than repeat. studies which have been reported here. Certainly, more systematic, well-conceived and well-executed research is required. There is a challenge to increase the volume of research while maintaining its quality and relatedness to the real world. Moreover, there is the exciting prospect that much valid research evidence will accrue to influence educational theory and practice.

R. P. Tisher
University of Queensland

CONTENTS

RESEARCH ASSOCIATED WITH A.S.E.P.

A MODEL FOR CURRICULUM EVALUATION

Application of a general model for formative evaluation to the evaluation of ASEP units.

N.L. Baumgart

By way of explanation

Although this paper considers a general model for curriculum evaluation, the principles are applied in particular to the evaluation of the ASEP units. This interest stemmed initially from a visit to ASEP Headquarters by a team of three consultants from the School of Education at Macquarie University. The report furnished by this team and subsequently distributed by ASEP concentrated on "a critical survey of the current methods of evaluation of the ASEP-produced materials, the development of a model for the study of variables in future ADEP evaluations, some observations about educationally significant contributions made by ASEP, and some thoughts on future developments".

At previous conferences of the Australian Science Education Research Association, interest has been expressed, both formally and informally, in the conducting of research allied to the ASEP Project. In a subsequent paper, Dr. Power will discuss one such project already initiated in Brisbane. One research area in which ASEP staff have had a heavy commitment and in which researchers from other institutions have expressed interest is that of curriculum evaluation.

The Consultants' Report by the Macquarie team who visited ASEP referred to the thorough procedures adopted by ASEP staff in building-in the evaluation of materials and practices as an integral part of the "38 stages in the development of a unit". However, the Report also noted the restricted facilities and manpower available to ASEP at present and suggested benefits which would accrue from the involvement of other institutions in the evaluation program.

However, if a number of people from different institutions were involved in such a project, it would be desirable for them to work within a consistent research paradigm and to use a similar conceptual framework. In this way, their findings would be not only directly comparable, but also cumulative. Recognizing that there could be a number of people who would be interested in the evaluation of ASEP units, this paper considers a model for curriculum evaluation and then applies the model to ASEP evaluation.

Why Evaluate Curricula?

"While you and I have lips and voices which are for kissing and to sing with, Who cares if some one-eyed son of a bitch invents an instrument to measure spring with."

C.E. Cummings

While we might appreciate the sentiment expressed by Cummings with respect to seasonal change, to apply a similar philosophy to school curricula is, in the opinion of the writer, to court disaster. Nevertheless, it is just this sin of omission (or failure to evaluate our curricula) to which we in Australia must plead guilty in many instances. We have been so concerned with appraising students through our systems of external examinations, and so concerned with appraising our teachers through our inspectorial systems, that we have failed to come to grips with the more extensive problem of curriculum evaluation. It might even be suggested that our philosophy with regard to curriculum evaluation could be expressed by an appropriate modification of the verse above:

> While teachers and students have schools and
> books which are for teaching and to learn with,
> Who cares if some one-eyed son of a bitch
> invents an instrument to measure curriculum with.

It must be recognized that some educators would argue strongly for just this view-point. They would claim that the really important objectives for school learning concern general transfer to real life situations. In addition, many of these real life situations are met years after the school experience so that measurement of responses to them is not feasible. Besides, they would claim, even if one could measure such things, what criteria would be used to evaluate the measures?

These are legitimate questions deserving careful consideration. Nevertheless, there is evidence to suggest that our failure to evaluate curricula in Australia has had far-reaching consequences. On a few occasions the results have been so disastrous that drastic measures have been necessary. How many curricula in Australian schools have undergone systematic or even haphazard modification based on evidence collected from planned evaluation. Most decisions relating to our curricula have been to label some aspects as "non-examinable" or to abandon the entire venture in favour of an imported model. Moreover, most of these decisions have been made against a background coloured by a tirade of letters to newspaper editors or the militant demands of an harassed teaching profession.

Nevertheless, we might be encouraged by recent trends in which at least an increasing awareness of the need for planned and systematic curriculum evaluation is apparent. For example, at the "Guidelines Conference" of the Australian Science Education Project held at Monash University in January, 1970, considerable time was devoted to papers and to discussion on the evaluation to be undertaken of ASEP. The use of local and national trials by ASEP, and the trialling of other science courses in several Australian States are further evidence of the recognition of the need to evaluate curricula.

A General Model for Curriculum Evaluation

In this paper, *curriculum* will be considered in the "narrow" sense – as a particular program of studies – rather than in the broader sense as the sum total of all of the child's experience in the school. *Evaluation* involves some judgment after a set of measures or observations are compared with a set of criteria. Initially this involves an estimate of the extent of CONGRUENCE between the set of measures or observations and the set of criteria (Figure 1). Judgments then have to be made about the adequacy of this extent of congruence.

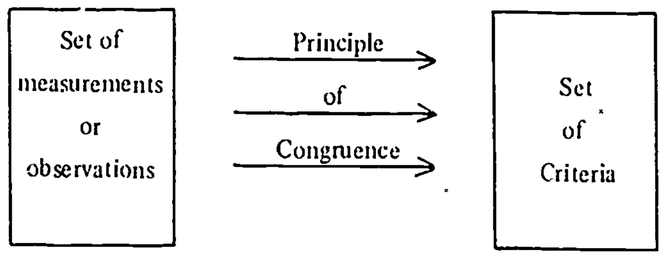


Fig. 1

Schematic representation of a simplified evaluation model

Although this simplified model forms the basis for what follows, in this form it leaves much unanswered. *What* measurements or observations should be made? *Which* criteria should be used? Attempts to answer these questions will result in refinements to the basic model.

The criterion problem

In general, there are two possible bases for judgment. The first is a *relative* standard in which the criterion becomes the outcomes of another curriculum, or some control group. The second possible base for judgment is an absolute standard. In most instances this will be particular course objectives or intended outcomes. These will include objectives relating to student behaviours, attitudes, and skills; intentions for the use of aids or apparatus in classrooms; intentions for methods to be adopted by teachers; possibly even intentions for the preservation of apparatus, and so on. Most often this kind of evaluation will include a feedback loop. Knowing the extent to which objectives have been attained, appropriate modifications can be effected provided adequate diagnostic information on the implementation of the curriculum is available.

Writers on curriculum have disagreed as to which of these deserve greater encouragement. Two examples will serve to illustrate the different emphases. Cronbach (1963) considered that generalizations which were made from comparisons between different curricula were hazardous and poor research investments. He therefore advocated curriculum "case studies" in which the aim was extensive measurement and thorough description. However, measurements so obtained should not be combined to produce single scores but judgments should be made about separate outcomes when these are considered with respect to particular course objectives.

On the other hand, Scriven (1967) advocated the use of a relative standard. He questioned whether absolute standards could be adequately defined or prescribed. He saw the educator or administrator as one who wanted to know which curriculum was best and he further considered that the curriculum evaluator was the one who must take responsibility for discriminating between poorer and better programs.

The kind of distinction made above is reflected in the differentiation which is commonly made between FORMATIVE and SUMMATIVE evaluation. Formative evaluation refers to the assessment of a curriculum during its course of development. It implies that evidence will be obtained on aspects of the course as it undergoes trial. On the other hand, summative evaluation consists of an appraisal of the final product. Although this could be done with respect to predetermined standards or objectives (absolute criterion), in practice it more often involves comparisons with other curricula (relative criterion).

In this paper attention is focused on formative evaluation and the model is described in terms of it. Reasons for this emphasis are given below. However, before leaving summative evaluation, mention will be made of a few examples. Williams (1968) gives several examples of summative evaluation in the United Kingdom. For example, he reports on a survey by Biggs published by N.F.E.R. into "methods" of teaching arithmetic. The survey reported on a wide variety of outcomes for students taught under five different methods:

> "traditional methods; 'uni-model' structural methods, like the Stern and the Cuisenaire, in which only one representation of mathematical situations is used; a 'multi-model' structural method, in the form of the Dienes method, in which a variety of representations of mathematical situations are used; 'motivational' or 'activity' methods; a mixture of motivational and traditional methods."

In as much as the survey attempts to determine how these methods relate to certain outcomes, it is an example of summative evaluation. Within Australia, examples are provided by several studies which have compared achievement and other variables of students following "new" science courses with similar scores for students following traditional courses. (See, for example, Lucas (1969), Baumgart (1969), Mackay (1971).

Although it might be argued that summative evaluation is essential to the administrator when he has to make decisions about implementation of new curricula, it does have many limitations. Summative evaluation does not make provision for feedback and hence for change. Its concern is efficiency rather than effectiveness. Thus it looks only at the final PRODUCT. Where this product is faulty it gives no indication of weaknesses in the PROCESS. If the evaluation is based on relative standards (comparison of two or more courses), then use of summative evaluation is likely to conceal as much as it reveals. An end-of-course examination is a delayed criterion measure and, as such, is subject to contamination from several sources. Moreover, scores on such an examination have resulted from an interaction between a whole complex of variables. Are we justified in giving a gross description (for example, "traditional") to a series of classrooms in which a whole constellation of teacher, learner, and environmental variables are interacting? Is our experimental set of classrooms sufficiently homogeneous and independent to be able to be distinguished from a second set of control classrooms? Siegel and Siegel (1967) consider that "it is precisely these kinds of interactions that we suspect are partly responsible for producing cancellation effects in mean-performance comparisons between grossly described classroom groups". The huge volume of null findings which have stemmed from "methods" experiments in educational research lends weight to their hypothesis.

One further point needs to be made with respect to the Australian situation. We do not have extensive local curriculum development agencies able to provide completed and competing products for summative evaluation. It is true that some overseas products are available. But here it would seem to be even more important to undertake the kind of evaluation which allowed for modification and change. In short, whether we are concerned with the development of curricula in Australia or whether we prefer to modify curricula developed overseas, it is essential that we undertake the kind of evaluation which allows for explanations of differences and subsequent feedback and modification to the curricula. The rest of this paper will consider an extension of the basic model in Figure 1 to allow for this.

Extensions to the Model

Robert Stake (1967) suggests that the basic model from Figure 1 can be considered on three levels. The model suggested by Stake (with some modifications) is shown in Figure 2.

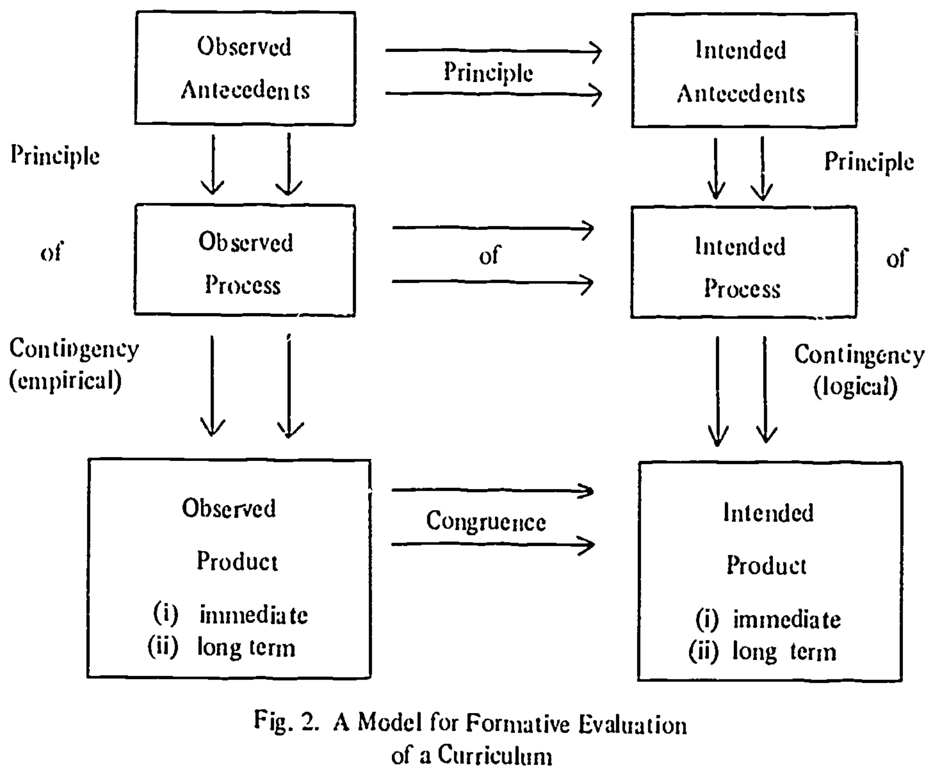


Fig. 2. A Model for Formative Evaluation of a Curriculum

The three levels in the model are labelled antecedents, process, and product.

Antecedents refer to any conditions existing prior to the teaching-learning process. Thus we include here prior knowledge, abilities, motivations (in general, "entering behaviours") of students, equipment and aids available, attitudes of teachers and administrators, prior training of teachers and so on.

Process refers to what happens in the classrooms, laboratories, libraries, or at home, which directly relate to the curriculum. It includes such things as the time spent on different sections, the kinds of questions teachers ask, what students do in the laboratory and so on. That is, not only *what* is taught, but *how* it is taught is an important part of process.

Product refers to outcomes. It includes, for example, pupil achievement, attitudes, cognitive style, the numbers of books read, the quantity of broken apparatus, and the attitudes acquired by teachers. Product might also be considered as "immediate" and "long term".

In this kind of model a number of relationships are apparent. Firstly, the principle of CONGRUENCE still applies. That is, at each level, it is possible to look for correspondence between the things "intended" to happen and the "observations" made of what actually happened. The observation data will be collected using a variety of instruments to make the measurements. They will include achievement tests, psychological tests, interviews, direct observation techniques, check lists, opinionnaires, and inventories.

When observations have been made at each level, it is then possible to describe the extent of congruence between "intentions" and "observations". For example, what equipment was intended for a particular experiment? Was it available? What kinds of questions were teachers expected to ask? What did a classification of teacher questions reveal? What pupil achievement was expected for a particular section? What was found?

A second kind of relationship is also present in this model and this has been described by a principle of CONTINGENCY. This is a type of cause and effect relationship applying between the vertical levels in the model. For the intentions, the contingencies are "logical". That is, students need to know X in order to do Y in order to learn Z. Or apparatus A is needed to perform experiment B so students will learn C. Or apparatus should be made of material P (say metal) so that when it receives treatment Q (say rough) it will finish up in condition R (say not broken).

A similar kind of contingency exists between levels of observations, in this case "empirical" contingency. Observations may be made or measurements taken of the extent to which students did know X initially, the extent to which they performed Y, and the extent to which they learnt Z.

It is the addition of this principle of contingency which gives the model a power of explanation or interpretation which is not present in a summative model. (Indeed, the summative model might be considered as a particular case of this more general model with only the third level (Products) included. (See Figure 3). Notice that the principle of congruence still applies but that the contingency relationship has been lost.

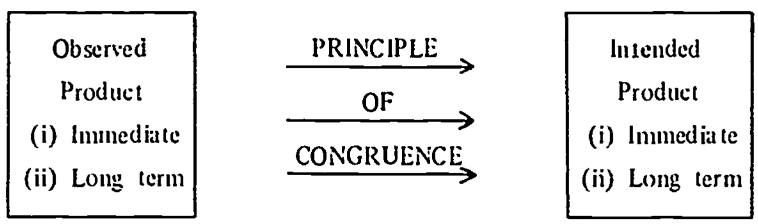


Figure 3. The summative evaluation model is a special case of the general model.

The Role of Judgment

Our model is now far more complex but is still rather a model for *description* of intents, observations, and the relationships between them than it is a model for curriculum evaluation. To complete the model, it is necessary to build in the factor of JUDGMENTS. When the extent of congruence between observations and intents has been established, a value judgment is still required as to the adequacy of the achieved outcome.

Finally, where the level of congruence is judged as inadequate, explanations should be sought through the principle of contingency.

It is not sufficient to know that a particular objective was not achieved. To correct the inadequacy of congruence between intention and observation, it is necessary to know what to correct -- the students' prior knowledge? Or the objective? Or the method of instruction? The model offers promise of this kind of analysis. That is, provided the contingency principle has been correctly applied, lack of congruence can be traced to its likely origin, and explained in terms of combinations of antecedent and process variables.

When such explanations are possible one final aspect of our model can be included. This is the role of FEEDBACK. (Figure 4.) Provided the observations are made systematically, the curriculum can become SELF-ADJUSTING. Corrections can be made and their effects can be similarly evaluated. Teachers can be encouraged to experiment and their experiments can be appraised. Adjustments can be made for changes in antecedents with a new intake of students. In short, the curriculum is now dynamic, not static.

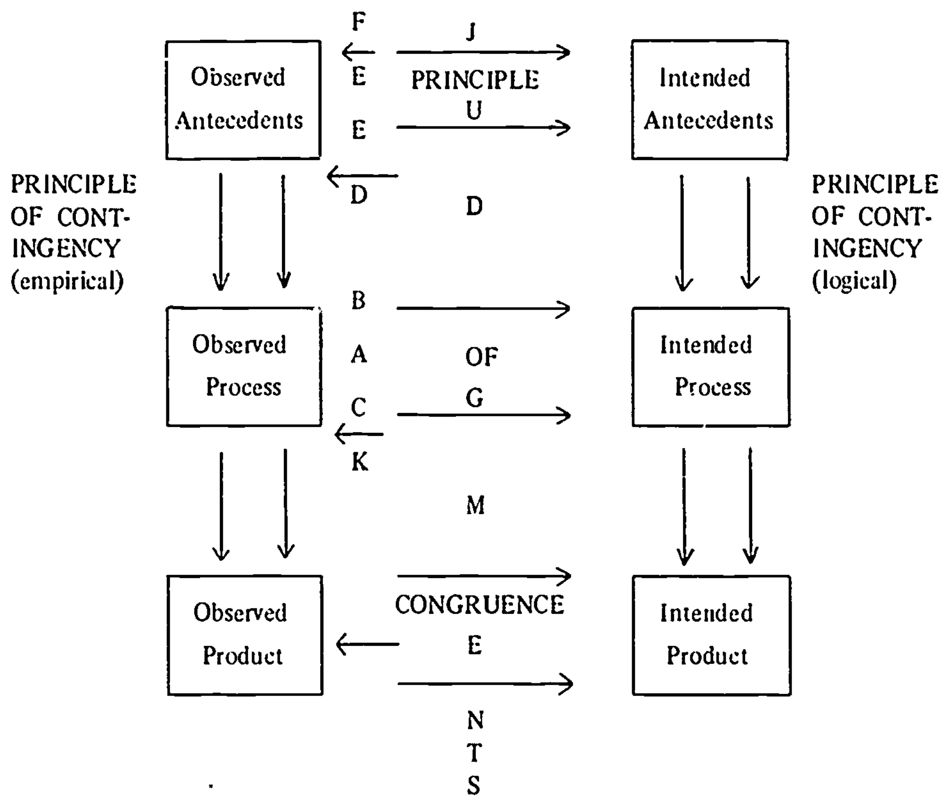


Figure 4. Judgments and feedback as part of the evaluation model

The Model Applied to the Evaluation of ASEP units

The Consultants' Report by the Macquarie team considered the application of this model to the national trials evaluation. The following is an extract from that Report:

> "It seems fair to assume that the major task of this trial is one of formative evaluation. Apart from the desire to improve the materials before final printing, the practical requirement of having to trial separate units imposes restrictions on the kinds of long-term outcomes which might be sought from summative evaluation.
>
> As an attempt at formative evaluation then, the use of control groups is less important than the classification of intentions or objectives. Also, if this kind of model is to be applied, observations or measures need to be made at each of the three levels. This means that pre-tests will be used, but the major purpose of such tests is to assess the entering behaviours of students and the preparation of

teachers and materials. The evaluation at this level then involves judgments about the extent to which these meet the intended entering behaviours or preparation. By contrast, in a summative model, major interest would focus on changes in behaviour from pre-test to post-test, or pre-test scores might be used as a covariance adjustment for post-test scores. While this allows a judgment of overall adequacy or success, it does not supply the diagnostic analysis which is at the heart of formative evaluation.

The formative model must also be prepared to observe or measure process variables; that is, an attempt must be made to find out what really happens in ASEP classrooms. While the cost of systematic observation by trained observers, or of video recordings, would be quite prohibitive both in terms of time and money, it is hoped that some of the techniques already employed in first trials could again be used. Some of the evaluation sheets have attempted to get at process variables. The use of the time-lapse camera also offers exciting possibilities. Perhaps a log book (or wall chart as previously used) would also be helpful here.

Finally the product measures (post-tests) should reflect the ASEP objectives for the particular unit under trial. It would seem that pencil-and-paper instruments will have to carry the major burden of measurement at this level. However, some of the objectives are such that these instruments will require ingenious preparation and careful validation. For students, the instruments will be concerned with a range of cognitive achievements, as well as attitudes and interests. However, the attitudes of teachers and administrators and the welfare of the materials should be considered as additional outcome measures."

An attempt was also made in the Report to classify the kinds of variables which could be measured in the evaluation. Table 1 provides this classification with sub-headings for pupils, teachers, and materials.

TABLE 1. Some Variables for Study in ASEP Evaluation *

Stage of Model Classification of variables	Antecedents	Process	Product
Pupils	1. Characteristics – age – sex 2. Ability level – intelligence – Piaget stage 3. Knowledge of unit 4. Skills – social – communication 5. Attitudes 6. Interests	1. Time spent on different activities – individually – in group – at home 2. Options worked	1. Achievement of range of cognitive skills 2. Skills – – – 3. Attitudes 4. Interests
Teachers	1. Characteristics – age – sex 2. History – training – professional activity 3. Knowledge of ASEP unit 4. Attitudes to science, science teaching	1. Organization of class activities 2. Time allocation when class working on activities 3. Interaction with pupils 4. How unit introduced	1. Attitudes to unit 2. Attitudes to material (science) within the unit
Materials	Covered by local evaluation	Observations of use of materials, films etc.	1. Ratings of materials by teachers 2. By pupils

* Extracted from Consultants' Report by Macquarie team who visited ASEP.

Implications for Cooperative Research on ASEP Evaluation

It is feasible that a number of institutions could co-operate in an evaluation of classrooms using ASEP materials, particularly since these materials will be used in several Australian States. It is also feasible that, if a model similar to the one detailed in this paper was used, different institutions might focus on different levels of the model. This would enable researchers to capitalize on their particular interests. Thus, for example, those currently engaged in classroom research might focus on process variables while those interested in test development might focus on product variables.

However, what is important is that variables from one level of the model bear a one-to-one correspondence to variables from other levels. This has obvious implications for the ways in which variables are conceptualized and the ways in which they are defined operationally. Failure to meet this requirement would mean that the principle of contingency within the model could not be applied. To ensure contingency between corresponding variables in different levels of the model, prior planning by cooperating researchers in relation to these factors would be necessary.

Another implication stemming from the model is that relevant data collected at all levels of the model would need to be available to researchers working at any particular level. The strength of the model as an interpretive framework is dependent on the ability to invoke measures from all levels to explain lack of congruence. This, then, would require cooperation in terms of the exchange of data and has additional implications for formats in which data might be collected and stored.

BIBLIOGRAPHY:

Australian Science Education Project. *Report of the Guidelines Conference*, Monash University, January, 1970.

Baumgart, N.L. "The discovery approach and educational set", *Australian Science Teachers' Journal*, 43, Vol. 15, No. 1, May 1969, 41-50.

Cronbach, Lee J. "Evaluation for course improvement", *Teachers College Record*, 1963, 64, 672-683.

Lucas, A.M. "The effect of teaching for content-free objectives in BSCS-type biology", *Australian Science Teachers' Journal*, 43, Vol. 15, No. 1, May, 1969, 51-57.

Mackay, L.D. "Changes on Cognitive Objectives During Two Years of Physics Study in Victorian Schools". *Australian Science Teachers' Journal*, Vol. 17, No. 3, October, 1971. 74-79.

Scriven, M. "The methodology of evaluation", *AERA Monograph Series on Curriculum Evaluation*, No. 1 Rand McNally, 1967, 39-89.

Siegel, Laurence and Siegel, Lila Corkland. "The instructional gestalt" in Siegel, Laurence (ed.), *Instruction: Some Contemporary View Points.* Chandler, 1967.

Stake, R. "The countenance of educational evaluation, *Teachers College Record*, 68, April, 1967, 523-540.

Williams, J. "The curriculum: some patterns of development and designs for evaluation" in Butcher, H.J. (ed.), *Educational Research in Britain,* University of London Press, 1968.

EDITOR'S COMMENT

During the discussion session following this paper several salient points were made. They included the suggestion that "expected outcomes" or "expected products" might well be an appropriate feature to be included in the evaluation model. The term "expected outcomes" was used to distinguish them from the "intended products" referred to in figure 2. Information about expected outcomes can be obtained from parents, employers, politicians and pupils and the congruence between observed and intended products and expected outcomes assessed. It was averred that any process of curriculum evaluation should take account of the "expectations" of persons and groups other than the curriculum designers.

THE AUSTRALIAN SCIENCE EDUCATION PROJECT AS A STIMULUS FOR RESEARCH – A PROGRESS REPORT OF A STUDY OF TEACHING STRATEGIES USED WITH A.S.E.P. MATERIALS

C.N. Power

In their book "The Academic Revolution", Jencks and Riesman (1968) claim that a large proportion of research "exhibits no genuine concern with answering real questions or solving important problems; it is simply a display of professional narcissism. Instead of recording a struggle with methodological and substantive issues that actually matter to either teachers or students, it is simply a roller-coaster ride along a well worn track." It is with some sadness that one must admit that roller-coaster responses to the higher degree stimulus have generated a good deal of useless research. If the Australian Science Education Research Association (ASERA) provides no more than an opportunity for us to become absorbed in the beauty and precision of our own research efforts, we can hardly expect to convince science teachers that systematic research and disciplined intellectual effort can help them solve their problems.

Unfortunately research workers rarely are rewarded by their colleagues for dirtying their hands by working *with* teachers and students on their problems, but rather for working *on* them for some esoteric purpose. Little wonder that research has contributed little to educational theory and had even less impact on practice. This is not to deny the need for "pure" research, but rather an attack on the sterility of "pure" studies which ignore (a) what goes on inside the "black boxes" of the classroom and the school, and (b) the complexity of persons and their environments.

If research is to contribute to the practice of teaching it must provide information on the consequences flowing from the use of particular teaching strategies in particular situations. It follows that multivariate naturalistic studies of the effects of different strategies used by teachers on various types of pupils are needed, along with multivariate experimental studies which explore impact of new combinations of curriculum materials and strategies. However, the existing evidence suggests that no one strategy or combination can be expected to be equally effective for all teachers, with all students for all time. There is a need for research studies in which the teachers themselves form part of an action research team engaged in finding the types of combinations and adaptations which work for them and their students.

In this paper, the progress made in a study concerned with the effects of teaching strategies in micro and mini teaching situations using A.S.E.P. materials is reported. The study does attempt to incorporate some of the features described in the foregoing paragraph.

Origins of the Project

In the last decade much research has been concerned with the effects of teacher behaviour, with developing and evaluating micro-teaching techniques, and with curriculum development. Rarely have these three interests been combined in a teacher education program. To the investigators, the advantages of combining these interests and using them in a teacher education program seemed to be many and include the development of trainees' understanding of new curricula and the provision of opportunities to practice a variety of teaching strategies. Furthermore, the study of the effects of teaching strategies in a micro-teaching situation allows for greater control of the strategies, while being not too far removed from the "real world".

Of the new curricula in science, few provide a greater stimulus for research into teaching strategies than self-paced, semi-individualised courses like A.S.E.P. The production of these materials has focussed attention on an enormous number of ideas and questions which can be taken up by those interested in research. The materials themselves can serve as a research stimulus. Ramsey (1971) has already provided a list of formative evaluation research questions and Moritz will undoubtedly add more. As well as providing a stimulus for curriculum research, A.S.E.P. provides a stimulus for those interested in research into classroom interaction, teaching strategies and teacher education. The bulk of the research in these areas has been based on the assumption that the traditional modes of classroom instruction will persist. However, there can be little doubt that the rules which govern the behaviour of A.S.E.P. teachers and pupils are quite different from those operating in conventional science classrooms. It also follows that the results of research concerned with the effects of teacher and pupil classroom behaviour on achievement is largely irrelevant. There is a need for descriptive studies of teacher-pupil-material interactions in A.S.E.P.-type classrooms to provide answers to the question, "What is the nature of the interactions occurring in these classrooms?" Then we may go on to ask: How are these behaviours related to antecedents, to outcomes? Which teaching and learning strategies maximize the learning of various types of students? How effective are attempts to train teachers to use these strategies? Does it make any difference if the teachers have values similar to those of A.S.E.P. staff? Does it make any difference if the materials are tightly structured or relatively unstructured? There are many other research questions relating to teaching, curriculum evaluation, testing, sequencing of activities, interpersonal relations, and teacher education which A.S.E.P. has raised simply because it has dared to break with some traditions but not with others.

The foregoing elements, coupled with a concern for the quality of living and learning in science classrooms, provided the motivation for the project; the AACRDE provided the funds. The purposes of the project were as follows –

(a) to train student-teachers and teachers in the use of selected A.S.E.P. materials,

(b) to develop criterion measures for selected A.S.E.P. sequences,

(c) to study the effects of (A.S.E.P.) teaching strategies on pupils in mini and conventional teaching situations.

(d) to develop a simplified classroom behaviour classificatory system for teachers and student teachers.

(e) using the system, to provide feedback to teachers and student-teachers on the nature of their teaching behaviour.

Aims (a), (d) and (e) relate to the training and action research functions of the project, and aims (b) and (c) relate to the pure research function.

It is appropriate to state that the study of the effects of teaching strategies was, and is, an essential feature of the proposed study. As indicated earlier, much remains to be discovered about the teaching-learning process as it occurs in classrooms. Because classrooms are complex places and the research models guiding classroom research have been inadequate to the task, we have been largely unable to state what happens to pupils of type X when teachers behave in manner Y. One difficulty has been the variability of content and of strategies from one lesson to another. The use of A.S.E.P. materials and of micro-teaching situations introduces an element of control. Consequently the association between the teaching strategies and outcomes will be more amendable to study. Furthermore, the effects of newer, varying combinations of "strategies" for given content will be open for study and the results, it is hoped, will contribute to theory. At the same time, the Diploma in Education students involved in the project ought to become sensitive to a variety of teaching strategies and their effects. Also, they will use the procedures in schools, thus influencing educational practice.

At the time the proposal was put forward, the investigators were particularly interested in following up questions and hunches about teaching strategies generated out of their own research and experience with micro-teaching. However, it seemed critical that they should take into account the questions which A.S.E.P. staff and trials teachers were asking. Accordingly in February, they began to share their ideas with trials teachers and visited A.S.E.P. to discuss the project with project staff. In every case, the individuals approached gave unstintingly of their time and the investigators were most appreciative of the assistance and constructive criticism given.

Interestingly enough the issues raised by A.S.E.P. staff and teachers in relation to teaching strategies appropriate in self-paced, semi-individualized programs proved to be surprisingly close to those of interest to the investigators. These included questions about the degree of structure which was appropriate, the diagnosing of pupil difficulties; the nature, frequency and informativeness of teacher feedback; the sequencing of teaching strategies; and organization.

A great many science teachers expressed tremendous interest in the project and were keen to work on what were to them real problems. In addition, A.S.E.P. staff were very much concerned with the impact of A.S.E.P. on existing teachers and with the effectiveness of inservice as well as pre-service training. The study will, hopefully, provide answers to real questions.

Design

The basic design used in this study is a pre-test/post-test control group design. The research involves two phases, one involving student teachers (Diploma in Education students) and one involving practising teachers. The phases are as follows:

(a) *Effectiveness of teaching strategies employed by student teachers in mini-teaching situations using A.S.E.P. materials*

In this phase of the research the effects of three conditions will be explored. The conditions are as follows –

1. *Degree of Structure:* In half of the 8 mini-teaching classes (i.e. classes 10-15 students), a highly structured A.S.E.P. unit and teaching strategy will be employed. In the other half, a relatively unstructured version of the same unit will be employed.

2. *Individualization:* In half of the classes, students will work alone at their own pace. In the other half, students will work in small groups of 2 or 3 at the same pace.

3. *Unit:* Half of the classes will use one A.S.E.P. unit, and half another unit. The . its selected are "light forms images" and "Pigments and Acidity".

In this phase of the project, eight Diploma in Education students will serve as teachers. Students in the mini-classes will be required to complete a short personality test based on H.S.P.Q., and pre and post criterion-referenced science tests. Classroom interaction data will also be collected, that is, educational inputs, outputs and operations will be measured. This data will be used in the evaluation of teaching strategies and the A.S.E.P. units as well as in hypothesis testing (Astin and Panos, 1971).

(b) *Effectiveness of teaching strategies employed by teachers using A.S.E.P. materials*

This part of the study is concerned with evaluating the effectiveness of inservice education with teachers "committed" and "uncommitted" to the approach adopted by A.S.E.P., as well as the effectiveness of the strategies used by these teachers with different "types" of students identified in a previous study (Power, 1971). Essentially, the conditions are as follows –

1. *Degree of Commitment:* Of the eight teachers selected, half will include those who react favourably to the A.S.E.P. materials to be used, who have educational values and practices similar to those of A.S.E.P. staff. Other teachers selected will include those who are less congruent with A.S.E.P. values.

2. *Training:* Half of each of the above groups will be trained to use A.S.E.P. materials and to use teaching strategies appropriate to "individualized" instruction. The other four teachers will be untrained.

3. *Unit:* Teachers will use one of two A.S.E.P. core sections of the units mentioned above.

Progress, Problems and Proposals

At present the project is in its early phases. The major data collection period for practicing teachers is July and for Diploma in Education students, September. To date, the investigators have administered the A.S.E.P. attitude scale (27 items) and a teaching practices questionnaire (30 items) to science teachers in twelve schools, to our Science Diploma students, and to A.S.E.P. staff (N = 157). Using this data we hope to map the values and practices of those sampled. The initial intention is to identify those "committed" and "uncommitted" to A.S.E.P. values. However, this is perhaps an artificial classification. A sounder approach might be to systematically classify all individuals and to determine how the clusters identified do, in fact, differ. The procedure being followed is represented below.

COMPONENT ANALYSIS
(determine dimensions
measured by attitude scale
and practices questionnaire)

→

H GROUP
(clustering of individuals
in sample)

↓

DISCRIMINANT ANALYSIS
(determine how clusters differ)

Having determined the component factors measured by the attitude scale and practices questionnaire, a clustering procedure (H GROUP) will be used whereby individuals with similar values (as represented by factor scores) are grouped together. The precise nature of the differences among the groups thus identified can be determined by discriminant analysis techniques. For a description of these procedures and of the computer programs used, the reader is referred to Veldman (1967).

In the principal component analysis of the attitude scale and the practices questionnaire, a total of 18 factors, accounting for 69.6% trace were extracted (an eigenvalue cutoff of 1.0 was used). The first five of these factors were as follows:

FACTOR 1 LABORATORY PROBLEM CENTRED APPROACH

8.3 Emphasis on pupils devising and conducting their own experiments (.78)
9.3 Emphasis on first hand experience with phenomena (.72)
7.3 Emphasis on problem-centred approach (.71)
3.3 Emphasis on different activities carried out by different groups (.61)

FACTOR 2 SUBJECT MATTER EMPHASIS

21 Science course produced by A.S.E.P. should be separated into major branches of science (.36)
15 An understanding of interpersonal relationships has a place in the science course (– .71)
16 The development of such personal attributes as cooperation, persistence . . . can be fostered by a suitable science curriculum (– .70)
4 As one of its major objectives, the science course should aim to contribute to the personal and social development of the child (– .64)

FACTOR 3 EMPHASIS ON TEACHER RULES

4.1 Emphasis on rules established by the teacher (.72)
5.1 Emphasis on rules stating how pupils must behave (.71)
1.3 Emphasis by teacher on disciplining and controlling pupils (.54)
6.3 Emphasis on effective use of punishments (.47)

FACTOR 4 EMPHASIS ON CONCEPTUAL SCHEMES

1 There is a fundamental core of scientific knowledge which all students should know, and which should form the basis of the science course (.74)
14 It is of great importance that the science course deals with the main theories of science (.72)
2 The science course should reflect the structure of the major scientific disciplines (.61)

FACTOR 5 LECTURE – DEMONSTRATION

27 Pupils gain little of value from carrying out their own investigations (.72)
26 A teacher's time is better employed in giving lectures and demonstrations than in preparing for laboratory work (.63)
17 Knowledge which is most relevant to the everyday life of students should be given preference in the course (.40)

Having determined if and how teachers differ in attitude and practices from one another and from A.S.E.P. staff, the researchers can proceed to select teachers representative of different orientations towards teaching for the first phase of the project and to train half of these. For training purposes a 20 minute videotape has been compiled. It is designed to highlight some problems facing the A.S.E.P. teacher and to focus on teacher strategies for organising equipment and for acting as a guid counsellor and trouble-shooter. In addition, work is progressing in a self-paced, semi-individualized "unipak" for use in a training program.

In July, it is planned to administer a criterion pre-test, a personality test, and an ability measure to selected classes; to videotape three lessons for each class; and to administer a criterion-referenced achievement test at the completion of a short teaching sequence. The data consist of multivariate measures of antecedent variables, classroom variables and outcome variables and will provide information on the impact of A.S.E.P. on the behaviour of teachers who differ in training and orientation, and of the impact of teachers, teaching strategies and A.S.E.P. materials on different "types" of students.

One problem, not yet fully resolved, is that of collecting and analysing classroom data in A.S.E.P. classrooms. Preliminary attempts to record and classify teacher-pupil-material interactions in A.S.E.P. classrooms have been complex because of diversity of activity and interactions; the difficulty of obtaining an audio-record which is decipherable; and the shortcomings of existing schemes for classifying verbal interaction in settings other than the "conventional". If changes in behaviour (learning) occur when input is received from the classroom environment and if the input interacts with a student's cognitive, personality and motivational systems – it becomes important that to obtain a representative sample of the direct and vicarious input received by pupils. How much the student learns may depend on the nature of the questions asked by the materials and the teacher, the responses made and the feedback received. Learning may also depend on the sequence of activities in which pupils are engaged and on how the teacher structures the learning environment. Of course it is impossible to obtain a meaningful record of every teacher and pupil action. One possible plan is to focus one camera on the teacher and to use a radio controlled microphone to record his talk. A second camera could be used to systematically sample the behaviour of each pupil group during the 40 minute period, by moving the camera and a microphone from one group to the next according to a pre-determined schedule. In addition to this dynamic videotape data, student record books, student diaries and responses to the classroom activities questionnaire can provide valuable data relating to what pupils and teachers do in A.S.E.P. classrooms.

One final problem is that of unstructuring the fairly structured cores of the selected A.S.E.P. units – "Light Forms Images" and "Pigments and Acidity". Instructional design has emphasized adult defined behavioural objectives, careful selection and sequencing of activities, frequent testing and feedback. However, some newer curricula like the Environmental Studies Project and some A.S.E.P. units, reject behavioural objectives; rigid frame-works; semi-programmed, self-paced learning sequences designed by adults to lead to objectives adults hold to be important. These materials are more open-ended and allow considerable freedom for pupils to explore those aspects of environment the pupils sees as important and relevant. The emphasis is on intrinsic rather than extrinsic evaluation. In other words, the pupils assume some of the responsibility for deciding what shall be learned, and how and how well it has been learned. However, "destructuring" may prove to be a matter of degree rather than of kind – and how does one measure the extent to which materials have been destructured, and the extent to which teachers have provided structure?

Concluding Comment

In the introduction to this paper it was implied that research might be more productive if investigators work closely with teachers on mutual problems and provide opportunities for comment and attack by being open about the struggles in which they are engaged. Hopefully, this study meets these requirements.

REFERENCES

Astin, A.W. and Panos, R.J. The evaluation of educational programs. In Thorndike, R.L. (Ed.) *Educational Measurement* American Council on Education, 1971.

Jencks, C. and Riesman, D. *The Academic Revolution* Garden City, N.Y. Doubleday, 1968. p. 535.

Power, C.N. The effects of communication patterns on student sociometric status, attitudes and achievement. Unpublished Ph.D. Thesis, University of Queensland, 1971.

Ramsey, G.A. Formative evaluation and the Australian Science Education Project. In Tisher, R. (Ed.) *Research 1971*. Australian Science Education Research Association.

Veldman, D. *Fortran Programming for the Behavioural Science.* Holt, Rinehart, Winston, 1967.

SOME SUGGESTIONS FOR RESEARCH RELATED TO ASEP

K.W. Moritz

The following suggestions for possible research tasks relate, in the main, to those areas of research which are seen by members of ASEP staff to be of interest and value to a curriculum development project.

The suggestions range from those for quite minor pieces of research to those which would require major resources. No attempt has been made to assess the feasibility of any of the research tasks.

The accompanying chart, which gives a summary of ASEP's evaluation procedures, is a convenient frame-work for considering research possibilities related to ASEP. The chart divides the topics of evaluative concern into four areas -- materials, students, teachers and the instructional process -- and in what follows possible areas for research are considered under these headings. The chart shows the emphasis which ASEP is obliged to place on its formative evaluation of the materials it is producing, and, not unexpectedly, this emphasis is reflected in the number of research tasks suggested under the headings of Materials. By the same token, the chart also points up the extent to which ASEP is restricted in the amount of effort it is able to expend in gathering and investigating data concerning students, teachers, and the instructional process, where such data do not bear directly on the task of evaluating the materials. The chart also indicates some of the techniques of evaluation used by ASEP, and lists some of the instruments used. It may be that consideration of these features of the chart by those working outside ASEP could lead to ideas for research not mentioned below.

Some of the research suggestions listed have been made with some feeling. Often we in the Project have keenly felt the absence of reliable research data which could have given us direction or guidance. For example, the readability of our materials is a matter of prime importance to us, and we are making considerable efforts to ensure that the students for whom the materials are intended have minimum difficulty in understanding what is written. The principal way in which we try to achieve this is by making the material conform to the appropriate grade level by using a Flesch readability formula and Flesch's grade ratings. Flesch devised his formula and suggested his grade ratings for American conditions some thirty years ago. How valid are they for Australian students in the 1970s?

Again, various stages of mental development in children have been described by Piaget, and ASEP materials are designed to cater for students at three different stage levels concrete, transitional and the formal thinking stage. Obviously it is important in using ASEP materials for teachers to be able to decide the stage of mental development which their students have reached, but how is this to be done? What kind of test would best enable a teacher to do this?

A third example of a deeply-felt need for more research data occurs in the area of attitude measurement. ASEP has undertaken to try and develop certain attitudes in students, and regards this aim as a most important one. What is the best procedure to adopt in an attempt to find out if the materials are in fact developing these attitudes in students? Can this effectively be done during trials, when of necessity no classes are exposed to ASEP materials for any protracted period of time, or must this wait for summative evaluation, when the final ASEP product is in the schools?

As has already been indicated, most of the research tasks suggested are those whose findings would be of direct value to ASEP staff in their present task of producing materials for science in junior secondary classes, and in the formative evaluation of these materials. Unfortunately any research begun now would be most unlikely to produce results which would be of value to ASEP in either of these respects. However, research begun now using ASEP as a framework would almost certainly be of value to any future curriculum development project, and perhaps in this regard this tentative list of suggestions may prove of some value.*

In what follows, 'suggestions' is the operative word. Additional details of what Project staff have in mind may be obtained by approaching the Project direct.

List of Research on Materials

Readability

The validity of readability measures when used to rate science materials, as distinct from other kinds of material.

The validity of Flesch ratings as applied to materials for Australian students in the 1970s.

The effect on students of materials which are written at a relatively high level of reading difficulty. Does this lead to improvement in reading, or to frustration and lack of interest? Does it depend on circumstances? If so, which ones?

How interest, readability, understanding, fatigue are affected by

(i) size of book

(ii) size and kind of type

(iii) space

(iv) cartoons

(v) activity frames.

Effect of material written at a low level of reading difficulty on students who are superior readers.

ASEP's system of writing materials compared with other systems. (To illustrate: in ASEP, writers are drawn from the ranks of teachers; what they write is guided and evaluated by experts in education and science. In another well-known system, initial manuscripts written by scientists and other subject-matter experts are rewritten by teachers and others for use by students.)

The effectiveness and usefulness of wall charts.

Analysis of content

Analysis of ASEP materials to find out if they are 'activity-based', 'inquiry-oriented', 'student-centred', etc., and to compare them with other materials.

The extent to which the materials appear to reflect ASEP's aims. For example, to what extent do the materials overtly aim at giving students an understanding of the nature, scope and limitations of science?

The extent to which the materials are 'structured'. Should this depend on the stage of development of the student for whom the materials are intended?

Teaching of key concepts

How can concepts such as 'control', 'model', 'operational definition' best be taught?

How can the understanding of such concepts best be tested?

Tests

('Diagnostic tests' are those provided in the units as an integral part of the instructional procedure. 'Pre-post-tests' are those given to experimental and control classes both before and after a unit has been done. The latter are designed solely to provide feedback to ASEP.)

Design of performance test items for diagnostic and pre-post-tests.

Validation of pencil-and-paper test items by performance or other tests.

Design of test items to measure attitude changes. Can attitude changes be expected within the space of time taken to trial a unit? Can attitudes be categorized as lists of observable behaviours, and checklists compiled to measure students' attainment of attitudes?

Design of pre-tests where material tested involves a definite hierarchy, or a key concept. For example, a series of test questions may involve an understanding of some concept. If, on the first question, a student shows that he is quite unfamiliar with the concept, further questions on such a concept may be unnecessary.

Design of pre-tests to avoid sensitization of students to the material of certain units.

ASEP does not supply tests designed to allow the teacher to grade his students. What effects does this have?

Remedial loops

The kind of remedial loops, if any, which should be provided as part of the diagnostic test material.

The extent to which the diagnostic tests as presently constituted are instrumental in helping students to achieve the objectives of a unit.

Record books

The justification for student record books. (Should students write things down?)

Objectives

The form in which objectives of a unit are best stated. Should students be informed of the objectives they are supposed to achieve?

List of Research on Students

Student characteristics

Relationship between a student's subject preferences and his reaction to an ASEP unit.

Relationship between a student's preferences for various activities (science and non-science) and his reactions to an ASEP unit.

Differences between students of different age groups and different school systems with respect to

(i) preferred ways of tackling experiments

(ii) preferred sizes of groups in which to work in a science class.

Characteristics of students who cope well in an ASEP class compared with characteristics of students who do not.

Relationship between students' interest in the materials and their achievement of the objectives.

Entry behaviours necessary for success in ASEP units e.g. ability to read scales, graphical skills, mathematical skills.

Tests

Tests to determine the Piagetian stage level of students.

List of Research on Teachers

Teacher characteristics

Comparison of characteristics of ASEP trial teachers compared with characteristics of teachers as a whole.

The kinds of personality traits and other attributes which best equip a teacher to deal with ASEP material.

Comparison of a trials teacher's self-image with the assessment of the teacher by visiting Project staff.

Teachers and ASEP

The effect ASEP workshops have on the teacher's handling of an ASEP unit in his class.

The differential effect of ASEP materials on younger and older teachers, qualified and unqualified teachers, etc.

List of Research on the Instructional Process

The nature of interactions which go on in an ASEP class. How these interactions affect the learning process. Does ASEP material foster any special types of interaction?

The relationship between students' assessments of a unit and teacher variables.

The formation and dissolution of groups of students.

The teacher's roles in respect of different-sized groups.

The differences in the effectiveness of ASEP units as between large and small classes.

Other Areas for Research

The administration and organization of ASEP as one example of a curriculum development project.

The effects on students of learning sequences in ASEP materials.

The nature and utilization of evaluation feedback.

Analysis of interaction in the ASEP classroom.

Teaching strategies in the use of ASEP materials.

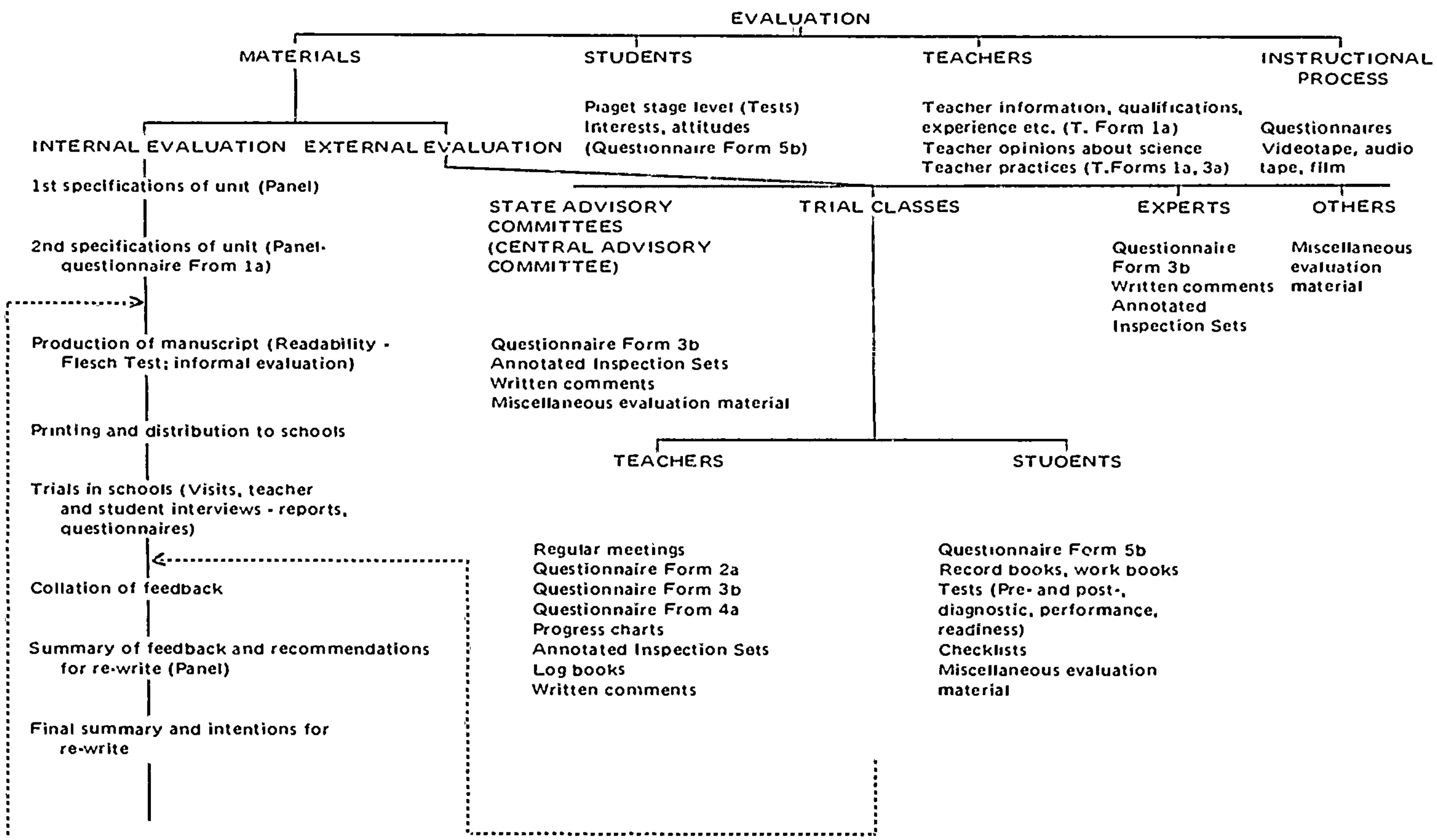
Long-term effects of using ASEP materials.

Investigation of micro-teaching situations using the ASEP class.

Measurement of student progress through ASEP units.

Summative evaluation of the final ASEP product.

SUMMARY OF ASEP'S EVALUATION PROCEDURES



A CHECKLIST FOR ANALYSING THE 'STYLE' OF INSTRUCTIONAL MATERIALS

Editor's Comment:

This report is based on work undertaken by S.R. Shepherd of the staff of the Australian Science Education Project. Mr. Shepherd described his exploratory study during the A.S.E.R.A. Conference and the comments which follow are adaptations of his presentation and the paper he tabled. It is appropriate to point out that the project is an embryonic but, nevertheless, significant one. The study is reported here in the hope that other workers will co-operate with and build upon the work begun by Shepherd.

* * * * * * * * * *

Introduction

The materials that are being produced as part of the activities of ASEP will be classified in terms of

1. topics, i.e. the information presented,
2. techniques, i.e. procedures for which instructions are given,
3. degrees of prescription, i.e. whether pupils' tasks are "open-ended" – or structured,
4. group size, i.e. number of persons in a group for each pupil activity,
5. ancillary materials required, i.e. additional equipment, chemicals, references and audio-visual media that are required;

and 6. style of written material.

In an attempt to describe the "styles" or "approaches" used in the written ASEP material, a checklist was devised. This list, which is reported here, will hopefully provide some details about ASEP units and will allow comparisons to be made with other instructional materials which have characteristics somewhat similar to those of ASEP. What, then, are the relevant characteristics which will provide guidelines for the establishment of a checklist for style?

Characteristics of ASEP Materials

The main characteristics of the ASEP materials for pupils are that:

1. they are 'self explanatory', that is, all the information and instructions that pupils need are contained in the pupils' books. The ASEP materials do not necessarily rely on teachers or other reference sources to provide information and instructions.

2. printed words carry the main 'messages'. There are illustrations and other materials but these support, rather than replace, the printed words.

3. short sentences are used to increase the "reading ease" of the materials (Flesch, 1968). Most of the sentences contain 10 to 15 words and are "simple", i.e., they contain a single statement.

The Checklist

The sentences in the A.S.E.P. materials were taken as the basic unit of analysis and were classified into three broad groups designated:

statements,

questions, and

instructions and suggestions.

Statements were further subdivided into those providing *information* and those dealing with *explanation*. In a similar manner the category, "instructions and suggestions" was subdivided so that distinctions could be made as to whether these instructions were concerned with *equipment, handling things, observation, writing, discussion* and *reading*. The questions category was not subdivided, but, perhaps in subsequent classifications it may be appropriate to distinguish between recall and other types of questions. The various categories and the subdivisions are shown in the checklist in the Appendix to this paper.

Advantages of the checklist are that (a) it can be mastered quickly, (b) it can be scored readily and (c) it is quite reliable. Certainly some results indicate that different coders obtain a high measure of agreement when classifying similar materials. An example of such a reliability check appears in the table below. The data were obtained by 2 teachers, 2 clerical assistants and the ASEP staff writer when they classified the sentences in the ASEP unit, "Electric Circuits".

The results indicate that the scorers were fairly consistent in their classification for most types of sentences. Greatest discrepancies occurred for the categories designated 'explanation', 'discuss' and 'read'. It is proposed to produce a guide sheet for scorers so that greater consistency can be obtained on these few categories in the future.

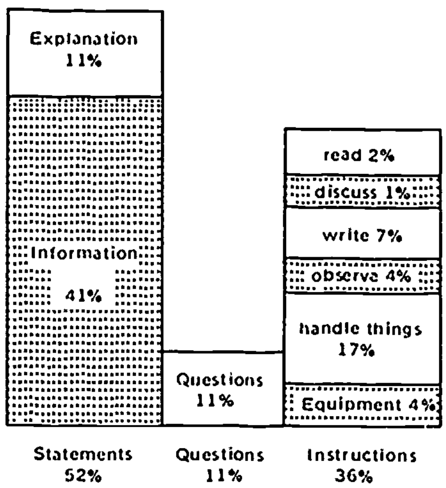
Type of Sentence	Number of sentences classified by SCORER					Mean	Standard deviation as a % of mean score
	JR	PC	ES	GH	SRS		
information	297	348	308	274	298	305	8%
explanation	122	61	111	127	128	110	23%
TOTAL STATEMENTS	419	409	419	401	426	415	2%
TOTAL QUESTIONS	101	95	95	110	98	100	5%
equipment	19	16	18	19	18	18	6%
handle things	208	241	205	201	187	208	8%
observe	48	38	49	33	55	45	17%
write	81	80	86	90	81	84	5%
discuss	12	12	11	28	13	15	43%
read	10	5	28	25	20	18	49%
TOTAL INSTRUCTIONS	378	392	397	396	374	387	2%
Grand TOTAL	898	896	911	917	898	904	1%

Application to Other Materials

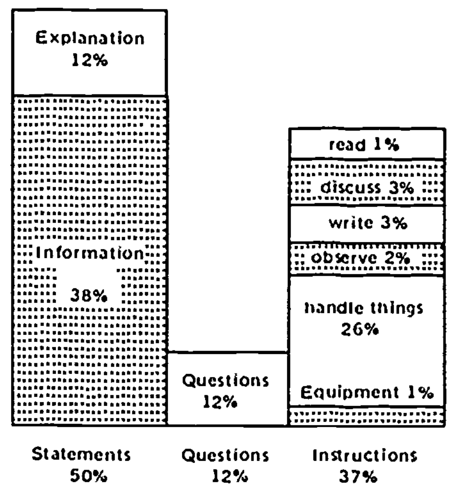
As a further application of the checklist a number of additional ASEP units, e.g., Life in Freshwater, Light Forms Images, and Mice and Men, sections of the Junior Secondary Science Project (J.S.S.P.), e.g., Green 5, *How Hot is it?* and Red 7, *How Mammals Function,* Chapters 1 to 4 in Volume 1 of the Intermediate Science Curriculum Study (I.S.C.S.) and Chapters 9, 18, 27, 36 and 45 of *Abridged Science for High School Students* (S.F.H.S.S.) were classified. The results are presented in the diagrams which follow.

It seems appropriate to note the diagram indicates that for A.S.E.P. materials the proportion of sentences in the different categories is remarkably similar to the proportions in the categories for the I.S.C.S. chapters. In both materials, for example, 11-12 per cent of all sentences are questions. Both I.S.C.S. and A.S.E.P. materials differ from J.S.S.P. in the number of times pupils are asked to write statements. Sixteen per cent of the sentences in J.S.S.P., compared to 7 per cent for A.S.E.P. contains an instruction to write. Another interesting observation is that, for the sections sampled, the project materials (A.S.E.P., J.S.S.P., and I.S.C.S.) present far more instructions than do traditional texts such as *Abridged science for high school students.*

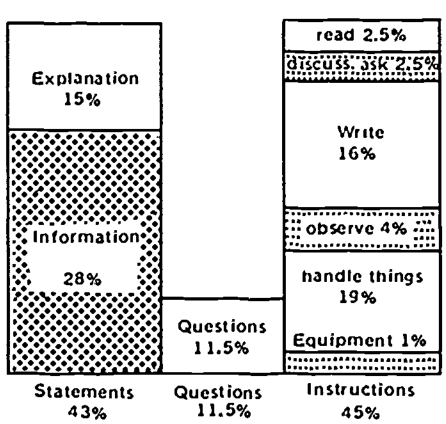
DIAGRAMS OF SENTENCE ANALYSIS



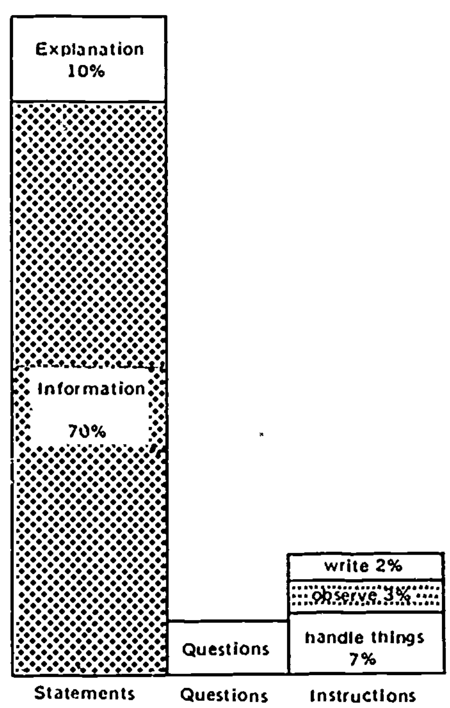
ASEP



ISCS



JSSP



SFHSS

Concluding Comments

Certainly, there seems to be a great similarity among the three project materials. If the claim that A.S.E.P. materials have a rather special or distinctive approach is to be substantiated, then additional research is required. This research could involve a more detailed and sophisticated analysis of the written materials and studies of the effects of A.S.E.P. on pupils.

Hopefully, the checklist described here will be modified and/or extended by other workers in studies of curriculum materials. If this report stimulates further research, then it has achieved one of its objectives.

REFERENCES

1. Australian Science Education Project, *Electric Circuits, Life in Freshwater, Light Forms Images,* and *Mice and Men.* National trial materials, Melbourne, 1972.

2. Intermediate Science Curriculum Study, *Probing the natural world,* Volume I, Morristown, New Jersey, Silver Burdett, 1970, Chapters 1 to 4.

3. Junior Secondary Science Project, Green 5 *How Hot is it?* and Red 7 *How mammals function,* Melbourne, Cheshire, 1968-69.

4. Flesch, R., *The art of plain talk,* New York, Collier books, 1968.

5. The Science Foundation for Physics School Certificate Integrated Science Textbook, Group of Authors and Editors (Chairman H. Messel), *Abridged science for high school students,* Volume 2, Sydney, Government Printer, Revised edition, 1970, Chapters 9, 18, 27, 36 and 45.

APPENDIX

Australian Science Education Project
<u>SENTENCE ANALYSIS OF UNITS</u>

UNIT:

VERSION: •local trial •national trial •final version •core •options •record book

Analyser: .... Date started ....

Rules for sentence analysis:

(1) Do not score the front cover or the copyright statement on the back cover.
(2) Include captions.
(3) Complex sentences can be scored twice e.g. handle + record.
(4) Regard colons or semi-colons as ends of sentences.

Type of sentence or wordgroup	Score	Total	%
INFORMATION			
EXPLANATION e.g. how to use the book, what the unit is about, introduction e.g. what you will be doing, reminders e.g. 'We have seen . . .' 'You have already found out . . .' Headings, contents lists			
QUESTIONS (other than headings)			
INSTRUCTIONS AND SUGGESTIONS Equipment lists			
handle things — includes . leave or store things . things you must not do . go to . . . . . collect, find, obtain			
observe, look at, notice, examine, measure, compare			
write, record, turn to your record book, draw, list, calculate, explain, describe (in writing)			
discuss, think about, ask your teacher. 'Suppose . . .'			
read, look up, check books, references listed.			
	Total		100%

REVIEW OF SOME RESEARCH IN SCIENCE EDUCATION

CLASSROOM INTERACTION ANALYSIS : THE NEW RELIGION

R.P. Tisher

"We are the interaction boys,
We analyze amid the noise
Of classroom struggles, though we die –
With failing hands we hold pens high . . . "

(Nicholas, 1968)

In the beginning

In this last decade educational research has been characterized by a new religion -- interaction analysis -- and many people regard Ned Flanders (1960) as its chief prophet. Be that as it may, he is but one of the many prophets and devotees to the new technique, and his particular scheme, is but one of over a hundred. Nevertheless, the Flanders "version" of interaction analysis has influenced many other workers, as will become evident in the ensuing discussion. That many persons have become converts to the new technique is indeed amazing. Perhaps, a more unfortunate feature is that many have embraced the religion with fervour, but without initial thought, and these fervent "thoughtless" individuals include science educationists. It is sad, too, to note that the revival movement after fifteen years has little to boast about: there have been no great leaps forward in knowledge, no revolutionary changes in educational practice, although there are now a multitude of classificatory schemes which seem to bear little relation to each other (Biddle, 1967; Meux, 1967). This state of affairs, among other things, seems to imply that this most ingenious technique, with its many commendable features, may be ruined by its own success (Mitchell, 1969). Certainly the technique is not a panacea, nor is it likely to be a bonanza for teacher education, as some writers claim (Campbell and Barnes, 1969). These remarks are not meant to imply that interaction analysis is of no value at all. On the contrary. However, they are intended to alert you to limitations in the technique and to guard against indiscriminate and thoughtless adoption. There is, in fact, much valuable work that science educationists can do with and for interaction analysis and hopefully this paper will, in addition to reviewing the state of the art, provide a fillip for new projects.

Interaction Analysis belongs to the study area of classroom research and it developed from the need and desire to discover what was happening in lessons and in classrooms. The term "interaction" implies an action-reaction or a two-way influence which may be between individuals (e.g. pupil-pupil, teacher-pupil, or teacher-target (Biddle, 1967), or between an individual and a group (e.g. teacher-audience) or between groups or between materials and individuals or groups. An interaction is usually inferred from the behaviour of persons in the environment being studied. This behaviour may be verbal or non-verbal and can be classified as being predominately cognitive, affective or controlling in nature. These categories do not exhaust all possible ones for the classification of classroom behaviours but they do indicate that interaction analysis is involved with many components and that classroom interaction is a multi-dimensional phenomenon with, undoubtedly, multi-dimensional effects (Mitchell, 1969). Unfortunately, classroom interaction has not frequently been treated as multi-dimensional and it is only in the last few years that the pleas for multivariate studies of classrooms have been heeded. One of these pioneer multivariate projects has recently been completed in Australia (Power, 1971). Generally, the focus in most interaction studies has been on verbal behaviour and in particular, on the teacher's verbal behaviour. However, studies of science classrooms and other inquiry-discovery lessons have highlighted the need to examine non-verbal behaviours also.

The preceding remarks provide a general introduction to this paper which has as its objectives to survey the "state of the art", particularly as it applies to science education and to suggest some new directions which research, involving classroom interaction, might take. It is assumed that readers will be familiar with or have access to several of the earlier reviews on interaction analysis (e.g., Biddle, 1967; Boyd and De Vault, 1966; Campbell, 1968; Flanders and Simon, 1969; Meux, 1967; Nuthall, 1968) and will attempt to read some of the current statements on classroom interaction research (e.g. Rosenshine, 1971; Rosenshine and Furst, 1971; Westbury and Bellack, 1971).

To achieve the aims of the paper, the ensuing discussion has been grouped under four headings which, logically, relate to the objectives. The first three relate to the current state of interaction analysis in science education and respectively deal with various classificatory schemes, the association between classroom behaviours and educational outcomes and the applications of interaction analysis to teacher education. The remarks under the final headings are intended to specify some of the new directions for classroom interaction analysis and classroom research. The ideas presented are by no means original. They are, in fact, modifications and translations from many sources, some of which will not be specified. In a short paper such as this one it seemed appropriate to resist the temptation to provide a comprehensive review and possibly give the reader interaction analysis diarrhoea.

The Classificatory Schemes

Seventy-nine of the hundred or so classificatory schemes are described by Simon and Boyer (1967) and many of these schemes appear to be modifications to or extensions of Flanders work. For example, V.I.C.S, the Verbal Interaction Category System (Amidon and Hough, 1967) and S.C.A.S., the Science Curriculum Assessment System (Matthews, 1970) are very similar to Flanders system and focus on verbal behaviour. One of the most sophisticated schemes is that devised by Bellack and his collaborators (Bellack et al, 1966) and though it focuses on verbal behaviour it has been used with success in science classrooms (Power, 1971). However, the study of verbal behaviour alone does not provide an adequate picture of what is happening in many laboratory or open ended enquiry situations or for that matter, in normal lessons. Consequently several workers have attempted to develop schemes for the classification of non-verbal behaviour (Galloway, 1967; 1970) and schemes that are particularly suited to science lessons. Some of these last mentioned systems are described, albeit briefly, in the following paragraphs. But before this is done, it is appropriate to note the confusion that occurs with the existence and use of a great variety of classificatory schemes (Biddle, 1967) with their different methods for data collection, "conceptual posture" and units of analysis. The choice of the unit of analysis is both a theoretical and a methodological issue and interaction investigators have solved this problem in a number of ways. Several, (e.g. Flanders, 1960; Hongiman, 1967; Matthews, 1970) have used an arbitrary unit of time while others have used selected, naturally occurring units such as an episode (Smith and Meux, 1962) an incident (Nuthall and Lawrence, 1966), a move (Bellack et al, 1966) and a venture (Smith et al, 1967). Some investigators (Barnes et al, 1971; Gump, 1967) have advocated and used "phenomenal units" which are natural breaks in the stream of classroom processes that may reasonably be assumed to be recognized by classroom participants. Each of these units has advantages. The arbitrary nature of time sampling, for example, is an advantage simply because of its mechanical character which facilitates the effort of an observer engaged in systematic observation. However, one difficulty with this unit is that it does not reflect classroom events *per se*. The naturally occurring units, on the other hand, do reflect classroom events *per se* but they are based on the sophisticated concerns of the investigator rather than of the classroom participants. Consequently these units contain an investigator bias even when they are established by "inductive procedures" (Evans and Balzer, 1970). McClellan (1971) in a critique of classroom research states that investigators are more concerned with studying *manners* in the classroom than the essential features of *what* is said, done and made. Certainly, his criticism applies to the majority of projects to date which concentrate on the relative frequency of events rather than on the *form* of the behaviours. Undoubtedly, frequencies are easier to calculate, but the number of higher cognitive questions, for example, gives no indication of their *psychological significance* to each and every pupil.

The fact that there are such a variety of classificatory schemes in use means that it is very difficult to compare and contrast studies and to accumulate a reliable body of knowledge about teaching. There is a need, now, to use one or two schemes consistently in a large number of studies and also, in other projects, to check whether several classificatory schemes used together provide the same description of the lessons. It may be possible by these procedures to reduce the number of viable schemes (in the interests of parsimony) and to generate a reliable body of knowledge on science teaching. At the same time it should be possible to study the substantive aspects of classroom communication, and to examine new patterns of classroom behaviours. The classificatory schemes neglect the substantive aspects of a communication and there is a need to incorporate, in some instances, a study of the accuracy of information. Furthermore, the schemes are based on studies of teaching as it is i.e., on naturalistic studies. It seems appropriate, too, to design teaching situations where "new" or different *patterns* of behaviours occur (Meux, 1967).

The preceding comments are not intended to imply that the classificatory schemes for the analysis of classroom interaction are of no worth. On the contrary. The development of these schemes has provided teachers and researchers with a vocabulary to describe classroom events, thus providing greater power and control, and sensitising teachers to the variety of interactions which may occur. The use of the schemes has resulted also in many valuable descriptions of the patterning of behaviours in science lessons (Matthews. 1968; Parakh, 1967; Tisher, 1971). Furthermore, although there are limitations with the schemes and their units of analysis, the serious research worker has been provided with many significant techniques. Despite many difficulties the serious classroom analyst must use analytic units in order to study classroom procedures in depth (Biddle, 1967).

The researcher in science education requires schemes that will allow the classification of verbal and non-verbal behaviour and that will be applicable to traditional as well as pupil-centred classroom environments. Some schemes may be adapted for these purposes. For example, Frost (1972) adapted M.A.C.I. (Honigman, 1967) to study micro-teaching by science teachers in Brisbane. Some of the schemes that have been designed specifically for science lessons and that can be readily tracked down in the literature include ones by Evans and Balzer (1970), Fischler and Zimmer (1968), Matthews (1968, 1970), and Parakh (1967). Others, about which details are not readily available include those by the Mid-continent Regional Educational Laboratory (Bingham, 1969) and Friedel (1969). There are, too, several other techniques that have been used in studies of individualized instruction e.g. the PLAN Teacher Observation Scale (Quirk, Steen and Lipe; 1971) and in projects on raising pupils' level of thinking (Davidson, 1970) that may well be applicable to studies on science teaching. Be that as it may, the immediate objective here is to refer to some of those schemes that have been developed for and used in science lessons.

The Science Curriculum Assessment System (Matthews, 1970) is being used in Florida, but as yet few reports have appeared. It is sufficient at this stage to note that it is based on a Flanders scheme but that it allows for both pupils' and teachers' behaviours to be coded in enquiry situations. Furthermore, the scheme is designed so that interactions between the teacher and the whole class and the teacher and small groups are distinguished. Parakh's (1967) category system is also based on Flanders work but contains 17, not 10, categories which are grouped into four dimensions labelled Evaluative (Affective-Substantive), Cognitive (Giving or Seeking Substantive Information), Procedural and Pupil Talk (Giving or Seeking Information). The scheme was designed for Interaction Analysis in Biology classes and was used with some success even though some disquietening descriptions of biology classes were presented. Parakh found, for example, that teachers dominated the talk in lessons, most pupil talk was directed toward the teacher and little time was devoted to motivation, evaluation, pausing for reflection and to discussing the nature of science.

The Science Teaching Observation Instrument (Fischler and Zimmer, 1968) differs from the two last mentioned schemes in that the categories are not based directly on previous classificatory systems. The designers believed that behaviours should be classified into two broad categories, namely techniques teachers use to promote learning and questions teachers ask, with appropriate sub-categories (21 in all). Also, they believed that an attempt should be made to categorize the "characteristics" of teaching. They did this by rating a lesson on three continua designated concrete-abstract, practical-theoretical and directed-non-directed. Using these three broad categories and relevant sub-categories an investigator should be able to differentiate between teachers of general science – or so the designers state, and with some justification as Perkes (1968) study indicates.

One other interesting scheme is the Biology Teacher Behaviour Inventory (Evans and Balzer, 1970) which was developed inductively from video-recordings of thirteen biology lessons over a two month interval. The lessons were presented by eleven teachers who taught nine different biology courses e.g., B.S.C.S. Green, B.S.C.S. Blue, Physiology, Zoology and "traditional" biology. Seven categories and twelve sub-categories were established and behaviours were also classified as non-verbal or verbal and as pupil or teacher-centred. The research workers report that considerable success was obtained in categorization of behaviours of teachers (4 B.S.C.S. and 4 non B.S.C.S.) over long (3 months) periods of time. In particular, they showed that their scheme highlighted the significance of non-verbal behaviours in biology lessons. Over 65 per cent of the behaviours coded were non-verbal in nature. The analyses also revealed that over 44 per cent of all hebaviours were managerial ones and that 50 per cent were associated with content development.

When all the descriptive data obtained by using many of the schemes referred to are collated an unexciting picture of science classrooms emerges. They appear to be places where teacher talk predominates, pupils respond more to the teacher than to other pupils, pupils rarely initiate discussions, little time is given to motivation or to exciting interest in science and probing, extending or exploring ideas and evaluation are short-lived. Yet, despite all this, some learning appears to occur for some pupils, and a few of the relevant research findings are reported in the next section. Before passing to this next section, however, it seems appropriate to refer to a Class Activities Questionnaire (Steele, House and Kerins, 1971) which has been used with some success to determine four major dimensions of instructional climate. The instrument could probably be used in conjunction with Interaction Analysis Schemes since it gives the pupils perceptions of the classroom instruction. The CAQ is suitable for high school pupils, and group methods of instruction but is threatening to the teacher.

Classroom Behaviours and Outcomes

In recent years several comprehensive reviews have appeared on the association between classroom behaviours and outcomes (Campbell, 1968; Flanders and Simon, 1969; Nuthall, 1968; Rosenshine, 1971 a & b; Rosenshine and Furst, 1971). Perhaps the most comprehensive are those by Flanders and Rosenshine. Recently, the results of studies in science lessons have been reviewed also (Balzer, 1970; Evans, 1970), and these findings must be seen against the general background of results of research in teaching. For this reason a brief resumé which contains the distilled essence of the first mentioned review papers is in order. The resumé will be followed by specific reference to some of the studies in science classrooms.

(a) the "cognitive clarity" of the teacher's presentation,
(b) the teacher's use of a variety of procedures and behaviours, including extra equipment and displays.

(c) the vim, vigour or enthusiasm displayed by the teacher,
(d) the task-orientation, achievement-orientation and/or business-like manner displayed by the teacher,
(e) the allowance made for or "opportunity given" to pupils to learn the lesson material,
(f) the use made of pupils' ideas and contributions,
(g) the use of structuring statements throughout a lesson,
(h) the use of criticism,
(i) the use of a mixture of low and higher-cognitive type questions,
(j) probing, that is, use of responses which encourage pupils to elaborate on answers, and
(k) maintaining discourse at a "suitable" level of difficulty.

For (h) and (k) there was some evidence to show that increased use of criticism or increase in the level of difficulty was associated with lower achievement. However, the findings are by no means clear and the studies do not present unequivocal evidence. In fact when the research literature is reviewed it becomes clear that little is known about the relationships between classroom behaviours and pupils gains. Certainly, more research on teaching is required. But, a more disappointing feature is that no one category system has been shown to be a superior one in that the behaviours it identifies are significantly and consistently related to pupil gains.

A number of research workers in Australia have studied interaction in classrooms (Tisher, 1972) and in some of these projects science lessons were sampled. Tisher (1970, 1971) studied the patterning of behaviour in eight grade 8 classrooms in Brisbane using a modified Smith and Meux (1962) scheme to categorize teachers' demands. His hypothesis testing exercises proved to be interesting and worthwhile, although only a small number of hypotheses were confirmed. However, the "trends" which were observed were in line with the hypotheses, and a number of significant and interpretable results were obtained. For example, in classrooms where a medium level of cognitive demand was maintained (i.e., a mixture of recall and higher-cognitive questioning occurred) it was found that greater gains in understanding occurred for pupils who were more able or who were rated as high in prior-knowledge-in-science. These findings are in line with some other results of studies on questioning (Gall, 1970), but there is still much more to be done on the effects of different forms of questioning. Gall (1970) raises a number of pertinent issues for research. Recently, Ladd and Andersen (1970) added some additional interpretable findings of value to science educators. They used a modified Smith and Meux scheme to establish a scheme for classification of questions and were able to show that pupils made greater gains on criterion tests when they were taught by teachers who more frequently asked "high inquiry" (i.e. high cognitive demand) questions. What is needed now is an extension of this work on questioning to determine what effect changes in level of demand and thinking in a classroom have on pupils' responses and achievements. Perhaps the work by Davidson (1970) could serve as a guide and starting point for science educationists. Certainly, there is a need to break away from the Flanders scheme when studying inquiry and higher-cognitive questions in science lessons. A recent use of the Flanders scheme was reported by Citron and Barnes (1970), but they too focussed on ID ratios rather than level of cognitive demand.

Not all researchers in science education are enamoured with the Flanders scheme and its derivatives. For example, Evans (1970) in a comprehensive review of eight studies in science education, expressed a number of doubts about the uses of the scheme, the generalizability of results, the research techniques and the nature of the interpretations. Nevertheless, he does believe the scheme and the studies using it have made a contribution to research in science education. Some research workers have developed schemes of their own, suited to their peculiar needs, and others have relied on special schemes designed for use in science classrooms. Balzer (1970) has reviewed the results of a number of these projects and those by Kleinman (1965) and Perkes (1968), among others, are probably of particular interest to Australian science teachers. The Perkes (1968) study, especially, is intriguing because he uses the Science Teaching Observation Instrument (STOI) which was developed by Fischler and Zimmer (1968). He found that the achievement in science of junior high school pupils was related to behaviour patterns in classrooms. For example, pupil-oriented activities like laboratory discussions, hypothesizing and "use of applications in the role of clarifying principles" were related (correlated) with high scores on tests involving interpretation.

It is appropriate at this stage to refer to several studies being undertaken in Christchurch, New Zealand (Nuthall and Church, 1971; Wright and Nuthall, 1970). Although these deal with science lessons in primary schools they are significant projects for researchers in secondary schools because they illustrate a type of laboratory controlled classroom research. In both projects lesson content was tightly controlled by scripting and attention was limited to verbal interaction. In the earlier project (Wright and Nuthall, 1970) an attempt was made to identify the short-term effects of teacher behaviours in a set of three subject-matter (life and habitat of the black-backed seagull) controlled lessons of the discussion type common in most primary school classrooms. Seventeen teachers were used in the study, measures were obtained for pupils' ability, prior knowledge in science and achievement and multiple regression techniques were used to examine various relationships including that between achievement and ability and prior knowledge. Relationships between teacher variables and pupil achievement were also studied and the examinations of the data suggested that there were six major kinds of teacher behaviours related to pupil achievement. These kinds of behaviours were designated patterns of solicitation, type of solicitation, reciprocation and redirection, structuring, reaction to responses and recapitulation and revision. With respect to type of solicitation, for example, although the *frequencies* of closed and open questions were not significantly related to achievement, the *percentage* of closed questions was. Also, revision at the end of a lesson, rather than recapitulation of a previous lesson at the beginning of the next, was clearly related to achievement.

The preceding paragraphs do not cover all the pertinent studies of interaction in science lessons, e.g., the ones by Moon (1971) and Smith (1971). Some Australian studies are also omitted. These include those by Frost (1972) and Power (1971) in Brisbane and Young (1972) at Macquarie. The two last mentioned projects make use of the Bellack (1966) and the Flanders schemes respectively. Frost's study deals with the behaviours of experienced and less experienced teachers in micro-teaching situations and makes use of MACI (Honigman, 1967). Of the three projects the one by Power is the most sophisticated and it has been reported elsewhere (Power, 1972 a & b). Hopefully, other researchers will follow his lead and use multivariate statistical procedures to study person-environment interactions, rather than teacher-pupil interactions or verbal interactions.

Interaction Analysis and Teacher Education

The techniques of interaction analysis have been applied to pre-service and in-service education of science teachers with variable results. At present interaction analysis does not seem to be a bonanza for teacher education. Nevertheless, some valuable studies have been undertaken and they appear to fall into three broad categories e.g., those dealing with the effects of new science courses on teachers' behaviours, those on the effects of in-service education on classroom behaviours and those dealing with changes in teaching behaviour during practice teaching. In the last mentioned category are included those projects which examine the similarity in behaviours between student teachers and their supervising teachers. References to these various types of study follow.

The first two categories are not mutually exclusive, as evidenced by the following discussion. Be that as it may, Wilson (1969) in a review of the literature on the behaviour of teachers following new science curricula, concluded that these teachers were not only asking more questions of the higher cognitive type but were asking more questions in general. This conclusion was based on the work of Harris and McIntyre in Texas. They developed a classificatory scheme based on Bloom's Taxonomy – a system which allowed an observer to classify teacher demands (questions) into six types of which two were recall (lower cognitive) and the others were higher cognitive.

Hunter (1969) studied the effects of training in the use of new science programs upon the verbal behaviour of primary teachers as they teach science. She trained seventy five teachers in the use of six new science curricula and then in the year after the training program observed their behaviour in science lessons. The Revised VICS was used to classify classroom behaviours. The behaviours of the trained teachers were compared with those of teachers who had received no special training in the use of the new curricula. Several hypotheses were tested, e.g., "teachers in the experimental group will use a greater range of verbal behaviour than teachers in the control group", but only one was not rejected. It related to use of *materials* in the experimental and control classes and indicated that pupils used materials to a greater extent in the first mentioned classes. A comparison of the categories of verbal behaviour between the two groups of teachers indicated similarity in all categories – that is, the trained teachers did not vary from the untrained in their verbal behaviour. Hunter interpreted this finding as indicating that the new curriculum materials, contrary to their claims, did not encourage the use of divergent and evaluative thinking in classrooms. This interpretation is open to debate especially when the nature of sampling in the study is examined. Each teacher was only observed on two separate occasions.

Hall (1969), using a specially designed *Instrument for the Analysis of Science Teaching* (IAST), studied the effects of a new curriculum *(Science – A Process Approach)* on three groups of teachers. One group attended a five-day summer workshop before adopting the curriculum and another received in-service education during the year preceding adoption. Both these groups received supervisory help when they began teaching the new course. A third group implemented the curriculum without in-service education or supervisory help. An analysis of the teaching behaviours of the three groups indicated that there were differences among them. The teachers who attended the summer workshop appeared to be implementing the curriculum more effectively, i.e. they followed the ideals of the curriculum more closely. In another study involving *Science - A Process Approach* Newport and McNeill (1970) found that teacher-pupil verbal interactions changed after teachers had some training in the use of the new materials. They also studied the effects of training in Interaction Analysis on the classroom behaviour of teachers following a set science text. Of twenty-three measures investigated only one was observed to change significantly. Newport and McNeil argued, as a consequence, that interaction analysis, when used without any accompanying instruction on the philosophical and psychological bases of science education is relatively ineffective in bringing out changes in teacher-pupil interaction in lessons. Interaction analysis only indicates *what is* happening, not what *should* be.

The preceding studies used primary school teachers. Few projects appear to have involved ones from secondary schools. One of these was reported by Rentoul (1972) in a previous Australian publication. This project involved 45 teachers who were participating in an in-service course for Harvard Project Physics. In another study Matthews (1968) used eighteen secondary student teachers and their co-operating teachers. His study, which differs from Rentoul's, was an attempt to identify changes in the classroom behaviour of the student teachers during practice and to determine how similar students' behaviours were to those of the co-operating teachers. Although the student teachers behaviours appeared to become more like that of the teachers in some respects there were many marked differences. In terms of the Flanders System which was used to categorize behaviour the similarities were in the occurrence of "silence or confusion", length of statements and occurrence of pauses following teahcers' questions. But of what use is this information unless it is linked with discussions on the nature and purposes of science education? Certainly, Matthews appeared to have some reservations about the value of his project for he listed twenty-three propositions for study in future research.

New Directions

A wealth of knowledge has been derived from interaction studies but a basic difficulty in using the knowledge is how to put together the findings from different studies. At present interaction analysis appears to be bedevilled by a phenomenon akin to "speaking in tongues". Each investigator tends to use his favourite or esoteric language for naming variables and a synthesis of different studies is extremely difficult. As a consequence it is well nigh impossible to organize the knowledge into coherent theories. Of course, it may be that classroom research is extremely difficult and the problems to be untangled extremely complex. On the other hand, perhaps the approach that has been adopted is fundamentally unsound. Certainly, Travers (1971) suggests that this may be the case. He believes that a more careful study of particular classes of teaching events is required in specially contrived "laboratory type" situations. This research may involve classroom-like situations but these situations will need to be more highly controlled than those situations involved in previous classroom research. Perhaps the type of situations Travers had in mind were the ones used by Nuthall and Church (1971) in their studies of teacher behaviour in New Zealand. In these studies there was a systematic and controlled variation of teacher behaviours. For example, lessons were prepared word-for-word and scripted. The lessons involved the use of prepared materials, and the scripts consisted of the information moves and questions which the teacher was to ask, the comments and subsidiary questions he could use when certain kinds of pupil responses occurred, and a predetermined method of addressing questions around the class. The controlled variations which were introduced were the amount of time taken to complete lessons and the "coverage of content" in lessons.

However, some research workers believe that the laboratory type studies of the kind that Travers sees as essential cannot produce the forms of knowledge we need to understand "teaching" as "a potentially meaningful real-world act" (Westbury, 1971). These workers are turning to the writings of people such as Gump (1971) and Smith and Brock (1970) who are regarded by Westbury (1971) as providing "thrusts not yet fully developed, to new research genres". Certainly Smith's work is less rigorous and systematic than that of many of the more accepted prophets of interaction analysis, but it is most intriguing. His classroom is less homogeneous in texture, with many strands to it and many actions. Also, he stresses that particular attention should be given to the conceptualization of the "latent" variables of classrooms. Whether we do this, or whether we attempt to conduct laboratory-type research, there is, in addition a need to synthesize many of the classificatory schemes, to study the effects of different patterns and sequences of behaviour, to conceptualize studies in multivariate terms and to use multivariate techniques to analyse data. These four last mentioned suggestions appeared in earlier sections of this paper. Furthermore, there is a case for studying person-environment interactions as Power (1971) has done.

The preceding suggestions do not contain all the possible directions that interaction studies may take. There is, too, a need to incorporate interaction analysis in projects of the action research type. Certainly, if the results of educational research are to be incorporated readily and rapidly in educational practice, a case can be made for more action research. One such intriguing project, which amalgamates the techniques of interaction analysis, some elements of control as advocated by Travers, and uses A.S.E.P. materials is in progress in Brisbane. Details are provided elsewhere.*

Although the new religion of interaction analysis has produced more chaos than organized knowledge, the situation is not a hopeless one. On the contrary, there are many exciting new directions to be taken. But, the challenge for the research worker is to maintain a sense of perspective.

---

* Dr. C.N. Power discusses the project in the second paper of this publication.

REFERENCES

Amidon, E. & Hough, J. (Eds) (1967). *Interaction Analysis : Theory, Research and Application*, Reading, Addison-Wesley.

Balzer, L. (1970). "Review, Appriasal, and Recommendations Concerning Research on Classroom Behaviour in Science". Paper presented at 43rd Annual Meeting of National Association for Research in Science Teaching, Minneapolis (Mimeo).

Barnes, D. et al (1971). *Language, the learner and the school*, Middlesex, Penguin Books.

Bellack, A. et al (1966). *The Language of the Classroom*, New York, Teachers' College Press.

Biddle, B.J. (1967). "Methods and Concepts in Classroom Research". *Review of Educational Research, 37*, 337-357.

Bingham, R. (Ed) (1969). *Inquiry Objectives in the Teaching of Biology*, Mid-continent Regional Educational Laboratory, Vol 1, No. 1, Kansas City, Missouri.

Boyd, R.D. & De Vault, M.V. (1966). "The Observation and Recording of Behaviour", *Review of Educational Research, 36*, 529-551.

Campbell, J. & Barnes, C. (1969). "Interaction Analysis – A Breakthrough", *Phi Delta Kappan, 50*, 587-590.

Campbell, W.J. (1968). "Classroom Practices", *New Zealand Journal of Educational Studies, 3*, 97-124.

Citron, I. & Barnes, C. (1970). "The Search for More Effective Methods of Teaching High-School Biology to Slow Learners Through Interaction Analysis. Part I. The Effects of VArying Teaching Patterns. *Journal of Research in Science Teaching, 7*, 9-19.

Davidson, R. (1970). "The Use of Interaction Analysis in Studying Teacher Influence on Pupils' Levels of Thinking", *Classroom Interaction Newsletter, 6*, No. 2, 26-39.

Evans, T. (1970). "Flanders System of Interaction Analysis and Science Teacher Effectiveness". Paper presented at 43rd Annual Meeting of National Association for Research in Science Teaching, Minneapolis. (Mimeo).

Evans, T. & Balzer, L. (1970). "An Inductive Approach to the Study of Biology Teachers Behaviours", *Journal of Research in Science Teaching, 7*, 47-56.

Fischler, A. & Zimmer, A. (1968). "The Development of an Observational Instrument for Science Teaching". *Journal of Research in Science Teaching, 5*, 127-137.

Flanders, N.A. (1960). *Teacher Influence, Pupil Attitudes, and Achievements.* U.S. Department of Health Education and Welfare, Office of Education Cooperative Research Project No. 397, University of Minnesota, Minneapolis.

Flanders, N. & Simon, A. (1969). "Teacher Effectiveness", *Classroom Interaction Newsletter, 5,* No. 1, 18-37. (Reprinted from *Encyclopedia of Educational Research,* edited by R. Ebel, New York, Macmillan, 1969, pages 1423-1431.

Friedel, A. (1969). "A Procedure for Observing Teacher and Pupil Behavior in the Science Classroom". Paper presented at 43rd Annual Meeting of the National Association for Research in Science Teaching, Pasadena.

Frost, R. (1972). *Applications of the Multidimensional Analysis of Classroom Interaction in Mini-Teaching Situations,* B.Ed. (Hons) Thesis, University of Queensland (in preparation).

Gall, M.D. (1970). "The Use of Questions in Teaching", *Review of Educational Research, 40,* 707-721.

Galloway, C. (1967). "A Model of Teacher Nonverbal Communication", *Classroom Interaction Newsletter, 4,* No. 2, 12-21.

Galloway, C. (1970). "Galloway System" in Simon, A. & Boyer, G. (Eds), *Mirrors for Behavior II. An Anthology of Observation Instruments. Volume B.* Philadelphia, Research for Better Schools, Inc.

Gump, P. (1967). "Setting Variables and the Prediction of Teacher Behavior", Paper presented to the American Educational Research Association, New York.

Gump, P. (1971). "What's Happening in the Elementary Classroom", in Westbury, I. & Bellack, A. (Eds), *Research into Classroom Processes,* New York, Teachers' College, Columbia University.

Hall, G. (1969). "A Comparision of the Teaching Behaviors of Second Grade Teachers Teaching *Science – A Process Approach* with Second Grade Teachers not teaching a Recently Developed Science Curriculum," Paper presented at 42nd Annual Meeting of the National Association for Research in Science Teaching, Pasadena.

Honigman, F. (1967). "Multidimensional Analysis of Classroom Interaction" in Simon, A. & Boyer, G. (Eds). *Mirrors for Behavior,* Philadelphia, Research for Better Schools, Inc.

Hunter, E. (1969). "The Effects of Training in the Use of New Science Programs upon the Classroom Verbal Behavior of First Grade Teachers as they Teach Science", *Classroom Interaction Newsletter, 4,* No. 2, 5-11.

Kleinman, G. (1965). "Teachers' Questions and Student Understanding of Science". *Journal of Research in Science Teaching, 3,* 307-317.

Ladd, G. & Andersen, H. (1970). "Determining the Level of Inquiry in Teachers' Questions", *Journal of Research in Science Teaching, 7,* 395-400.

Matthews, C. (1968). "Classroom Verbal Behavior of Science Teachers and their Cooperating Teachers", *Journal of Research in Science Teaching, 5,* 193-201.

Matthews, C. (1970). "Science Curriculum Assessment System" in Simon, A. & Boyer, G. (Eds), *Mirrors for Behavior II. An Anthology of Observation Instruments, Volume B,* Philadelphia, Research for Better Schools, Inc.

McClellan, J. (1971). "Classroom-Teaching Research: A Philosophical Critique", in Westbury, I. & Bellack, A. (Eds), *Research into Classroom Processes,* New York Teachers' College, Columbia University.

Meux, M. (1967). "Studies of Learning in the School Setting", *Review of Educational Research, 37,* 539-562.

Mitchell, J.V. (1969). "Education's Challenge to Psychology: The Prediction of Behavior from Person-Environment Interactions", *Review of Educational Research, 39,* 695-721.

Moon, T. (1971). "A Study of Verbal Behavior Patterns in Primary Grade Classrooms During Science Activities", *Journal of Research in Science Teaching, 8,* 171-177.

Newport, J. & McNeill, K. (1970). "A Comparision of Teacher-Pupil Verbal Behaviors Evoked by Science - A Process Approach and by Textbooks", *Journal of Research in Science Teaching, 7,* 191-195.

Nicholas, A. (1968). "Interaction Analysis", *American Educational Research Journal, 5,* 285.

Nuthall, G. (1968). "Types of Research on Teaching", *New Zealand Journal off Educational Studies, 3,* 125-147.

Nuthall, G. & Church, J. (1971). "Experimental Studies of Teacher Behavior", Paper presented at 43rd Congress of Australian and New Zealand Association for the Advancement of Science, Brisbane, Australia.

Nuthall, G. & Lawrence, P. (1966). *Thinking in the Classroom,* Wellington, New Zealand Council for Educational Research.

Parakh, J. (1967). "A Study of Teacher-Pupil Interaction in High School Biology Classes. Part II. Description and Analysis", *Journal of Research in Science Teaching, 5,* 183-192.

Perkes, V. (1968). "Junior High School Science Teacher Preparation, Teaching Behavior, and Student Achievement", *Journal of Research in Science Teaching*, 5, 121-126.

Power, C.N. (1971). *The Effects of Communication Patterns on Student Sociometric Status, Attitudes and Achievement in Science.* Ph.D. Thesis, University of Queensland, Brisbane.

Power, C.N. (1972a). "The Effects of Pupil Involvement in Classroom Interaction of Science Achievement", in Tisher, R. (Ed), *Research 1971*, Australian Science Education Research Association, Melbourne, A.S.E.P.

Power, C.N. (1972b). "The Effects of Communication Patterns on Student Outcomes", in Bellack, A. (1972). *The Language of the Classroom*, New York, Teachers College Press (in press).

Quirk, T., Steen, M., & Lipe, D. (1971). "Development of the Program for Learning in Accordance with Needs Teacher Observation Scale", *Journal of Educational Psychology*, *62*, 188-234.

Rentoul, J. (1972). "Self Instruction Using Microteaching and Videotape Feedback During an In-service Education Course for Teachers of Harvard Project Physics", in Tisher, R. (Ed) *Research 1971*, Australian Science Education Research Association, Melbourne, A.S.E.P.

Rosenshine, B. (1971a). "Teaching Behavior Related to Pupil Achievement: A Review of Research", in Westbury, I. & Bellack, A. (Eds) *Research into Classroom Processes*, New York, Teachers' College Press.

Rosenshine, B. (1971b). *Teaching Behaviours and Student Achievement*, Slough, Bucks, National Foundation for Educational Research in England and Wales.

Rosenshine, B., & Furst, N. (1971). "Research on Teacher Performance Criteria", in Smith, B. (Ed), *Research in Teacher Education*, Englewood Cliffs, Prentice Hall.

Simon, A. & Boyer, G. (Eds), (1967). *Mirrors for Behavior*, Philadelphia, Research for Better Schools, Inc.

Smith, B. et al (1967). *A Study of the Strategies of Teaching*, U.S. Department of Health, Education and Welfare, Cooperative Research Project No. 1640, Urbane, Bureau of Educational Research, University of Illinois.

Smith, B. & Meux, M. (1962). *A Study of the Logic of Teaching*, U.S. Department of Health, Education and Welfare, Office of Education, Cooperative Research Project No. 258, Urbana, Bureau of Educational Research, University of Illinois.

Smith, J. (1971). "The Development of a Classroom Observation Instrument Relevant to the Earth Science Curriculum Project", *Journal of Research in Science Teaching, 8,* 231-235

Smith, L. & Brock, J. (1970). "'Go, Bug, Go!': Methodological issues in classroom observational research". CEMREL Occassional Paper Series No. 5, St. Ann, Missouri: Central Midwestern Regional Educational Laboratory.

Steele, J., House, E., & Kerins, T. (1971). "An Instrument for Assessing Instructional Climate through Low-Inference Student Judgments". *American Educational Research Journal, 7,* 447-466.

Tisher, R.P. (1970). "Association between verbal discourse and pupils' understanding in science", in Campbell, W.J. (Ed), *Scholars in Context,* Sydney, John Wiley & Sons.

Tisher, R.P. (1971). "Association between Classroom Behavior and Pupils' Understandings: An Application of the Smith and Meux Technique", *Classroom Interaction Newsletter, 7,* No. 1, 30-39.

Tisher, R.P. (1972). "Studies of Behavior in Australian Classrooms", *Classroom Interaction Newsletter, 8,* No. 1 (in press).

Travers, R. (1971). "Some Further Reflections on the Nature of a theory of Instruction", in Westbury, I. & Bellack, A. (Eds), *Research into Classroom Processes,* New York, Teachers' College, Columbia University.

Westbury, I. (1971). "Problems and Prospects", in Westbury, I. & Bellack, A. (Eds), *Research into Classroom Processes,* New York, Teachers' College, Columbia University.

Westbury, I. & Bellack, A. (1971). *Research into Classroom Processes,* New York, Teachers' College, Columbia University.

Wilson, J. (1969). "The 'New' Science Teachers Are Asking More and Better Questions", *Journal of Research in Science Teaching, 6,* 49-53.

Wright, C. & Nuthall, G. (1970). "Relationships between Teacher Behaviors and Pupil Achievement in Three Experimental Elementary Science Lessons", *American Educational Research Journal, 7,* 477-491.

Young, R. (1972). "An Experimental Study Comparing a Guided Discovery Method with Two Other Methods of Teaching a Primary School Science Unit", in Tisher, R. (Ed), *Research 1971,* Australian Science Education Research Association, A.S.E.P. Melbourne.

ANTECEDENT VARIABLES AND LEARNING IN SCIENCE

Editor's Comment:

The three papers in this section deal with variables which may be classified as antecedent ones that are associated with learning in science. Professor Fensham deals with a pupil's prior knowledge, P. Gardner with vocabulary and attendant difficulty of words and Dr. Cohen considers the characteristic designated as creativity. Gardner's contribution is a significant one for Australian teachers and researchers for he presents an extensive list of words with their respective "difficulty levels" for pupils in secondary school forms I, II, III and IV.

PRIOR KNOWLEDGE – A SOURCE OF NEGATIVE FACTORS FOR SUBSEQUENT LEARNING

P.J. Fensham

The significance of prior knowledge for subsequent learning has been recognised by theorists concerned with cognitive learning of verbalised or symbolic material.

One categorisation of situations in which prior learning is significant is that of lateral, sequential and vertical transfer.[1] Empirical investigation has been aided by the distinction in learning task implied by these categories and there is, for example, encouraging evidence that massive general transfer can be achieved by appropriate learning.[2] In the light of this finding the attention of educators and curriculum designers has been given to the provision of strategies for solving classes of problems, of powerful ideas that constitute the structure of a discipline of study[3] and of arrangement of learning material in ways that seem to coincide with cognitive organisational principles.[4]

In science and mathematics the degree of sequential dependence of much of the content is so great that these questions of how and what transfer of learning occurs assume very great importance. It has been argued by Gagné[5] (and in a more tightly defined way by White[6]) on the basis of empirical studies that some of these sequences are so definite that without mastery of a prior step there can be no further progress in the learning of subsequent ones. Gagné and the others who have followed his lead pay close attention to the series of small learning skills that lead up to the intended new learning and on their sequence in relation to each other. Such a hierarchy of learning steps is then a useful guide for the arrangement of the new material and also for developing checks to see just where learning is breaking down.

This approach has had very valuable outcomes for teachers as it has led to both "Readiness" and Diagnostic" types of tests. The former check the prior knowledge that is assumed by the new material[7] and inform the teacher which learners need extra attention to being them to a common starting point with the others. The latter are used by teachers after a period for new learning has occurred.[8] They probe the steps in this learning and indicate to teacher and learner where breakdown has occurred. This information is certainly not available if only testing of terminal performance occurs, particularly when objective test items are used.

The work of Bruner[3] and Ausubel[4] with respect to the relation between prior knowledge and subsequent learning are both more global than this atomistic approach of Gagné. However, all three authors emphasise the positive links between what has been and what is to be learned. They argue that because there is a logical and psychological link, the possibility of sequential and vertical transfer of learning will be enhanced by suitable arrangements and presentations. If this is so, as seems highly likely, then the possibility will also exist, given other arrangements and presentations, that the prior knowledge will have a negative rather than positive effect on subsequent learning.

Unexpected and "undesirable" transfers may occur even though they are not intended. This may clearly be so if prior knowledge if faulty but "Readiness Tests" should deal with this. However, it may also happen if the prior knowledge is hitherto quite sound, but the wrong parts of it are used in the sequential or vertical transfer.

If we consider the schematic arrangement shown in Figure 1 of a learner's prior knowledge and the new learning, it is clear that there are some parts of the prior knowledge that have positive significance for the learning task.

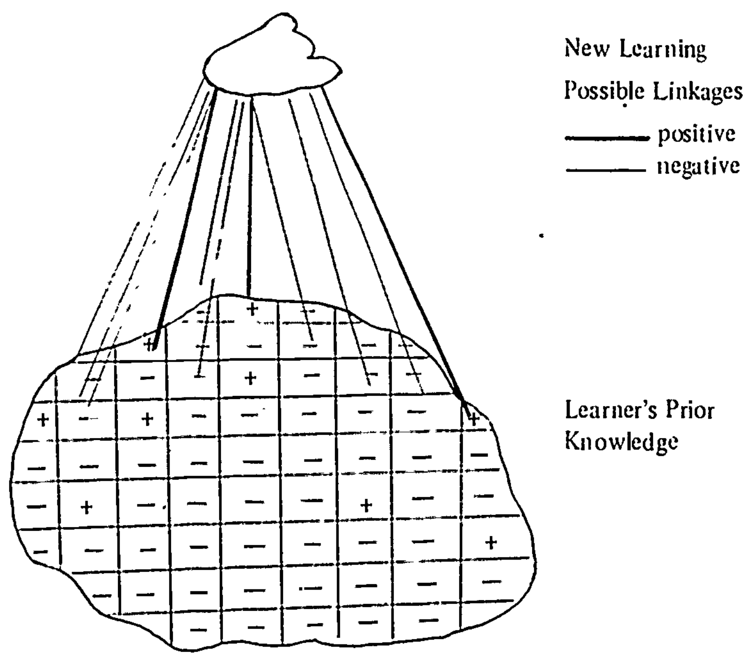


Figure 1: Possible linkages between New Learning and the Learner's Prior Cognitive Structure.

However, the whole set or prior knowledge will for any learner be enormously larger than those elements of it which positively relate to a new task. As the authors referred to above, it is part of the skill of teaching to know which are the positive elements and so to present material to the learner that he will have and make use of these facilitating elements.

In many cases of teaching and learning situations this clarity is lacking and the teacher unwittingly, or the learner quite innocently, may use linkages that have a negative effect.

There appears to be surprisingly little attention in the literature to these negative aspects of prior learning. Ausubel[9], in suggesting the way in which he believes organising principles can operate in learning, does get close to discussing negative interaction. He claims that two consequences seem likely if ideas specifically relevant to new meaningful material are not available in the learner's cognitive structure. These are rote learning, or learning which makes use of tangential or less relevant ideas from the prior knowledge. If the second should occur it could be expected to be poorly anchored in the learner's cognitive structure; that is, it would be learned with ambiguous meaning with little longevity. A third possible consequence can be added, namely, learning occurs which is a distortion of what was to be learned. If this distortion changes the learning so that it is consonant with the tangential ideas from the prior knowledge, it may, in its distorted form, be well assimilated into the cognitive structure and have considerable permanency.

These three possible learning outcomes will, I believe, be familiar to many teachers, particularly at the tertiary level. Here the teaching and learning modes are such that there is little chance to observe or correct the step-wise phase of the learning. Only terminal performance after long unchecked learning sequences is displayed. Furthermore, the lack of feedback en route and the traditional nature of lecturing, as distinct from other teacher-learner modes, makes a psychological learning arrangement of new material less likely than a logical ordering which is retrospective from the opposite position of mastery.

To overcome these problems Ausubel[8] urges teachers in their verbal or literary presentations to use suitable organising ideas in the introductory stages of new learning.

However, as yet there is little empirical evidence to support what seems to be reasonable hypotheses of what occurs in learning when the situation of Figure 1 is uncontrolled or directed for the learner.

Method of Approach

The approach to positive linkage in learning that has been used by Gagné *et al* is to ask for a given skill in the new learning. "What immediately prior skill or skills does the learner need to have if he is to achieve mastery of this one?" This question is repeated until a possible hierarchy or sequence from prior knowledge to new learning is postulated for empirical investigation.

By analogy, a possible method of exploring negative linkage is to confront the learner with the new learning in its most succinct form and ask him the question, "What do you think, from what you know now, is relevant to this topic?" This approach can be used in a quite open-ended fashion or in a more structured and limited set of options. Both approaches need to be explored and there are probably different gains to be had from each. Since high level learning usually takes place in a clearly defined context, most of the preliminary studies were done using a limited set of options.

The Second Law in Chemistry – an example of new learning for all

In his famous introduction to the continuing "Two-Cultures" debate, C.P. Snow gave the Second Law of Thermodynamics as one of his examples of highly significant science that is not understood or appreciated by otherwise highly educated, non-scientists. Whether because of this reference or for some other reason, it is interesting to note that a great number of curriculum reformers in science education in the last 10 years have attempted not only to improve the teaching of this topic in its traditional tertiary science setting,[11] but also to include it in the scope of new science courses in the secondary school.[12]

Snow might, of course, have added that this topic in science has also been little understood by the majority of science students who are exposed to it in tertiary courses of physics or chemistry. The reasons for this difficulty have not been systematically explored, and accordingly the various recent attempts to teach the topic in new ways can only be said to be based on the experience and intuition of particularly innovative teachers.

Experience of this sort over a number of years also suggests to the author that rote learning or "distorted" learning occurs very frequently in relation to this topic compared with some others in physical chemistry. On Ausubel's hypothesis this is one criterion that may be used as a pointer to the influence of absent or inappropriate and less relevant organising ideas in the learner's cognitive structure. If this topic is now to be taught to secondary students with any meaning at all, then it would appear that great care needs to be exercised to optimise the learning situation.

Previous Cognitive Structure and the Second Law

What aspects of the learner's cognitive structure are likely to interact with the new learning material of a topic such as the Second Law in chemistry? To explore this relationship two techniques have so far been used.

The first was based on an examination of the approach suggested as the teaching mode to the subject in five recent curricula. This revealed that each of these approaches leant heavily on a particular physico-chemical system with, in each case, a plausible hope that this concrete or quasi-concrete (second-hand data) would provide a good base for the generalised abstractions of the formulae of the Second Law. The phenomenological systems used were the electrochemical cell,[13] a gas piston,[14] mixing energies and the energy levels of molecules,[15] physical work situations[16] and the thermodynamic data for a variety of chemical reactions with as solution.[17]

Accordingly, a list of twenty-five physico-chemical phenomena was made which included all of the above, (sometimes worded in more than one way). The phenomena were chosen so that they were not entirely strange to senior secondary students of science, but they had not been covered in depth by the courses at that level of science education.

Students in five schools and at two levels (fourth and sixth years of formal study of chemistry and physics) were asked to indicate from their existing knowledge of the subject areas, whether they would expect the phenomenon to be treated in detail in a physics course, a chemistry course, or whether both were likely.

The task was not difficult and almost all the students were able to make a prediction about each of the phenomenon. There were no uncertainties about several of the items. For example, the "production of ammonia from nitrogen and hydrogen" and "the patterns formed when light passes through a grid of very small slits" were rated "chemistry" and "physics" respectively by all respondents. A few items were much less certainly located: for example, the "absorption of heat at constant temperature when a liquid boils".

The items which related to the phenomena of the various curricula were rated as shown in Table 1.

Phenomenon	Particular Curriculum	Students' Expectation		
		Physics	Chemistry per cent	Either or Both
Electrochemical cell	Chemistry	87	4	9
Gas piston	Chemistry	76	4	20
Energy levels of molecules	Physics	0	60	40
Work situations	Chemistry	85	3	12
T.d. data of chemical reactions	Chemistry	0	90	10

Table 1. Students' expectation of the subject identity of various phenomena.

Only the last phenomenon is perceived by a majority of students as associated with the subject area in which it is used by the curriculum planners as a key initial phenomenon for the development of a difficult piece of learning. Armed with these findings and aware of the limitations of transfer across subject barriers, it certainly appears likely, in the terms of Ausubel's model, that the relevant organising ideas will not emerge or be used by the learner unless considerable care and attention is given by teachers to building appropriate bridges between existing knowledge and the content of these new curriculum materials.

As a second exploration of the possibility of organisers emerging in the absence of explicit identification of relevant ones by the teacher, the fact that "some mixtures of chemicals react and some do not" was simply brought before the students (Form 4 and 6 as above). For this dichotomous situation concerning chemical mixtures they were asked to indicate which of a list of possible reasons may have had some relevance. The list of reasons was drawn up after consultation with experienced teachers who considered what had gone before this stage in the chemical education of the students.

Again, most students had little difficulty replying and all indicated at least several reasons as potentially relevant. Table 2 gives those items which were positively checked by more than 35% of the respondents.

Reason	Percentage Respondents
A. The atoms in the products can get closer together	> 50%
B. Reactions always give out heat	> 50%
C. Reactions lead to greater mixing of the atoms	> 50%
D. The atoms in the products attract each other more strongly	> 50%
E. The products are less like pure substances	35-50%
F. The products occupy less volume	35-50%
G. Reactions occur if a gas or precipitate can form	35-50%
H. The energy is shared more evenly as a result of reaction	35-50%

Table 2. Reasons for chemical reactions that were perceived as potentially relevant.

The full significance of these replies from substantial proportions of the sample would require much more detailed investigation. However, the appearance of reasons like B and G, which may be regarded as irrelevant or erroneous ideas for the new learning, are not surprising when the aims and experience of earlier learning in chemistry is considered. Very effective instruction at that stage may well result in a cognitive structure that would seek to apply these ideas to a topic like the thermodynamics of reaction direction. Similarly, the number of reasons (A,C,D) that relate to an "atomistic" model of reaction processes is a not unexpected consequence of earlier learning. They are, however, a quite irrelevant idea for the way in which the new learning topic is usually presented.

Conclusion

These two preliminary approaches to negative aspects of the learner's prior cognitive structure appear sufficiently promising that it is now proposed to explore this field more systematically. Several open-ended forms of questioning will be compared with the closed approach used here. When a suitable technique for displaying the organising ideas being used by learners has been established, we should then be in a position to empirically investigate Ausubel's very important hypotheses.

In the meantime, the existing techniques already provide material for a new section of the teacher's guides to various curricula. This would consist of an estimate of the percentage of learners with certain expectations about the topic of a lesson or series of lessons. Where this is discrepant from the suggested teaching mode, suitable bridging steps could be suggested.

REFERENCES

1. D.P. Ausubel and F.G. Robinson, *School Learning*, chap. 6, New York, Holt, Rinehart & Winston Inc., 1969.

2. J.S. Bruner, *The Process of Education*, Cambridge, Mass., Harvard U.P., 1960.

3. J.S. Bruner, "Some Theorems on Instruction" in *Theories of Learning and Instruction*, 63rd Year Book, Nat. Soc. Stud. Educ. Part 1, Chicago, J. of Chicago Press, 1964, pp. 306-335(b).

   and G.C. Finlay, "Secondary School Physics: the PSSC." *Amer. J. Physics*, 1960, *28*, pp. 574-581.

4. D.P. Ausubel, *Educational Psychology*, chap. 4, New York, Holt, Rinehart & Winston, Inc., 1968.

5. R.M. Gagné, "Contributions of Learning to Human Development", *Psych. Rev.*, 1968, *3*, pp. 177-191.

6. R.T. White, "Learning Graphical Skills in Kinematics", unpublished Ph.D. thesis, Monash University, 1972.

7. For example, P.L. Gardner, "Biology Readiness Test Materials", A.C.E.R., 1969.

   L.D. Mackay, "Physics Readiness Diagnostic Tests", A.C.E.R., 1967.

8. For example, "Diagnostic Chemistry Test Series", A.C.E.R.

9. D.P. Ausubel, *Educational Psychology*, op. cit., p. 131.

10. C.P. Snow, *The Two Cultures and a Second Look*, 1963, Cambridge U.P.

11. See articles in J. Chem. Education by M.L. McGlasham, *43*, pp. 226-232, 1966, and by N.C. Craig, L.E. Strong and H.F. Halliwell, and L.K. Nash in *47*, No. 5, 1970.

12. In particular, various Nuffield Foundation Science Projects, see ref. 13, 15, 16.

13. Nuffield Chemistry, Advanced Level, Penguin Education, 1971.

14. B.A.W. Coller. I.R. McKinnon and I.R. Wilson, "Physical Chemistry - a behavioural view", Monash University, 1970.

15. Nuffield Physics, "Advanced Level", Longmans, 1972.

16. Nuffield Chemistry, "Orindary Level", Longmans/Penguin, 1966.

17. J.M. Millen, private communication re course at University College, London.

DIFFICULTIES WITH NON-TECHNICAL VOCABULARY AMONGST JUNIOR SECONDARY SCHOOL STUDENTS: THE WORDS IN SCIENCE PROJECT

P.L. Gardner

Difficulties with language represent a major barrier to educational progress for a significant minority of Australian students. Various sources have estimated that 15% of students entering secondary school are functionally illiterate, with reading abilities less than those of the average eight-year-old child. Other children read with moderate fluency, but their comprehension is limited by their restricted range of vocabulary. These difficulties naturally pose major problems for a science curriculum development project such as ASEP's.

The vocabulary problem can be attacked by identifying what vocabulary is required in the learning of science, and by investigating the extent to which the words so identified prove difficult to secondary school pupils. This is the rationale of the present research.

Between 1968 and 1971, the author directed a project which set out to identify important non-technical words used in the teaching of science at Forms III and IV level in the Territory of Papua and New Guinea, (Gardner, 1971). In the first phase of the project, a comprehensive, relevant, non-trivial and non-technical word list was produced. Comprehensiveness was desirable so that all words likely to prove difficult would appear on the list, and was achieved by starting with the Thorndike-Lorge word list of the 20,000 most frequently used English words. Relevance was determined by having a panel of Victorian teachers and educational researchers, followed by a similar panel in Papua-New Guinea, systematically select and rate words on the Thorndike-Lorge list for their importance in science; only words regarded as essential or valuable were included for testing. Non-triviality was attained by deleting nearly all words rated as AA or A in the Thorndike-Lorge list; these words (e.g. 'blood', 'circle', 'up') occur with high frequency, denote simple concepts, and are usually learned early in the pre-school or primary school years. Non-technical words were included, and technical terms excluded, on the grounds that the latter would be specifically taught if they were part of the curriculum. Technical terms would include such things as physical concepts (mass, force, energy ...), names of chemical elements, minerals, plants, organs, processes, apparatus, etc. For the purposes of the project, it was felt that a list of words of 'ordinary English' frequently used in science teaching but seldom explicitly taught by the science teacher would be more useful. A list of 599 words meeting the various criteria was produced.

Multiple choice items testing the comprehension of these words were written, tried out, assembled into fifteen tests, and administered in T.P.N.G. These tests were then modified slightly for use in the present project, known as the *Words in Science* Project. Various types of items were employed: most require synonym recognition, others picture recognition; a few require sentence completion, or correct sentence usage. *Sample items are given in Appendix I.*

The *Words in Science* tests were administered during Term 3, 1971, to a sample of Form I, II, III and IV Victorian secondary science pupils, drawn from schools participating in the ASEP trials. Although not a strictly random sample of all Victorian secondary schools, the schools are widely representative of many types of secondary schools:- various school systems and various socio-economic areas are represented. The sample included 18 high schools, 8 technical schools, 7 independent non-Catholic schools and 6 Catholic schools. There were 270 classes in the sample; 7567 students provided data by responding to two tests drawn at random from the fifteen-test battery. The sampling procedure ensured that all classes and schools were equally represented on all tests. The procedure controls for the effect of mean differences in vocabulary ability amongst schools and classes. It also permits the difficulties of different items to be legitimately compared, even though different pupils provided the data.

A full report of the project (Gardner, 1972, in press) describes the project in detail, lists the test items used, and presents the findings of the project. Data available are the percentage correct on each item, distractor popularities, and a word list organized on the basis of difficulty level. These data are available both for the total sample, and for the four form levels taken separately. Space does not permit reproduction of all these data here; however, an alphabetical word list of all the words included for testing, the contexts in which these words were tested, and the difficulties of the items, is presented in Appendix II.

REFERENCES

Gardner, P.L. *Project SWNG: Scientific Words – New Guinea.* Report to the Department of Education, Territory of Papua & New Guinea, 1971.

Gardner, P.L. *Words in Science.* Australian Science Education Project, 1972 (in press).

WORDS IN SCIENCE

APPENDIX I: TYPES OF ITEMS

TYPE I: VERBAL RECOGNITION

(IA) Synonym recognition

The dog attacked the snake

A. The dog began to fight the snake.
B. The snake bit the dog.
C. The snake was stuck to the dog.
D. The dog ran away from the snake.
E. The dog saw the snake, but did nothing.

The bird was audible.

A. It made a noise which could be heard.
B. It was very beautiful to look at.
C. It was a long way from home.
D. It was flying very slowly.

The children talked about the composition of the brick.

A. its length, width and height.
B. its shape.
C. what it was made of.
D. what it could be used for.
E. where it was.

(IB) Instance recognition

Which of the following is mobile?

A. a mountain.
B. a bicycle.
C. a school.
D. a house.
E. a tree.

Which of the following is a sense organ?

A. the brain
B. the heart
C. the tongue
D. the legs.
E. the stomach.

Which of the following shows the numbers 1, 2, 3, 4, in a different sequence?

A. 2, 3, 4, 5.
B. one, two, three, four.
C. 5, 6, 7, 8.
D. 2, 3, 1, 4.
E. 10, 20, 30, 40.

(IC) Formal classification recognition.

Seaweed is

A. a grass growing along the sea shore.
B. a plant growing in the sea.
C. a kind of fish.
D. a tree which grows near the sea.

(ID) Functional classification recognition.

A clamp is used for

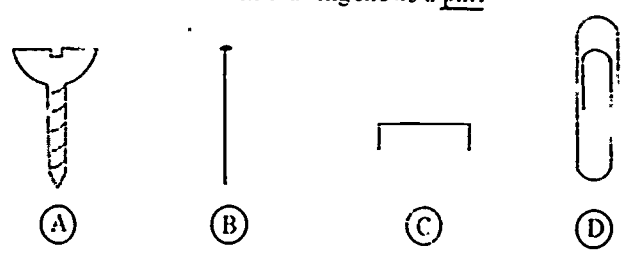
A. holding things.
B. measuring volumes.
C. weighing things.
D. measuring time.

A pump can be used to

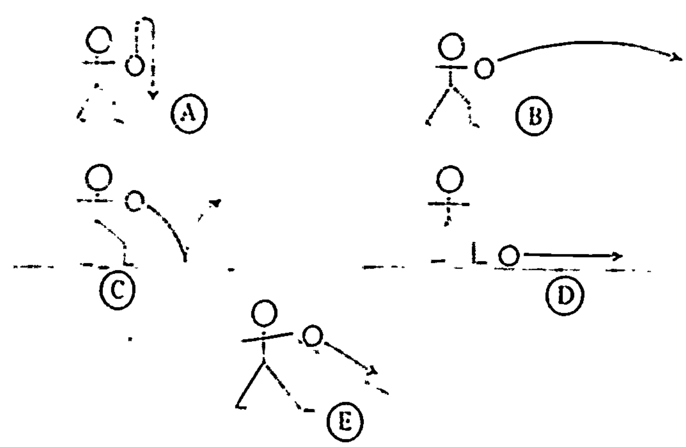
A. clean a liquid.
B. move a liquid.
C. heat a liquid.
D. weigh a liquid.

TYPE II: DIAGRAM RECOGNITION

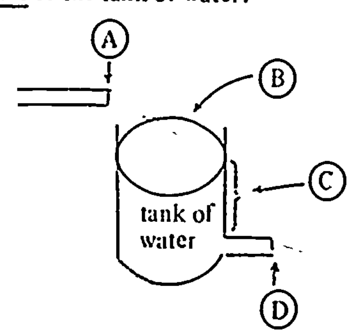
Which drawing shows a pin?



Which drawing shows a ball bouncing?



Which part of the drawing (A, B, C or D) shows the outlet of the tank of water?



TYPE III: CONTEXT USAGE

For Questions ____ to ____ you are given a word, and then given some sentences which use the word. Pick out the sentence which uses the word correctly (Example given).

Which sentence uses the word aware correctly?

A. The man in the shop charged Jim one dollar too much, so Jim was very aware of the man.
B. When Bill was old enough, he went aware from school to find a job.
C. Jack was aware of a noise coming from the trees behind him.
D. Tom looked very aware in his brightly coloured clothes.
E. The teacher said that Dick was very aware for doing such good work.

Which sentence uses the word clue correctly?

A. The letter in the dead man's pocket gave the police a clue about the reason for his death.
B. The teacher told his science class to look at the clues running around the box.
C. The hot gases passed through a clue in the side of the building.
D. The clue became hard after an hour, and the pieces of wood were stuck together.

Which sentence uses the word convenient correctly?

A. Mary found it convenient to do her homework at night, when the house was quiet.
B. The cook made some soup which tasted very convenient.
C. The books were sold to the children; the books did not cost much and so the children were convenient.
D. Bill was sick, but after a week he felt convenient again.

TYPE IV: SENTENCE COMPLETION

Questions ____ to ____ are different to the earlier questions. There is no underlined word. You still have to choose the best answer. (Example given)

Jim's eyes hurt after looking at the sun, because of its:

A. clarity.
B. brightness.
C. similarity.
D. visibility.
E. permanence.

At school, Jim is always thinking of new ways to do experiments. At home, he often makes things that he has thought of all by himself. This means that Jim is a boy who:

A. can think logically.
B. has persistence.
C. has imagination.
D. is sensitive.

When the plant was put inside the jar of gas, its leaves changed from green to yellow. The leaves were:

A. adjusted by the gas.
B. deflected by the gas.
C. displaced by the gas.
D. resisted by the gas.
E. affected by the gas.

TYPE V: GAP FILLING

For Questions ____ to ____ you are given some sentences. In each sentence a word has been left out. From the list of words at the right, pick the word that best completes the sentence. (Example given)

The school held __________ to see which boy could kick a football the longest distance.	A. an argument
	B. a substitution
Bill and Jack both did an arithmetic problem; they had different answers, so they had _______ over who was right.	C. a competition
	D. an effort
Mary was not very good at spelling, but she made __________ and became better.	E. an error

66

WORDS IN SCIENCE

APPENDIX II: Alphabetical Word List

Abbreviations used in following pages:

corr.sent.us.	–	correct sentence usage
diag.	–	diagram
ex.	–	example
recog.	–	recognize
sent.comp.	–	sentence completion
**	–	rates as essential
*	–	rates as valuable

Word	Percentage Correct Total	Form 1	Form 2	Form 3	Form 4	Context
ability	82	71	79	88	90	sent.comp.: some chemicals have a. to make plants grow
abnormal	90	85	87	93	95	a. rainfall = unusual
absence	79	74	78	80	85	a. of light = there was no light
absorb	95	91	94	97	99	water a.ed = soaked up
abundant	85	72	82	93	94	a. supply of pears = more than enough
accumulate	77	61	70	87	92	clouds a.d = came together
accurate	84	75	79	88	92	a. description = correct in every detail
action	60	46	57	64	76	sent.comp. = a. of gas on metal
adjacent	86	73	83	95	96	a. seats = next to one another
adjust	72	64	75	77	73	ex. of a.ing a ladder = making it longer by turning screws
adopt	89	88	82	93	92	a. a plan = think it is a good idea
adult	98	97	97	100	98	ex. of a. = 25 year-old
advantage	90	85	89	95	92	a. over others in games = always better
affect	90	87	93	88	91	sent.comp. : leaves change colour = a.ed by gas
agent	83	72	80	87	93	a. in rusting = cause of rusting
agriculture	86	75	82	93	96	a.al research = helps grow better crops
airtight	90	82	87	96	94	a. jar = doesn't let air in
algebra	64	48	58	71	81	sent.comp. : a + 2b = 6 found in book on a.
alternate	82	64	82	91	92	a.d as from captains = one, then the other etc.
analysis	79	68	78	81	92	sent.comp. : a. of rock showed . . .
ancestor	89	84	87	91	92	ex. of man's a. = grandfather
angle	88	80	85	93	95	recog. diag. with largest a.
annual	82	70	78	90	90	a. examination = once a year
aperture	77	67	73	83	87	a. = opening
apply	89	83	88	93	94	sent.comp. : truck stops when brakes a.ied
appropriate	81	70	79	91	83	sent.comp. : most a. unit of length
approximate	84	81	79	86	89	a. distance = roughly correct
argument	94	92	92	95	97	sent.comp. : a. over answer to problem
arid	75	53	72	86	92	a. = dry and bare
arithmetic	95	93	95	97	95	a. = working with numbers
artificial	95	90	94	97	98	a. pineapple drink = tastes like pineapple, but . . .
ascend	89	83	88	92	92	balloon a.s = goes up
ash	90	84	90	93	94	a. = grey dust under fire
assemble	96	93	95	98	96	a. = put together (model aeroplane)
assignment	89	80	87	95	94	teacher gave a. = job to do
assist	94	91	93	95	97	Mary a.ed Anne = helped
associate	77	70	78	78	84	malaria a.d. with tropics = often found in tropics
assume	85	76	85	91	87	sent.comp. : mother a.d that book was borrowed
attach	96	90	97	98	98	a.ed p[illegible]r to window = stuck on
attack	95	93	95	97	95	dog [illegible] [illegible]nake = began to fight
attract	92	87	90	95	96	insects [illegible]ed by light = moved towards
audible	62	45	60	69	75	bird was a. = made noise which could be heard

Word	Percentage Correct					Context
	Total	*Form 1*	*Form 2*	*Form 3*	*Form 4*	
automatic	85	78	85	88	92	clock strikes a.ally = without help from any person
available	96	94	94	97	99	tools a. = can get when needed
average	52	35	54	60	61	recog. a. of three given numbers
avoid	91	89	93	90	94	sent.comp. : driver a.ed hitting tree
aware	92	88	92	93	94	corr.sent.us. : a. of noise
balance	93	91	93	96	94	recog. diag. of two blocks b.ing
balloon	99	99	99	99	100	recog. diag.
basic	93	87	93	96	97	ex. of basic thing = eating
basin	95	93	97	97	94	use of b. = wash hands in it
behave	93	90	91	94	96	sent.comp. : animal b.d. in a funny way
bench	98	97	97	99	98	b. like a table
bend	97	97	95	99	98	recog. diag. of object with b. in it
bind	94	86	95	97	98	b. hands = tie together
blank	97	95	97	99	98	recog. diag. of b. box
bounce	85	79	87	87	89	recog. diag. of ball b.ing
boundary	85	76	84	91	91	recog. b. of school from diag.
brake	97	96	96	97	98	b. on truck = stops it
breed	89	82	86	93	95	b. mosquitoes = grow and multiply
breeze	92	89	93	94	94	b. = light wind
brightness	92	91	94	90	94	sent.comp. : eyes hurt because of sun's b.
brittle	86	75	86	92	94	glass is b. = breaks easily
bubble	96	93	95	100	98	b.s in water = small amounts of air
bud	92	89	91	94	94	recog. b. grows into leaf or flower
bulb	90	84	92	95	91	recog. diag. of plant b.
bump	89	88	86	92	93	ex. of b. = two marbles hit
calculate	84	68	83	92	96	boy c.d volume = found out by arithmetic
camouflage	91	85	88	95	95	c. man in jungle = wear same colour clothing
canal	92	83	93	97	95	c. = tube or passage for liquids
cancel	89	83	91	88	95	recog. ex. of arithmetical c.ling
candle	97	95	98	98	99	recog. diag.
capable	84	76	84	92	85	sent.comp. : dogs c. of biting
capacity	91	82	88	96	97	c. of tank = how much liquid it holds
capture	92	89	92	96	93	c.d gas = caught it and stopped it escaping
caution	91	86	92	92	95	add liquid with c. = carefully
cave	90	81	89	95	94	c. = large hole in ground
centre	97	95	98	96	100	recog. c. of circle from diag.
century	95	90	94	98	98	c. = 100 years
chalk	98	97	99	99	97	use of c. = write on blackboard
characteristic	83	73	79	88	91	eggs had c. smell = unlike any other
chest	98	96	98	99	98	which in c. ? = heart
chew	92	88	90	96	95	mouse c.ed paper = tore with teeth
chip	88	82	88	89	93	c. of rock = piece size of fingernail

Word	Percentage Correct					Context
	Total	*Form 1*	*Form 2*	*Form 3*	*Form 4*	
clamp	90	80	90	95	96	c. used for holding things
classify	93	90	91	96	95	rocks c.ied = put in groups
climate	88	80	89	89	93	c. of Port Moresby = hot and wet
clue	94	91	96	93	96	corr.sent.us. : letter gave police a c.
coarse	85	76	84	92	92	tool was c. = rough
coil	93	89	91	97	96	recog. diag.
coincide	80	60	77	90	93	two things c. = happen at same time
collapse	93	88	94	95	96	building c.d = fell down
collection	98	98	97	99	100	rock c. = pieces of different rocks
collide	95	93	95	96	98	bird c.d with three = flew and hit
column	69	63	63	73	79	recog. c. of numbers in a table
combination	92	86	91	94	97	c. = putting things together
commence	92	86	91	96	95	c. a test = begin
comment	95	90	96	96	99	made a c. = said or wrote something
communicate	95	94	95	94	96	some animals c. = understand sounds made by others
compare	93	89	91	94	98	c. two trees = say how alike
competition	96	96	96	94	99	sent.comp. : football kicking c.
complex	68	52	64	77	82	TV is c. equipment = many parts
complicate	95	94	91	97	98	ex. of c.d job = building a school
component	69	49	63	76	87	c.s of a car = parts of which it is made
composition	70	49	61	81	91	c. of brick = what it was made of
compress	90	83	88	93	95	c. = push together and make smaller
concept	52	32	45	61	74	no c. of cat = no idea of what a cat looks like
conception	53	31	44	63	74	sent.comp. : baby grows after c.
conclusion	85	76	86	88	91	given description of experiment, recog.c.(boiling pt.=90^{0})
conical	74	62	72	79	83	recog. diag. of cone
consecutive	78	68	78	81	88	ten c. days = every day for ten days
consist	94	88	93	97	91	cloth c.s of wool and cotton = made of
consistent	81	68	79	87	92	results c. = all about the same
constant	84	71	79	92	96	weight c. = stayed the same
constituent	59	44	51	63	81	c. of a cake = one of things from which made
construct	92	86	93	95	95	c. a school = build it
consume	67	47	62	76	85	bird c.d bread = ate it
contact	93	87	92	97	96	recog. diag. of two rods in c.
container	86	80	84	88	92	which not a c. ? = a table
contents	92	87	92	95	96	c. of a jar = water
continent	67	54	64	71	79	Australia a c. because . . .
continual	93	89	91	96	94	c. noise = goes on all the time
contract	62	39	59	70	80	metal c.ed = become smaller
contrast	58	51	53	60	68	teacher c.ed trees = said how they were different
contribute	76	58	72	86	89	salt c.s to rusting = one of causes
convenient	92	83	91	97	97	corr.sent.us. : c. to do homework at night

Word	Percentage Correct					Context
	Total	*Form 1*	*Form 2*	*Form 3*	*Form 4*	
convention	75	61	72	82	85	c. for women to cover faces = usual
converge	75	58	74	81	87	lines c. = get closer together
converse	39	38	39	38	41	c. of a statement : if acid, then litmus red
convert	84	72	84	88	91	c. something = change it into something else
co-operate	94	91	94	96	97	two men c.d = word together
co-ordination	89	81	89	92	94	police and army c.ed = worked well together
cord	92	87	93	94	93	use c. to tie up dog
core	90	85	93	92	90	c. of pineapple = middle part
cork	98	98	97	99	99	recog. from diag.
correspond	49	35	44	58	61	corr.sent.us. : fingers on hand c. to toes
crack	98	97	97	98	99	recog. diag. of block with a c. in it
creation	93	88	93	95	98	c. = making something new
crest	66	50	66	72	74	c. of a hill = top
criticise	75	66	72	79	83	sent.comp. :'drawings good, writing not' - c.ing work
crude	65	42	62	71	87	c. sulphuric acid = not very pure
crush	95	92	97	95	97	c. a cigarette by = standing on it
crust	93	87	92	95	97	c. = hard outside layer
cultivate	86	76	84	92	92	c.d. tomatoes = planted and looked after
cure	86	82	85	88	90	medicine c.d man = made well again
curve	97	96	96	98	98	recog. diag. of a c.d line
damage	94	91	97	96	90	something d.d = has something wrong with it
damp	95	92	97	96	95	paper d. = a bit wet
decay	81	72	79	83	90	ex. of d. = dead dog
decimal	93	89	93	93	95	ex. of d. = 1.2
decrease	91	80	92	96	97	d. wages = less than before
define	85	74	78	90	97	d. a word = say what it means
definite	95	93	94	96	96	d.ly going swimming = sure
deflect	76	64	75	80	83	ex. of ball d.ed = rolls, hits wood, changes direction
degree	78	66	78	84	85	distinguish correct, incorrect use as a unit
demonstrate	94	91	93	96	97	teacher d.d experiment = showed how
dependent	92	85	92	95	99	animal d. on plants = must have plants
depth	89	82	90	90	95	d. of harbour
derive	80	74	73	84	91	sent.comp. : medicine d.d from tree
descend	85	69	85	92	93	d.ed a mountain = climbed down
descendant	48	32	47	52	62	man's d.s = children and grandchildren
description	91	84	93	95	92	d. of rock = saying what it looks like
design	97	96	96	97	99	d.ed a house = drew a plan
detect	73	61	71	79	83	sent.comp. : just hear a noise = d. the noise
develop	96	94	95	99	99	seeds d.ed = grew into plants
device	81	67	81	87	88	corr.sent.us. : hammer a d. to knock nails
devise	68	53	65	69	84	d.ing an experiment = thinking about how to do it
diagnose	83	69	79	90	95	d.ing a disease = finding out what it is

Word	Percentage Correct					Context
	Total	*Form 1*	*Form 2*	*Form 3*	*Form 4*	
diagonal	79	69	75	85	86	recog. diag. containing a d.
diagram	97	94	97	98	99	d. of a fish = drawing
dial	91	88	91	92	93	ex. of something with d. = clock
diameter	70	48	67	81	86	recog. diag. of circle with a d. drawn
dimension	88	77	86	93	95	d.s of a box = length, width and height
disc	87	83	85	92	88	ex. of d. = one cent coin
discuss	92	84	92	95	97	d. something = talk about it
disintegrate	26	18	22	27	38	boiled meat d.s = breaks into pieces
disperse	77	62	74	85	90	gas d.d = spread out
displace	73	67	72	73	82	liquid d.d air = pushed out
dissolve	95	90	95	97	99	ex. of d. in water = sugar
distinct	85	71	85	90	93	veins of leaf d. = clearly seen
distinguish	96	93	95	97	99	d. two stamps = could tell they were different
distribute	88	79	85	91	96	d. chemicals = give them to different people
disturb	88	89	89	87	88	corr.sent.us. : d. bird, sleeping man, bottle on table
diversity	62	49	58	66	78	recog. set of diag. having greatest d.
dominant	71	51	67	81	88	d. = stronger than
doubt	90	82	91	94	95	d. = don't really believe
doubtful	93	90	94	97	90	d. about swimming in race = did not know if in it or not
drain	91	87	90	93	94	ex. of d.ing something = sink of water
drought	90	85	88	95	93	d. occurs = no rain for long time
duplicate	94	89	93	98	96	d. notes = make a copy of
edible	86	77	85	91	93	coconuts e. = can be eaten
effect	56	45	46	62	70	sent.comp. = heat is an e. of electricity
efficient	61	48	56	68	72	corr.sent us. : large brooms more e. for sweeping
effort	94	93	94	94	94	sent.comp. : made e. and got better
elastic	83	74	86	86	86	sent. comp. : string described, is e.
emit	62	46	58	65	79	recog. diag. of star e.ting radio waves
enable	76	62	76	83	83	corr.sent.us. : aeroplane e.d man to reach town
enlarge	95	93	93	96	97	picture e.d = made bigger
equipment	93	89	93	95	96	school boy's e. = pen, pencils, ruler
equivalent	92	88	92	93	97	Australian, N.Z. dollar e. = the same as
erect	86	75	85	91	93	recog. diag. of e. flower
error	94	92	93	95	97	recog. ex. of arithmetical e.
essential	75	63	71	82	86	sodium e. to life = must have sodium
estimate	81	76	76	82	89	e.ing length of string = guessing
evacuate	87	78	85	94	93	e. a glass jar = empty it completely
exact	82	70	85	88	85	recog. bar which is e.ly equal to length of ruler (diag.)
exception	89	79	87	93	96	all boys with good results, with the e. of . . .
excess	83	70	79	90	94	builder has e. wood = more than needed
excite	79	69	77	80	89	muscle e.d = made to move quickly
exclude	83	68	79	91	95	e. water = not allow water in

Word	Total	Form 1	Form 2	Form 3	Form 4	Context
	Percentage Correct					
exert	77	64	68	88	87	sent.comp. : pushing a car = e.ing a force
expand	92	89	91	93	96	metal e.ed = became bigger
expel	84	77	82	87	93	e. a gas = push it out
experience	92	88	93	92	97	my e.s = what has happened to me
explode	90	86	88	93	94	chemicals e.d = loud noise, broke beaker
expose	81	70	76	89	90	paper e.d to light = light reached it
exterior	94	91	92	97	96	e. of house - outside
external	73	53	68	82	90	e. skeleton – outside
extinguish	94	93	93	96	96	e. a fire = put it out
extra	90	87	88	93	91	insect with e. leg = seven instead of six
extract	95	93	97	96	96	e.ed oil from leaf = took it out
extreme	88	79	88	92	93	sent.comp. = temperature of 120^{o} is e.
factor	58	36	52	67	76	sent.comp. : drinking drivers one f. in causes of accidents
feather	97	96	97	98	96	which covered with f.s ? = bird
female	98	97	98	99	98	ex. of f. = mother
film	97	95	98	97	98	f. in camera = roll of paper in back
film	74	64	72	75	84	recog. diag. of f. of oil on water
fin	79	82	79	81	75	recog. from diag. of fish
flexible	95	93	95	95	96	f. = bend without breaking
float	95	93	97	95	97	block f.s = stays on top of water
fog	75	66	79	75	81	f. in a jar = thick gas
formation	93	88	93	96	96	f. of rocks = how made and put together
frog	99	99	99	99	99	recog. diag. of f.
function	93	88	94	96	95	f. of stomach = work it does
fundamental	67	53	66	73	75	f. laws of physics = most important
gap	95	92	95	97	95	g. in clouds = space between
generate	80	67	81	86	87	machine g.s steam = produces steam for use
globe	91	88	91	91	95	ex. of g. = ball
grain	38	31	35	39	48	g. of rock = how big the pieces in it are
grind	86	79	85	89	90	to g. corn = rub it between stones
hatch	85	82	86	84	87	h. an egg = keep warm until chicken comes out
hook	97	97	98	98	97	recog. diag.
horizon	91	85	91	92	95	sun on h. = when it rises and sets
hygiene	93	88	92	97	98	h. = how to keep clean and healthy
identical	90	85	87	94	94	ex. of i. = two nails made by same machine
identify	90	82	90	93	95	i. a leaf = tell which tree it is from
ignore	96	92	95	97	98	i.d what was said = did not listen to it
illuminate	73	49	71	84	89	i.ion = light
illustrate	65	56	59	70	76	sent.comp. : 6 + 3 = 9 is an i.ion of the rule . . .
imagination	85	78	82	88	91	sent.comp. : boy thinks of new ideas etc. has = i.
immerse	69	52	62	80	85	i. an object = cover with liquid
impact	79	66	77	85	87	ex. of i. = truck crashing into tree

Word	Percentage Correct					Context
	Total	*Form 1*	*Form 2*	*Form 3*	*Form 4*	
incident	53	42	54	54	62	light i. on metal = hit the metal
incline	54	47	47	61	63	recog. diag. of i.
index	83	78	84	81	89	recog. given extract of book as i.
indicate	92	85	91	94	96	sent.comp. : colour i.d water there
industry	93	87	93	97	96	i.ial town = one with factories
infinite	95	92	94	94	98	i. list of numbers = goes on for ever
inflate	82	77	83	85	83	i. a ball = blow air into it
influence	76	63	73	79	90	sent.comp. : i. of magnet on nail
information	97	95	98	99	98	ex. of giving i. = man on radio reading news
inhabit	89	84	84	91	95	i. a forest = live in it
inhale	93	88	94	96	96	i. smoke = breathe it in
initial	47	35	45	47	64	i. step of experiment = thing done first
inject	94	91	93	97	96	doctor i.ed medicine = use of needle
injure	94	91	93	97	96	i. hand = hurt it
inquiry	92	86	92	93	97	making i. = asking questions
insert	92	88	90	95	93	i. something = put it inside something else
instant	93	91	93	94	95	do something i.ly = as soon as asked
instantaneous	87	77	84	92	96	i. reaction = very quick
instrument	91	88	88	94	94	recog. that book is not an ex. of i.
intake	82	71	82	85	92	sent.comp. : 6 pints is large i. of water
intelligence	90	89	89	91	94	dog is i.t = learned to do many things
interfere	93	91	93	95	94	storm i.d with radio = did not allow to be heard clearly
interfere	89	84	87	90	94	chemicals i. with plant growth = slow down or stop
internal	88	79	86	93	94	ex. of i. part of body = heart, stomach, bones
interpret	71	56	67	77	83	i. results = explain the meaning
intersect	94	89	94	98	95	recog. diag. in which lines i.
interval	94	88	96	96	97	corr.sent.us. : five minute i. between races
introduce	96	94	95	97	98	i. class to science = first lesson in science
invent	90	89	88	95	91	i. a machine = first person to think of it
invert	61	45	54	68	79	i. a beaker = turn upside down
investigate	97	96	98	98	98	people making i.ion = trying to find out things
involuntary	82	67	80	89	92	i. movement of leg = could not control it
irritate	80	69	77	89	86	sent.comp. : smoke i.d eyes
isolate	82	63	77	90	97	mouse was i.d = alone
jar	98	98	98	98	99	recog. diag.
jaw	96	94	96	98	98	recog. from diag.
juice	98	98	95	99	98	tomato j. = liquid of tomato
junction	82	74	78	87	92	recog. diag. of j. of two lines
kit	94	90	92	96	96	tool k. = set of different tools
label	95	95	93	96	98	recog. l. of bottle from diag.
laboratory	94	91	94	97	96	l. = place where experiments done
latitude	45	35	45	47	52	recog. that two points on diag. of earth have different l.s

Word	Percentage Correct					Context
	Total	*Form 1*	*Form 2*	*Form 3*	*Form 4*	
law	60	39	51	72	78	recog. ex. of scientific l.
layer	87	77	83	95	95	recog. diag. of something displaying l.s.
leaf	88	86	88	87	91	gold l. thin sheet
leak	91	87	86	94	95	l. in tank of water hole
level	96	95	96	96	97	l. piece of ground – flat
level	85	79	84	91	87	given diag., recog. that two surfaces at same l.
liberate	75	54	71	86	90	gas l.d = given off
lightning	93	92	91	96	94	l. when there is a storm
limit	82	69	81	86	91	speed l. 40 mph – not more than 40 mph
linear	66	52	59	77	80	recog. diag. of l. graph
liquid	95	95	93	97	97	recog. milk as ex. of l.
locate	94	93	93	96	96	l. a house = find
logic	71	61	69	70	78	sent. comp. : l.al thinking
lubricate	59	37	45	70	84	l.s an engine = puts oil in
magnitude	72	50	74	85	80	m. of a tree = how big it is
majority	91	86	88	96	95	m. of people voted = more than half
male	95	93	95	97	96	recog. father as ex. of m.
manufacture	84	74	82	91	91	recog. that wood is not m.d
margin	92	87	92	94	96	piece of metal with extra m. = slightly larger piece
match	93	91	94	94	94	sent.comp. : socks and tie m.ed well
mate	90	84	90	92	95	m. for male rat = female rat
mathematics	96	95	96	99	97	m. = study of space and number
mature	85	74	82	92	92	ex. of m. = 30 year old man
maximum	77	63	74	84	90	m. temperature 75^0 = not above 75^0
measure	88	87	88	88	89	sent.comp. :m.d length of a table
medium	90	87	90	89	94	given information, identify m.-sized packet of cigarettes
method	90	82	87	95	95	m. of experiment = way to do it
migrate	94	90	93	94	98	m.s to Australia = makes it new home
mild	85	75	85	90	89	m. reaction = not very much happened
milk	95	93	96	93	96	recog m. as ex. of something people drink
minimum	70	52	67	82	81	m. temperature yesterday 75^0 = fell to 75^0, but not below
minus	91	90	86	91	95	6 m. 2 = 4
mirror	98	97	98	98	98	use m. = look at face
mobile	92	87	90	96	97	ex. of m. = bicycle
model	95	95	92	95	97	m. of aeroplane = looks real, only smaller
modify	70	57	61	76	87	m. an aeroplane = make changes to it
moisture	91	89	89	94	94	soil contains m. = water
molten	79	64	75	86	92	iron m. = hot and liquid
mud	93	94	91	96	93	m. = soil and water
multiple	79	70	72	85	90	m. car smash = more than two cars
multiply	91	86	89	94	97	kangaroos m. = more and more of them
naked	72	65	68	75	80	n. flame = not covered

Word	Percentage Correct					Context
	Total	*Form 1*	*Form 2*	*Form 3*	*Form 4*	
narrow	96	95	94	96	97	recog. diag. of n.est container
needle	97	99	98	98	94	recog. that n. is thin and sharp
negative	62	39	61	73	77	n. temperature = colder than $0^{0}C$
neglect	88	78	86	95	92	n. weight of metal = take no notice of
negligible	61	52	68	62	73	rainfall n = very little rain
neutral	70	53	67	78	82	petrol is n. = not acid, not alkaline
normal	93	91	92	96	95	n.ly goes to school = usually goes
nought	97	97	95	98	98	recog. 6 − 6 = n.
nourishment	96	94	97	98	93	people need n. = enough to eat and drink
obey	97	97	98	99	94	dog o ed = did as he was told
observation	82	74	77	92	88	sent.comp. : what was seen in an experiment = o.
obvious	77	67	73	83	84	ex. of something o. = sun rises every day
occasiunal	93	89	93	98	92	o. noise = sometimes
occupy	94	91	94	98	94	gas o.ied jar = took up all the space
occur	87	84	83	90	93	malaria o.s in tropics = happens
udd	93	94	91	98	91	results o. = unusual
odour	93	92	93	96	93	o. = smell
omit	64	42	59	74	81	o.ted to put water in = did not put water in
onion	99	99	98	98	99	recog. diag.
operate	91	88	92	95	90	o. a machine = switch it on and make it go
opinion	89	83	87	93	92	sent.comp. : "which do you like better?" = asking o.
oppose	76	68	71	84	82	recognize diagram of forces o.ing each other
origin	79	64	78	87	87	o. of life = how life first began
outlet	92	88	91	94	97	recog. o. of tank of water from diag.
outline	71	56	70	80	80	o. a problem = tell quickly what problem is
oval	97	97	96	97	97	recog. diag.
overcume	90	85	92	95	90	o. in fight = was stronger
overhead	89	88	91	90	88	ex. of something u. = aeroplane high in sky
parallel	94	91	94	98	92	recog. diag. of p. lines
partial	54	39	51	[illegible]	68	sent.comp. : p. explanation of failure at exam.
particle	87	82	84	94	86	p. of glass = very small piece
partner	95	94	95	98	93	p. in experiment = someone to work with
pattern	92	87	90	97	95	recog. diag. of o's and x's arranged in a p.
pea	96	96	97	98	93	recog. that p. is green and round
peak	96	94	94	98	97	recog. p. of mountain from diagram
pencil	97	97	98	98	94	recog. that p. used to write with
penetrate	86	76	84	91	94	spear p.d leg = went into leg
per	88	88	86	88	91	recog. that "births/year" is read "births per year"
percentage	49	41	51	49	55	half written as p. = 50
permanent	84	76	84	89	90	colour changed p.ly = stayed, not changed back
perpendicular	70	47	67	84	81	recog. diag. of p. lines
pest	95	96	96	96	93	recog. that farmer would try to kill p.
phenomenon	67	56	59	73	81	sent.comp. : . . . this p. is called an eclipse
pierce	91	87	90	95	91	p.d a tin = made a hole in it

Word	Percentage Correct Total	Form 1	Form 2	Form 3	Form 4	Context
pin	97	97	97	95	100	recog. diag.
pivot	88	82	87	91	92	recog. p. of seesaw from diag.
plot	80	55	81	90	95	teacher p.ted graph marked in points, connected them
plug	97	94	97	97	98	use of p. = stop liquid running out
plus	98	98	96	99	100	recog. 6 p. 2 = 8
poison	99	99	98	99	100	if man drinks p. - die
pond	98	97	98	98	99	p. in garden = put fish in it
positive	73	64	72	78	81	test for malaria p. = has malaria
positive	78	42	86	92	92	recog. + 3 as ex. of p. number
practise	92	87	91	98	93	boy p.d using a machine : to learn to use it better
precaution	92	88	90	95	95	list of p.s - things to be careful about
precise	95	91	96	97	96	recog. ex. of p. measurement
predict	85	71	85	93	94	p. something = say something will happen later
preparation	79	74	78	85	82	p. for an experiment = getting the materials
presence	85	74	84	92	92	p. of salt in water = there was salt in the water
preserve	94	93	92	96	96	meat p.d = kept for long time
previous	77	64	72	84	89	p. day = day before
primary	73	63	71	76	83	p. cause of death = recog. from given information
primitive	76	67	65	86	86	sent.comp. : aeroplanes were once p.
probability	63	52	58	68	76	recog. event with highest p.
procedure	89	81	88	94	94	p. of experiment = how to do it
process	79	66	76	84	91	sent.comp. : plants make food: p. uses water
profile	81	67	81	88	90	recog. diag. of p. of man's head
proof	85	77	87	87	90	p. = show that statement is true
propagate	69	67	69	70	71	sound p.d = carried along
propel	89	84	90	90	93	fish p.s itself by tail = uses tail to move through water
proportion	71	65	67	74	81	given data, recog. p. of water in body
pump	80	69	81	88	84	use of p. = move liquid
purify	89	81	92	92	93	liquid p.ied = made clean
rainbow	95	94	96	95	96	recog. diag.
random	25	12	18	36	35	recog. ex. of r. motion = fly flying inside a box
rapid	94	93	94	96	95	car travelled r.ly = very quickly
rare	94	91	94	96	95	r. animal = not many of that kind
rat	97	96	97	98	98	recog. diag.
rate	41	25	39	47	58	given data, recog. death r. in a village
recoil	65	52	63	71	75	snake r.ed = moved back quickly
reference	80	67	77	85	93	r. = name of book on the topic
refine	84	76	84	88	90	r.ing sugar = change into clean, white sugar
reflect	70	62	62	79	80	recog. diag. of light being r.ed
refrigerator	98	97	99	98	99	use of r. = keep things cold
regular	76	69	76	76	85	sent.comp. : church bell rings r.ly
regulate	67	50	61	75	82	corr.sent.us. : accelerator r.s speed of car

Word	Percentage Correct					Context
	Total	*Form 1*	*Form 2*	*Form 3*	*Form 4*	
reject	93	89	93	96	97	people r.ed money = would not take
relationship	89	79	89	92	94	corr.sent.us. : r. between smoking and illness
relative	60	41	58	68	74	diag., and sent.comp. : Ben is small, r. to Charles
relevant	78	66	75	84	88	distinguish r. from irr. statements about car accidents
reliable	92	88	91	94	95	measurements r. = made carefully
repel	75	58	71	83	88	wire r.led by metal rod = pushed away
replace	93	88	94	94	95	Joe r.d Bill in race = ran instead of
represent	75	59	74	83	86	Mr. J. r.d Bill at meeting = went in place of
research	93	86	94	96	96	r. in motor car factory = trying to find new things
residue	70	57	65	79	81	recog. r. (salt) in a process described
resist	73	55	72	79	90	air r.ed movement of balloon = tried to stop
resource	78	64	76	86	86	r.s of school for teaching = things teacher can use
respond	82	72	80	87	90	sent.comp. : insect r.ing to the light
retard	80	62	80	87	90	chemicals r. growth = slow down
reverse	95	91	96	98	97	r. a car = go backwards
revise	45	33	40	50	58	r. book = make changes to it
revolve	95	90	95	98	97	insect r.d around light = flew in circles
right angle	91	86	89	95	96	recog. diag.
rigid	74	60	68	82	85	recog. that i on bar is r.
rim	88	81	86	94	92	recog. r. of cup from diag.
ripe	95	95	94	96	95	r. banana = ready to eat
ripple	92	90	88	95	93	r. in water = little waves
rod	95	91	95	97	95	recog. diag.
root	98	96	98	99	99	recog. r. from diag. of plant
rot	96	92	96	98	98	apple r.s = goes soft and smells bad
rotate	90	83	91	95	93	recog. wheel of car as ex. of r.
row	53	53	49	55	57	recog. r. of numbers in a table
rub	83	71	83	90	88	recog. that piece of iron could be r.bed with wool
rubber	95	93	96	96	97	recog. uses of r.
rule	83	78	84	83	88	sent.comp. "Don't put water into acid" = r.
ruler	97	95	97	98	97	r. used to measure length
sample	96	92	98	98	95	s. of rocks = get some small pieces
scale	87	76	85	95	94	recog. s. of instrument from diag.
scent	98	96	98	98	98	flower s. = smell
scientist	95	94	95	96	94	s. person who = does experiments, finds out why things happen
scratch	91	88	88	95	93	to s. wood, use pin
screen	89	87	89	90	92	recog. s. from diag.
screw	97	96	97	98	98	recog. diag.
sea weed	91	88	92	93	94	s. = plant growing in sea
section	72	61	65	76	86	to get s. of leaf = slice with razor blade
select	93	88	92	96	96	s. a book = pick out a certain book
semicircle	90	85	89	93	93	recog. diag.

Word	Percentage Correct					Context
	Total	*Form 1*	*Form 2*	*Form 3*	*Form 4*	
sense	58	47	53	65	68	ex. of s. organ = tongue
sensitive	74	64	68	83	82	paper s. to light - changes in light
sensitive	82	75	79	86	88	beam balance s. instrument = can weigh small things
sequence	65	57	63	63	80	ex. of 1, 2, 3, 4 in different s. - 2, 3, 1, 4
separate	94	93	93	96	96	to s. salt and sand = put in different places
shallow	88	83	86	91	93	recog. diag. of s.est container
shelf	97	97	97	97	97	use s. = put bottles etc. on
shell	94	92	94	97	96	s. of egg = outside coating
shift	94	93	94	95	95	s. a car = move it
shrink	95	96	94	94	97	sock s.s = gets smaller
sign	87	79	87	90	92	sent.comp. : + and - are positive and negative s.s
signal	93	91	94	94	94	recog. situation in which people are making s.s
significant	89	83	86	93	94	given data, recog. most s. cause of car accidents
similar	95	94	94	96	97	my house is s. to yours = looks like yours
simplify	86	78	86	91	91	s. a problem = make it easier
simultaneous	67	54	66	71	77	two things happen s.ly - at th same time
sink	89	86	91	91	88	recog. diag. of block s.ing
skill	95	90	96	98	95	s.ed in cooking = cooks very well
slide	98	98	99	97	98	recog. diag. of something s.ing
slip	97	96	98	97	99	animal s.ped out of hands = could not hold onto
slit	96	96	97	96	97	recog. diag. of paper with s. in it
slope	95	92	96	96	98	recog. s. of hill from diag.
smash	98	98	99	97	97	s. a house = break into pieces
smear	84	74	80	89	93	ex. of s.ing paint = put drop on window and rub
smell	98	98	99	97	99	to s. something use = nose
soak	97	94	97	97	98	s. block of wood = put in water
soap	99	98	100	98	99	use s. = to wash yourself
solve	95	93	92	96	97	s. a problem = find the answer
sour	96	93	96	97	97	ex. of s. = lemon
source	89	79	89	94	94	recog. s. of light from diag.
spark	92	88	91	95	94	recog. that s.s jump from a fire
spill	92	91	93	92	94	s. water = knock over, water runs out
spin	97	97	96	99	97	s.ning = turning round and round
spiral	84	76	82	89	89	recog. diag.
splint	90	87	90	92	92	s. of wood = long, thin, straight
spontaneous	31	15	26	3[illegible]	45	s. burning = starts without being set alight
spoon	98	98	98	98	99	recog. diag.
spring	97	96	97	97	98	recog. diag. of helical s.
spring	94	89	9[illegible]	98	96	s. time = between winter, summer
stable	92	85	93	95	94	temperature s. = not changing
stagnant	74	63	69	79	87	water s. = not moving, going bad
stain	80	64	81	86	90	s. cells = colour them

Word	Percentage Correct					Context
	Total	*Form 1*	*Form 2*	*Form 3*	*Form 4*	
stalk	93	92	92	93	96	recog. from diag. of plant
standard	60	41	50	72	76	sent.comp. : 1 yard metal bar = s. of length
starve	95	91	98	96	97	s. if = not enough to eat
stationary	92	85	93	95	95	car s. = not moving
steady	91	87	93	91	94	s. speed = not changing
steam	96	94	96	95	97	s. made when = water boiled
steep	84	82	80	84	89	recog. diag. of s.est hill
stem	96	93	98	98	97	recog. from diag. of plant
stimulate	55	43	48	57	71	medicine s.d the heart = beat faster, more strongly
sting	91	86	90	95	94	sent.comp. : insect biting hand can s.
strucutre	93	88	92	97	96	s. of a house = parts and how put together
submerge	78	70	76	82	85	diver s.d in water = went down into water
substance	92	87	96	94	94	recog. ex. of s. = petrol
substitute	79	64	76	87	88	s. hydrochloric acid = use it instead of another
subtract	96	93	95	99	95	s. 2 from 10 = 8
successive	89	79	88	93	96	ex. of s. days = Tuesday, Wednesday, Thursday
suckle	73	63	69	81	78	cat s.s kittens = feeds them milk from her body
summary	91	85	89	95	98	s. of story = most important things
survive	96	94	98	98	97	animal s.d = stayed alive
suspend	76	62	72	84	87	recog. diag. of s.ed ball
sweat	95	92	96	97	96	if man s.ing = skin wet
switch	96	94	97	97	98	recog. s. on diag. of radio
symbol	75	54	73	84	91	recog. s. for oxygen
symmetrical	41	23	35	48	60	recog. ex. of s. figure
system	79	70	77	85	84	sent.comp. : water movement s. in trees
systematic	89	78	93	94	93	boat built s.ally = followed a plan
table	89	89	88	95	95	recog. ex. of a t. of numbers
tabulate	79	75	78	89	93	results of experiment t.d = written down on list
tank	55	43	56	56	65	recog. diag. of t. of water
tape	97	96	97	99	97	recog. diag. of paper t.
target	95	94	95	96	95	ball thrown at wall; t. = wall
technical	96	95	96	97	97	recog. ex. of t. work = make car parts
temporary	77	66	73	80	90	t. building = to be used for short time
tend	89	82	87	94	94	t. to be fat = usually fat
tension	83	75	78	87	93	wire under t. = pulled very tightly
test	94	90	95	96	95	t. a rock for copper = find out if copper in it
textbook	65	55	63	66	80	use of t. = to read and study
theory	74	67	78	78	74	sent.comp. : doctor's idea about a disease = t.
thirst	94	92	94	97	94	to stop t. = get a drink
thrust	79	69	77	83	88	t. of engine = push
thumb	95	95	94	97	96	recog. from diag.
thunder	83	82	86	84	81	sent.comp. : loud noise after flash of light in sky = t.

Word	Percentage Correct					Context
	Total	*Form 1*	*Form 2*	*Form 3*	*Form 4*	
tide	85	77	82	92	90	t. = rise and fall of sea
tight	98	97	97	98	99	recog. diag. of box with t. string
tilt	95	91	95	96	98	recog. diag. of t.ed tree
toe	96	96	95	95	97	recog. from diag.
tooth	97	98	96	98	97	recog. from diag.
topic	58	48	59	61	66	sent.comp. : light, oxygen, force = t.s
torch	96	95	96	97	97	recog. diag. of flashlight
trace	90	85	90	91	92	t. of copper in soil = small amount
transfer	97	95	96	96	99	t. something = move it from one place to another
transform	85	78	85	9[illegible]	88	iron atoms t.ed = changed into something else
transmit	90	85	84	95	96	disease t.ted = passed on from one person to another
transport	95	93	96	96	96	t. something = carry it from one place to another
tray	97	98	97	98	97	recog. diag.
treatment	77	67	75	79	86	t. of plants in an experiment = what was done to them
triangle	98	97	98	99	100	recog. diag.
trolley	94	91	94	95	97	recog. diag.
tropic	83	74	85	84	88	recog. part of Earth in the t.s from diag.
trough	89	85	86	91	95	use of t. = fill with water
trunk	96	94	98	96	98	recog. t. of tree from diag.
tube	85	77	83	91	90	recog. from diag.
tunnel	96	95	95	98	97	t. in mountain = large hole through mountain
tweezers	86	81	81	92	93	recog. diag.
twin	96	94	97	94	98	t. engines = two, both the same
twist	96	94	97	96	96	t. a lid = turn it
typical	88	77	87	93	97	t. plants = like others in that place
tyre	96	93	95	96	99	recog. diag.
underneath	96	95	96	98	97	recog. diag. of ball u. table
uniform	86	73	87	87	96	u. speed = same speed all the time
unique	82	70	82	87	90	something u. = only one of its kind
upright	94	90	93	97	96	recog. diag. of u. line
upset	92	85	94	94	96	u. a glass of water = knock it over
valid	55	30	50	65	79	said something v. = correct
variable	90	80	90	93	96	v. weather = changing
variety	93	87	93	96	97	v. of fruit = different kinds of fruit
vehicle	96	93	97	95	99	ex. of v. = motor car
vertical	76	64	72	84	87	recog. v. line from diag.
vessel	73	56	70	78	90	if farmer had v. = could put water in it
vibrate	86	80	83	91	91	hand v.ing = moving backwards and forwards quickly
vigorous	88	82	84	92	93	ex. of v. activity = running fast
violent	91	84	89	96	96	ex. of v. reaction = petrol exploding
violet	80	69	79	85	90	v. medicine = purple
visible	94	92	92	96	98	stars v. = able to be seen

Word	Percentage Correct					Context
	Total	*Form 1*	*Form 2*	*Form 3*	*Form 4*	
vision	95	92	94	96	98	boy's v. good = can see well
vital	78	61	75	85	91	some gases v. to plants = very important
vocal	84	71	83	91	94	man was v. = talked a lot.
web	98	96	97	98	100	recog. diag.
wedge	82	75	78	89	85	recog. diag. of w.-shaped object
weed	87	80	86	91	90	w.s = plants which grow wild and are not wanted
wilt	85	78	82	90	90	recog. diag. of w.ed flower
wipe	95	92	95	95	97	w. t ble = rub with cloth
zero	96	94	96	96	98	recog. from diag. of number scale

THE DEVELOPMENT OF A CREATIVITY TEST

D. Cohen

The Need for Creativity Testing

Tests comprising convergent-type items, typically seeking single responses scorable as correct/incorrect or as one point/zero, have been proliferated and are both readily available, and easily scorable with a high degree of reliability. Unfortunately, such tests reflect a failing of educators and psychometricians: what Liam Hudson (1966) has referred to as "the neglect of inconvenient evidence".

Contemporary trends involving diversification of educational objectives dictate a corresponding diversification of evaluative techniques. Voluminous literature attests to the current educational movement towards the "open classroom" – with greater emphases upon philosophical, psychological and physical freedoms (e.g., Nyquist, 1971). Concurrent increases in the open-endedness of learning experiences are heavily supported by so-called "inquiry" emphases in newer curriculum materials (e.g. those produced by ASEP, BSCS, and Harvard Project Physics). These seek to develop divergent abilities to a greater extent than previous experiences and materials.

It was in his model of the "structure of intellect" that Guilford (1959:469) first introduced a dimension of thinking labelled as "divergent". This dimension relates to the production of many and varied responses to a stimulus. Guilford distinguished divergent thinking from the convergent thinking required by conventional intelligence tests. Many writers have interpreted the two types of thinking as lying at opposite poles, but this was not intended by Guilford. Convergent thinking generates new information from known and memorized information, and leads to one right, correct, best or conventional response. Divergent thinking is most closely akin to creativity. It generally requires the production of a large number of responses but judgment and evaluation are frequently suspended.

Divergent thinking may be regarded as similar to what is sought in the socalled inquiry approach. This has great relevance in the ASEP context. Positions Document 38 of ASEP, "Use of the Inquiry Approach", indicates that ASEP is "aimed at encouraging enquiry", and also states that: 'For the student to be able to think and be creative he should be given opportunities requiring thinking and creativeness".

From a curriculum evaluation viewpoint, the need for evaluation of inquriy abilities and creativity becomes essential. Thus there exists an urgent (but unfilled) need for measures of ability to produce multiple and diverse responses, that is, for divergent tests.

Further evidence for such a need is provided by the fact that intelligence tests alone are not capable of predicting those who will make creative contributions in the arts and sciences nor of identifying those who have – even though intelligence plays a part in high-level creative thinking. For creative results in science, research by Torrance suggests a threshold effect that an IQ of 115 at least is required – but above that level, output bears no relation to IQ. Research studies have shown that while intelligence contributes to creativity, many with high IQ, while leading satisfactory professional lives, have been undistinguished creatively. There is also evidence that at this level, academic achievement is likely to be just as high among the high creatives as among the high IQs.

Personality traits such as persistence, open-mindedness and self-confidence, have been shown to contribute to professional distinction but these likewise are inadequate for prognostic purposes. Some intellectual aptitudes which are not assessed by traditional tests of intelligence appear to be involved. It is partly to provide supplementary (and indeed, alternative) evaluative data about such aptitudes and partly to dilute traditional overemphases on tests of a narrow-band of the spectrum of objectives, that the present writer has long been concerned (cf. Cohen 1957, 1963) to develop and validate new evaluative devices (e.g., devices to assess student progress in the development of laboratory performance skills (Cohen, 1966), of scientific attitudes (Cohen, 1971), and of creativity (Cohen, 1968)). There remains a desperate need to develop additional evaluative materials to match the emergent trends in educational objectives. Encouragement is derived from the need to dilute existing inadequate indices which have been undeservedly worshipped as the major indices of educational progress.

The Psychometrics – Creativity Dilemma

The aim of the present project is to develop a set of "creativity tests" – almost a contradiction in terms. In fact, the task represents a challenge to resolve an educational dilemma. In *test* development, psychometricians generally require the performance of standardised tasks which demand time constraints, and administration under precisely specified conditions. Even though meeting these conditions, the resultant test of creativity is likely to be subjected to many of the criticisms made by psychometricians of traditional educational measurement techniques. Such criticism include:

- the use of a global or umbrella term (in this case "creativity") to account for a composite of several complex interacting components;
- lack of comparability of nature or difficulty of separate items;
- non-additivity of scores derived from dissimilar items.

As if attempting to satisfy the demands of the psychometrician is not challenge enough, the researcher dedicated to the development of creativity tests is then faced with a set of contradictory demands, such as:

- creativity tasks should be open-ended;
- creativity can only be nurtured and incubated given freedom from time constraints and from specified administration conditions;
- creativity will most likely occur under anxiety-free conditions.

The above demands of psychometricians and creativity workers can be largely reconciled if the following assumptions are accepted:

(1) creativity does constitute a meaningful label for a global concept or characteristic of people, which subsumes many component parts;

(2) it is useful to evaluate creativity as a means of providing data about student performance upon educational objectives not otherwise adequately evaluated;

(3) it is possible to devise items to evaluate parts of creativity;

(4) it is valid to pool the items so devised to create a *creativity evaluation instrument;*

(5) the *administration* of creativity tests under standardized *conditions* does not invalidate the meaningfulness of the scores obtained;

(6) assigning a creativity profile to a student will be educationally meaningful and useful.

It is the purpose of this paper to examine these assumptions critically, in the context of the actual test development.

Approaches to the Study of Creativity

An examination of creativity literature, including research reports, indicates four approaches to studies, namely:

(1) The study of the *creative personality* in which there is a wide consensus of opinion about the traits, attitudes, etc. which characterise the creative personality.

(2) The study of the *criteria of creativity*, e.g., the end-products such as completed works of art, buildings, patents for inventions, theories, research results, etc.

(3) The study of the *environments* in which creativity is cultivated.

(4) The *development of tests* of creativity to assess some of its component intellectual aptitudes.

Mackinnon (1967) specified three conditions for creativity:

> "It involves a response that is novel or at least statistically infrequent. But novelty or originality of behaviour, while a necessary aspect of creativity, is not sufficient. If a response is to lay claim to being a part of the creative process, it must to some extent be adaptive to reality. It must serve to solve a problem, fit a situation or in some sense correlate with reality. And thirdly true creativity involves a sustaining of the original insight, an evaluation of it and elaboration, a sustaining and developing of it to the full."

Barron (1969) defined creativity as:

> "the ability to bring something new into existence . . . Since human beings are not able to make something out of nothing the human act of creation always involves a reshaping of given materials whether physical or mental".

Barron's definition emphasises invention via restructuring.

Hadamard (1945) described creativity via its preconditions – the first step being to produce many ideas – introducing a fluency or frequency criterion. Such a basis would be supported by some writers, for example, Parnes and Meadow (1959) who see the so-called brainstorming process as a stimulus to the generation of multiple creative ideas.

Thus fluency, flexibility and originality are highlighted as creativity components.

Devising a Creativity Test

In keeping with the above components, tests of creativity are designed to stimulate the production of as many responses as possible (fluency): tasks are open-ended (fluency and flexibility), there are no uniquely correct responses (flexibility), and inventive responses are encouraged (originality).

The technique used by Guilford for devising tests was to hypothesise the types of ability necessary and then see if they were identified by factor analysis. As a result, he isolated and named the following eight factors:

(1) *Word fluency* which refers to the ability to generate words that fulfil particular structural requirements.

(2) *Associational fluency* which is the ability to rapidly generate words that meet particular requirements of meaning.

(3) *Ideational fluency* which is the ability to generate, within limited time, ideas that will fil' particular requirements.

(4) *Expressional fluency* which is the ability to put rapidly into juxtaposition words that meet requirements of sentence structure.

(5) *Spontaneous flexibility* which is the ability to vary one's ideas over a wide range, even though not specifically called for.

(6) *Adaptive flexibility* which is the ability to vary ideas when specifically required to.

(7) *Redefinition* which is the ability to relinquish old ways of construing objects in order to make use of them for new purposes or to choose parts of objects to serve a function.

(8) *Originality* which is the ability to make responses which are unusual, only remotely associated with the stimulus, or judged to reach a level of cleverness.

(Wilson et al 1954)

These factors again highlight the fluency-flexibility-originality criteria.

Guilford subsequently revised his eight factors in hypothesising his "structure of intellect" model (1959).

Most of the investigators who have followed Guilford have not adhered to the factor analysis approach, but have adapted many of his ideas and adopted some of his terminology. His tests (Guilford & Hoepfner, 1971) have opened up a whole new area for testing and even his critics have used many of his tests as a basis for their own.

Another prolific writer and researcher on creativity, Torrance, has also used the fluency-flexibility-originality criteria as the bases of evaluation. In some cases, Torrance (1967) has added "elaboration", which he described as the subject's ability "to develop, embroider, embellish, carry out or otherwise elaborate ideas".

As might be expected in terms of their feasibility for classroom use, most creativity tests comprise group-administered items in which verbal stimuli are used to elicit verbal responses. Such an approach ignores art forms, music, and literature. The latest tests of Torrance incorporate responses to sound. Barron and Welch have devised art preference inventories in which the respondent is asked for a like/dislike rating of a variety of pictorial stimuli of varying degrees of symmetry/asymmetry and of complexity.

Many existing creativity tests are inadequate, for some of the following reasons:

- they lack validity in that they do not adequately tap the triple criteria of fluency-flexibility-originality;
- they cannot be administered in typical school or classroom conditions;
- they lack reliable scoring techniques;
- they lack reference contexts, in the form of expected performance ranges, especially in Australian circumstances;
- they rely solely on verbal stimuli and responses.

Consequently, it was decided to develop an evaluation device to attempt to overcome the above shortcomings. Although the psychometrician might argue that the *label* of the resultant instrument should suggest its contents (i.e. face validity), creativity workers would plead for anxiety-free stimuli, – and thus for the avoidance of the use of the word "test". The advice of Mehrens and Lehmann (1969:7) was: "the word *test* . . . should not be used in the title of non-cognitive measures". Thus the title of the present instrument became "How Creative Are You?".

Developing "How Creative Are You?"

The major characteristics to be incorporated into the creativity instrument include the following:

(1) *high level of construct validity*, with respect to the following grid:

	Fluency	Flexibility	Originality
Words			
Ideas			
Devices			
Designs			

Typical cues in the stems of items were "write *as many* words . . ." (fluency), "think of as many *different* ideas . . ." (flexibility), and "invent new and original devices . . ." (originality).

(2) *high level of scorer reliability*. It is a purpose of the resultant device that it will permit the comparison of creative behaviour

(a) of a number of people at the same point in time;

(b) the same person at different points in time.

The instrument could thus explore and compare the creativity growth of students exposed to specific curricula or educational environments (e.g. ASEP) in pre- and post- tests, or with the growth of students in control groups. Two parallel forms (A and B) were produced.

(3) *practicability* for use in junior secondary-school classrooms, implying that:

(a) administration details are clear, simple and unambiguous;

(b) administration of the test is feasible in terms of class group size and in terms of the typical duration of class periods (30–40 minutes) (so that the time limits for individual items had to be short);

(c) items are readable and comprehensible for intended population.

(4) respondents to the items within the instruments are to be asked to:

(a) *invent* or *construct* responses (rather than select from alternatives provided);

(b) produce a *large number* of responses;

(c) produce *unusual* and *novel* responses;

(d) produce *diverse types* of responses.

Items were written by the researcher and assistants to represent each of the cells in the creativity grid. Several of the items were adapted (with permission) from an American battery of seventy-two pencil-and-paper tests comprising the "Kit of Reference Tests for Cognitive Factors" which has been used experimentally for nearly a decade (French, 1963). This kit contains three tests each for twenty-four 'factors'.

Most items in "How Creative Are You?" were constructed in terms of perceived Australian requirements of the 1970's. In earlier forms of the test, item stem stimuli were devised to elicit invention of sound but scoring proved impossible. All other items have also been subjected to extensive trials and modifications in classrooms similar to those of the intended population. These trials led to the rejection of several items for which problems of administration or scoring became evident. In all, sixteen items or tasks were retained in each form of the test. These included three devised to stimulate responses of words, six to produce ideas, four for devices, and three for designs.

Following is a list of the grid areas for which items were devised, together with examples of the tasks required.

1. *Word Fluency:* Facility in producing words rather than ideas in response to a simple stimulus, e.g. producing all the words the subject can think of beginning with "tr-".

2. *Word Flexibility:* Facility in shifting from one class of words to another, e.g. word association is characterised by frequent shifts in the class of ideas governing the choice of words.

3. *Word Originality:* Facility in devising new words. This category is now omitted.

4. *Idea Fluency:* The facility to call up many ideas in response to instruction, e.g. writing down as many ideas as possible about a given topic.

5. *Idea Flexibility:* Facility in changing ideas from one class to another; e.g. in a task requiring the subject to suggest problems that could be presented by a glass of water, the emphasis is on describing many different types of problems not all related to the one aspect.

6. *Idea Originality:* The ability to produce uncommon ideas in response to instruction, e.g. in a task requiring the subject to think of the consequences if people were not able to tell the differences between colours, the emphasis is on unusual ideas.

7. *Devices Fluency:* Facility in producing responses when required to manipulate devices, e.g. in thinking of alternative uses for a 20 cent coin, one emphasis is on the number of responses.

8. *Devices Flexibility:* Facility in changing from one class of ideas to another when suggesting ways for manipulating devices, e.g. in suggesting alternative uses for a 20 cent coin, one scoring emphasis is on changes in the classes of response.

9. *Devices Originality:* The ability to perceive unusual ways of manipulating devices, e.g. in thinking of uncommon ways to improve a familiar object such as a broom.

10. *Design Fluency:* Facility in taking the components of a design and reproducing them many times in a design or a number of designs, e.g. in a task requiring the subject to draw designs using only circles and curves.

11. *Design Flexibility:* Facility in taking the components of a design and adding to it in different ways to produce designs, e.g. a task requiring the subject to incorporate two small arcs of a circle in a design.

12. *Design Originality:* The ability to take components of a design and combine them in new and unusual designs, e.g. in a task where the subject is presented with four different shapes to combine into designs.

The scoring of the design items (i.e., Categories 10, 11, and 12) is being validated through a separate study. A sample of student responses to the relevant items is being presented to a jury of creative artists, art teachers, interior designers and architects. The jury is being asked to rate the student responses on a five-point scale from high-creative to low-creative. In addition, raters are being asked subsequently to nominate the criteria which they used in their ratings. Both the designs and the rater criteria will then be examined to explore the influences of symmetry, complexity and other possible determining factors, in assessing design responses as "creative".

Some of the items may be scorable for all three factors (i.e. fluency-flexibility-originality), some for two of these, and others for only one factor.

The Resolution of Some Psychometric-Creativity Dilemmas

(a) *Administration conditions*

For the creativity test to have classroom relevance, it had to be administrable within a single class period (generally 30 - 40 minutes), yet its items must adequately sample the creativity test grid (i.e., a minimum of twelve cells each to be sampled at least once). The minimum number of items to achieve this appeared to be sixteen. Even applying time limits of from two to four minutes, it became necessary to split the test into two parts (labelled Part 1 and Part 2). This was arbitrarily done with a view to providing a balance in difficulty levels and item types. Allowing for time for distribution and collection of tests, the resultant parts were found to take approximately 35 - 40 minutes.

(b) *Effects of time limits*

The incorporation of time limits on each item introduced a new variable. Wallach and Kogan (1965) criticised the use of time limits in creativity tests. They claimed it was not conducive to creative performance, and that time limits made conditions too similar to those of IQ tests. In their own study, they abandoned time limits and instituted what they called a "game-like atmosphere". For young children, the tests were individually administered. As a result they found very little correlation between IQ and their creativity tests, and therefore concluded that their tests were more valid than tests with time limits.

In answer to criticisms about time limits, Guilford (1971:79-80) reported results of investigations using extended time limits. He found that item scores from the first two minutes gave a far better measure than scores on the last three minutes, and that responses during the latter period were very sparse. He does not report how originality was affected. It seems possible that with longer time, more original responses would be produced, as many studies have shown that they tend to occur later in the test period. A recent study by Cropley (1972:31) indicated that "highly uncommon responses were found in both the early and late responses", but nevertheless that "time scores (of originality) correlated substantially with untimed scores" (Cropley, 1972:34).

In any case, apart from the impracticality of school use for a test stripped of time limits, many uncontrolled variables are likely to enter. These could include boredom, opportunity to include suggestions from others, facilitating strategies and lack of perseverance. Those who are capable of producing the most responses in a short time might be expected to be those who are capable of producing the most responses in a longer time.

(c) *Test Instructions*

Vernon's (1971:245) recent findings confirm that a major difficulty in devising instructions and format for the test lies in trying to dissociate the tasks from the usual disciplined atmosphere of intelligence and achievement tests, conveying rather an atmosphere of freedom and challenge, still within a classroom setting. As Guilford (1971:77) points out, creativity tests cannot be completely freed from a competitive situation and most students would realise they were being assessed in some way. The "correctness" criterion must therefore be de-emphasised. This is stressed in the instructions which have been phrased in as free a manner as possible. The items are not referred to as "tests" but as "tasks", and subjects are told they are being challenged to give as many creative original responses as possible.

Numerous studies overseas have shown how much the actual conditions of testing and wording of instructions can influence results (cf Vernon, 1971). Suggestions for best results include a friendly and encouraging tester, a relaxed atmosphere, motivation of the subjects by interest and perhaps self-competition and comfortable physical surroundings. Once these suggestions have been made, the rest has to be left to the teacher administering the test. The actual instructions for administration, however, must be followed exactly, as slight changes can alter cues. It has been necessary therefore to pay considerable attention to their wording. A practice example is given before the test begins to give students an idea of the different approach required by creativity tasks, but time and feasibility constraints preclude the inclusion of practice examples with each new item. Also, trials have shown that where detailed examples are given, subjects tend to copy these, and so originality is inhibited. Other difficulties have also been revealed in trials, necessitating alterations in the instructions. It was found, for example, that the suggestion to use rulers or compasses for drawing lines or circles was severely retarding the production of designs. Younger students did not understand what was meant by the phrase "word association", and wording had to be found to suggest the concept clearly. In earlier forms of the instrument, some students were not familiar with some of the objects named in "devices" tasks, such as book-ends and cotton reels: substitutes had to be made. More important, however, were the difficulties in striking the right balance in cueing for frequency, flexibility and originality.

It might be argued that fluency and originality of responses should not be scored in the same item, and that the latter will suffer. However, reserach by Wallach (1970:1211) has shown that there tends to be a positive correlation between the two, the more original responses tending to occur at the end of a large number of responses. There are exceptions. Some subjects produce many responses with little originality; others, in spite of instructions to the contrary, tend to evaluate their responses first and do not produce as many as possible.

(d) *Scorer Reliability*

In seeking to maximize the objectivity (and reliability) of scoring, the following criteria were adopted:

fluency is assessed by the *number of responses* a subject makes to an item,

flexibility by the *number of shifts* in thinking or breaks from mental set, and

originality by the statistical *infrequency* of a response.

By contrast to tests of convergent thinking which are frequently machine-scored, scoring of creativity tests could be a long and tedious task. The *fluency score* presents little problem, since it represents a *tally* of all the responses made except for those that are inappropriate. (For example, if an item calls for a word beginning with "tr-", a word beginning with "ta-" would not be acceptable.)

Scoring of *flexibility* is more difficult. As defined, it requires a shift in thinking from one class or category to another and boundaries between classes are not always distinct; for example, responses to suggesting problems raised by a glass of water could include "testing the water for purity" and "testing the water to analyse chemical components". One scorer could argue that there is no shift, that both involve testing the water. Another scorer could argue that one test would be microbiological, while the other would involve chemical analysis, and a shift would be involved. As many examples and guidelines as practicable are to be given in the scoring guide. It should be noted that in this approach to scoring, there is no attempt to judge the *quality* of responses.

The scoring of *originality* depends upon unusualness, or statistical infrequency, rather than upon intrinsic merit. Guilford used the characteristics of remoteness and cleverness to score originality (Wilson et al, 1954), but it has been found that these practices introduced too much subjectivity into scoring, and research has shown that the validity of items has not been lowered by their exclusion. Cleverness in such items as suggesting titles for a story seems to rely on a fairly limited kind of creative word play.

In order to use the "statistical infrequency" approach as the basis of unusualness in scoring originality, it becomes first necessary to tabulate all the responses to the relevant item. The resultant method of assigning scores was first suggested by Cropley (1967:109). Those with a frequency of 1% or less receive a score of 4, those with a frequency of from 2% - 3% receive a score of 3, from 4% - 7% score 2 and from 8% - 15% score 1. All other responses score zero. A random sample of responses from one hundred students throughout Australia will be used to establish scoring categories for each originality item.

(e) *Reporting Student Performance*

It will be obvious that scores obtained on items for fluency, flexibility and originality will be widely disparate, and that it would be quite invalid to treat such scores as additive. Accordingly, a "creativity profile chart" is being developed, which keeps separate the scores obtained upon each of the twelve cells of the creativity grid. The profile will comprise a set of barometric scales, with a conversion scale for reading off stanine scores. The decision of whether to add raw or standardized item scores within the profile will be determined when data is returned from the national trials in progress. This decision will be based upon the extent of comparability of item variance.

Conclusion

The instruments described above are currently undergoing trials in a national, stratified (proportional-by-states-and-type-of-school), random sample involving more than 1000 students in eighty schools at the levels equivalent to Forms 1 to 4 in N.S.W. schools. Student performances in these trials may be used to help establish typical ranges or stages of creative performance.

> "Whatever the limitations of these (creativity) tests might be, they can help educators become aware of potentialities that might otherwise go unnoticed." (Torrance, 1967).

93

Acknowledgements

Acknowledgement is made for the help of the following:

Mrs. Fay Pettit (both as a Research Assistant, and for help in compiling this Report);

Mrs. Jan Mitchell (Research Assistant);

Mrs. Beverley Sheldrick (Typist);

Mr. R.E. Wilkes (Senior Research Officer, Test Services Division, ACER, for general advice and for assistance in the drawing of the sample for national trials).

REFERENCES:

A.S.E.P. (1970). "Positions Document 38: Use of Inquiry Approach". Melbourne: ASEP.

Barron, Frank (1969). *Creative Person & Creative Process*, New York: Holt Rinehart & Winston.

Barron, Frank and Welsh, G.S. (1952). Artistic Perception as a Factor in Personality Style: Its Measurement by a Figure Preference Test, *J. of Psychology, 33*, 199-203.

Cohen, David (1957). Lower Level Secondary Examinations, *The Educational Magazine, 14*, 186-192.

Cohen, David (1963). Multiphasic Assessment in Science, *VIER Bulletin, 9*, 1-12.

Cohen, David (1966). Organising a Science Performance Test, *South Australian Science Teachers' Journal*, 35-36.

Cohen, David (1971). "Curriculum Objective: Scientific Attitudes". Report of Research Project funded by Macquarie University (Mimeographed). Pp. 41.

Cohen, David (1968). Opportunities for Creativity, *Australian Science Teachers' Journal, 14*, 6-19.

Cropley, A.J. (1972). Originality Scores Under Timed and Untimed Conditions, *Australian Journal of Psychology, 24*, 1, 31-36.

Cropley, A.J. (1967). *Creativity*, London: Longmans Green & Co.

French, J.W., Ekstrom, R.B. and Price, L.A. (1963). "Kit of Reference Tests for Cognitive Factors", Princeton: Educational Testing Service.

Guilford, J.P. (1950). Creativity, *American Psychologist, 5*, 444-454.

Guilford, J.P. (1959). Three Faces of Intellect, *American Psychologist, 14*, 469-479.

Guilford, J.P. (1971). Some Misconceptions Regarding Measurement of Creative Talent, *Journal of Creative Behaviour, 5*, 2, 77-87:

Guilford, J.P. and Hoepfner, R. (1971). *The Analysis of Intelligence*, New York: McGraw-Hill.

Hadamard, J. (1945). *The Psychology of Invention in the Mathematical Field*, New York: Dover.

Hudson, Liam (1966). *Contrary Imaginations*, London: Methuen & Co.

Mackinnon, D.W. (1967). "Identifying and Developing Creativity" in Gowan, J.G. *et al.* (eds), *Creativity: Its Educational Implications*, New York: Harper & Row.

Mehrens, W.A. and Lehmann, I.J. (1969). *Standardized Tests in Education*, New York: Holt Rinehart & Winston.

Nyquist, E.B. (1971). Open Education: Its Philosophy, Historical Perspectives, and Implications, *The Science Teacher, 38*, 25-28.

Parnes, S.J. and Meadow, A. (1959). Effect of Brainstorming Instructions on Creative Problem-solving by Trained and Untrained Subjects, *J. Educational Psychology, 50*, 171-176.

Torrance, E.P. (1966). *Torrance Tests of Creative Thinking: Directions, Manual and Scoring Guide*, Princeton: Personnel Press.

Torrance, E.P. (1967). Uniqueness and Creativeness: The School's Role, *Educational Leadership*, 24(6), 494.

Vernon, P.E. (1971). Effects of Administration and Scoring on Divergent Thinking Tests, *B. Journal of Educational Psychology, 41*, 245-257.

Wallach, Michael A. (1970). Creativity, in Mussen, Paul H. (ed.), *Carmichael's Manual of Child Psychology*, New York: John Wiley & Sons, Inc.

Wallach, M.A. and Kogan, N. (1965). *Modes of Thinking in Young Children*, New York: Holt Rinehart & Winston.

Wilson, R.C., Guilford, J.P., Christensen, P.R. and Lewis, D.J. (1954). A Factor-Analytic Study of Creative Thinking Abilities, *Psychometrika, 19*, 297-311.

EDITOR'S COMMENTS:

1. During the presentation Dr. Cohen distributed samples of the test items and two of these are reproduced below.

   (i) (2 minutes)

   "In the time allowed, write as many pairs of words as you can starting with the given letters, and in the same order. The two words must make sense together. Try not to use the same word more than once.

   EXAMPLE: for the letters t .... g .... :-

   talented girl, tough guy, try glue, ...

   YOUR LETTERS ARE: p .... r ....

   (ii) (3 minutes)

   See how interesting you can make drawings or designs created from the repeated use of this shape.

2. A considerable amount of animated discussion followed the paper. The comments ranged over a number of issues and included statements on

   (a) whether the test described was clearly a norm referenced one rather than a criterion referenced one,

   (b) whether an adult, who is regarded as creative, can specify the characteristics of creativity, and

   (c) whether the term creativity test is a misnomer which should be replaced by "measure of fluency and originality".

3. A recent review by Crockenberg in volume 42, number 1 of *Review of Educational Research* should be read as a sequel to Dr. Cohen's paper.

GRAPHICAL SKILLS AND LEARNING HIERARCHIES IN SCIENCE

OUTLINE OF A LEARNING HIERARCHY INVESTIGATION

R. T. White

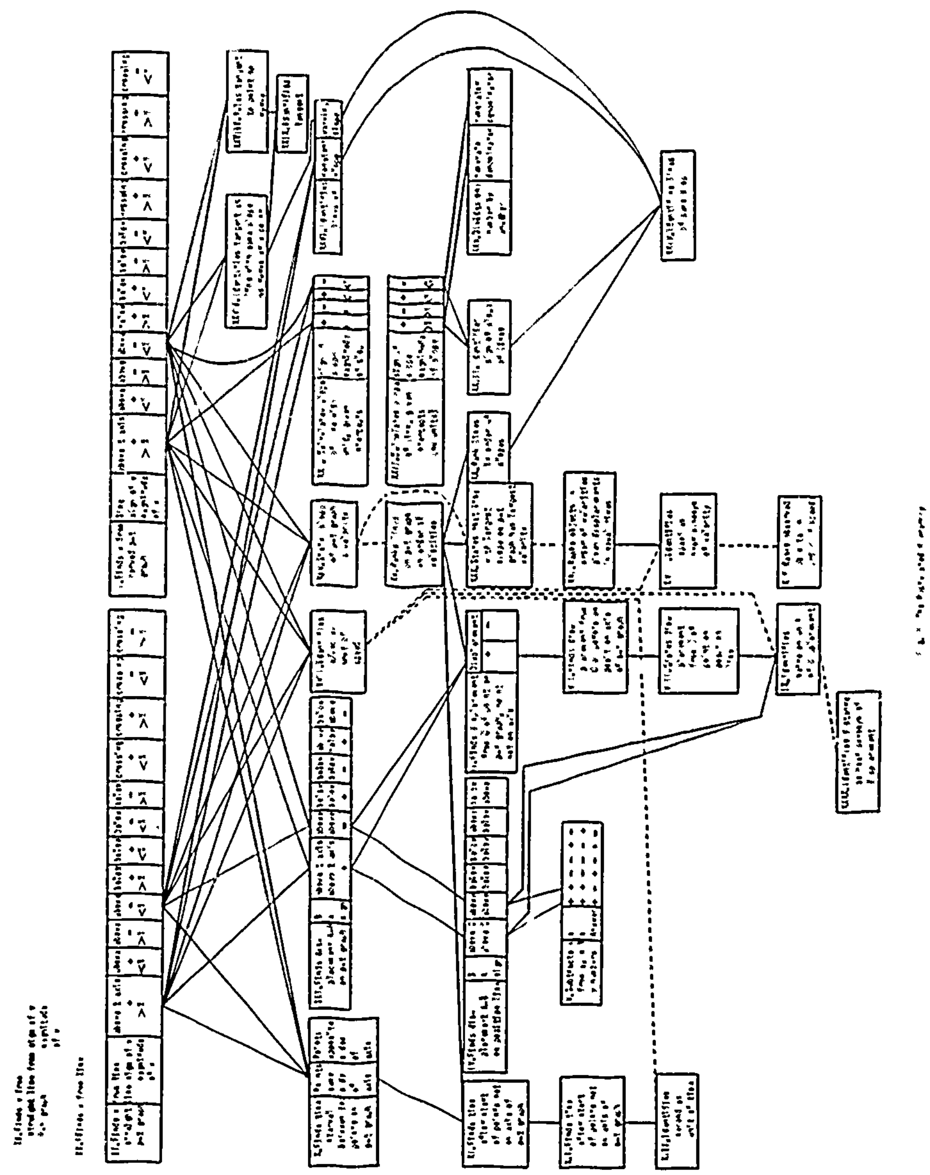
Cronbach (1966, p.91) once described the mass of people as rushing to and fro from one fashionable bandwagon to another, while the researcher stood on a small piece of firmly-tamped ground crying "wait for me!" as the herd galloped past. This paper describes an attempt to increase the area of the firm ground by a small amount. It does not contain the solution to all the teacher's questions, and deals with learning of subject matter.

The purpose of the research was to explore the accuracy of Gagné's theory of learning hierarchies (Gagné 1962), and to see whether it is limited to a particular class of subject matter. The basis of his theory is that any piece of knowledge can only be acquired by people who possess certain pre-requisite pieces of knowledge, which have their own pre-requisites in turn. Subsequently, after the original statement of the theory, Gagné suggested that it might apply only to a type of knowledge which he called intellectual skills, and not to another type which he called verbalized knowledge (Gagné 1968). An intellectual skill is the ability to perform a whole class of tasks, such as solving simple linear equations, and a verbalized knowledge element is a single statement, such as knowing that metre is a unit of length.

Gagné's suggestion that only intellectual skills are learned hierarchically needed to be tested empirically. Previous studies of hierarchies had failed to show whether any knowledge, of either type, is learned hierarchically, because of weaknesses of experimental design. Gagné and Paradise (1961) performed the first major investigation of a hierarchy and their methodology was used in most subsequent investigations. Some improvements to this basic design were made in this study. A hierarchy was developed for the subject-matter, "finding velocities of objects from position-time graphs", in the manner described by Gagné and Paradise, but then an attempt was made to define its constituent elements more narrowly. This led to several of the original elements being divided. For example, the element *Finds the displacement between two points on a position-time graph* was divided into cases where the displacement was positive and negative, and into further cases depending on the positions of the points above or below the time axis. The reality of these divisions was tested empirically. The final hierarchy is shown in figure 1, with some simplification – connections are drawn in only for a representative set of divisions of elements.

A learning program for the elements was written; and questions testing each element were embedded in the program, instead of being placed at the end of the program as was the case in all previous studies. Wherever possible, two or more questions were used for each element, to allow an estimate of the reliability of the questions. Only one question was possible for each verbalized knowledge element.

After a trial with a few subjects, the program was completed by a large number of subjects. For each postulated connection between a lower element and a higher one, the number of subjects who answered incorrectly all questions for the lower element but correctly all questions for the higher one was inspected. Wherever possible this number was compared



with the critical number for such subjects: the critical number is found from White and Clark's test of inclusion (in press). Where the observed number exceeded the critical number, the connection was judged invalid. The critical number could not be calculated when either of the elements was tested by one question only, or when all subjects answered correctly all the questions for either element. In those cases rejection of the connection was necessarily a matter of subjective judgement. In some cases it was difficult to tell whether the connection should be accepted or not, so no decision was made.

Nearly all of the connections involving a pair of intellectual skills were accepted as valid, while most of those involving one or two verbalized knowledge elements were rejected. The experiment showed that Gagné's original theory accurately represents one aspect of how intellectual skills are learned, but does not represent how verbalized knowledge is learned as Gagné himself suggested (Gagné 1968).

Much more research is needed to answer further questions, such as the causes of a subject's failure to learn even when he possesses all the apparently necessary subordinate elements. The present results have, however, some immediate implications for classroom practice. They suggest that whenever intellectual skills have to be learned, as is frequently the case in physical science and mathematics, teachers should take greater care than usual over planning sequences of instruction. Formation of a hierarchy would be a useful first step in this planning, and would be helpful in designing effective readiness, diagnostic, and mastery tests.

REFERENCES:

Cronbach, L.J. (1966). The logic of experiments on discovery. In *Learning by Discovery: a Critical Appraisal,* eds. L.S. Shulman and E.R. Keislar, Ch.5. Chicago: Rand-McNally.

Gagné, R.M. (1962). The acquisition of knowledge. *Psychological Review,* 69: 355-365.

Gagné, R.M. (1968). Learning hierarchies. *Educational Psychologist,* 6: 1-9.

Gagné, R.M., and Paradise, N.E. (1961). Abilities and learning sets in knowledge acquisition. *Psychological Monographs,* 75: 14 (whole number 518).

White, R.T., and Clark, R.M. A test of inclusion which allows for errors of measurement. *Psychometrika* (in press).

BASIC SKILLS OF GRAPHICAL INTERPRETATION

R. D. Linke

The use of graphical methods for the presentation of both scientific and statistical information has become increasingly important in recent years, since we have been regularly and progressively bombarded with a myriad of "facts and figures", which must be carefully organised and simplified to enable any real understanding or analysis. Moreover, the educational importance of graphical techniques in scientific communication is now recognised throughout the world, and since these skills are not generally acquired through incidental learning, they have now been specifically incorporated into every state primary curriculum in Australia, in addition to many similar courses overseas.

Courses of instruction in graphical techniques may be conveniently classified into two categories (a) General courses, usually incorporated in the mathematics curriculum, and (b) Service courses, usually associated with a physical, biological, or general science curriculum. Examples of General courses are included in the current state primary mathematics curricula for Victoria (5), Queensland (3) and South Australia (4), the S.M.S.G. Elementary Mathematics Program (12) in U.S.A., and in the Nuffield Mathematics Project (10, 11) in Britain. Service courses are produced by I.S.C.S. (9) and B.S.C.S. (2) in U.S.A., and by A.S.E.P. (1) in Australia to complement their respective scientific programmes. A comparative analysis of all these courses is outlined in Tables I - III.

Significant difference among the various General courses (Tables I and II), and between these and the Service courses (Table III) may be noted in several areas of comparison. Whereas the General courses are usually introduced in the early or middle primary grades and continued intermittently over several years, the Service courses are introduced at the High School level and generally completed as a single unit of study. Moreover, the General courses show two basically different *types of approach* – the Developmental approach, used in the Victorian curriculum (5) and one of the Nuffield courses (10), emphasises the gradual transition from operations with concrete objects to the abstract representation of statistical data, while the Hierarchical approach, as shown in the other General courses and in all the Service courses, concentrates on a logical sequence of progressively complex and more carefully defined graphical skills. It should be noted, however, that this hierarchical sequence is intuitively, rather than empirically based, and is by no means consistent from course to course.

The *construction* of graphs and the *interpretation* of graphs are generally thought to involve two different, though probably overlapping sets of skills, and thus a comprehensive treatment of graphical techniques must involve specific reference to both areas of learning. Despite the claim by Smith (13) that "the treatment of graphs in the primary mathematics program often lays the main stress on construction work", this analysis indicates that most General courses emphasise interpretation, while by contrast the Service courses are strongly oriented toward construction skills.

TABLE I

GENERAL COURSES (A)

(Australian State Primary Mathematics Curricula)

CHARACTERISTICS	VICTORIA	QUEENSLAND	SOUTH AUSTRALIA
Stage of Introduction	early primary grades	early primary grades	grade III
General Approach	Developmental	Hierarchical	Hierarchical
Skill Class Emphasis	interpretation via construction	interpretation	interpretation via construction
Initial Numerical Range	limited range of positive integers	positive integers (0-10)	positive rational numbers
Initial Specific Vocabulary	limited (5 terms)	limited (4 terms)	Moderate (9 terms)
Informational Models	extensive list of models suggested	several models used – no particular emphasis	several models used – no particular emphasis
Graphical Forms	Bar Graph Line Graph Circle Graph Pictograph	Bar Graph Line Graph Circle Graph Pictograph	Bar Graph Line Graph Circle Graph Pictograph

TABLE II

GENERAL COURSES (B)

(Overseas Primary Mathematics Curricula)

CHARACTERISTICS	S.M.S.G.	NUFFIELD (1) Pictorial Representation)	NUFFIELD (2) (Graphs to Algebra)
Stage of Introduction	Grade VI	early primary grades	middle primary grades
General Approach	Hierarchical	Developmental	Hierarchical
Skill Class Emphasis	interpretation	interpretation via construction	interpretation
Initial Numerical Range	positive and negative integers	limited range of positive integers	limited range of positive integers
Initial Specific Vocabulary	Exteasive (16 terms)	Moderate (9 terms)	Moderate (10 terms)
Informational Models	Population model and others	Several models used – no particular emphasis	Symbolic model and others
Graphical Forms	Bar Graph Line Graph Circle Graph Pictograph	Bar Graph Line Graph Circle Graph Pictograph	Line Graph only

The *numerical range* used in graphical exercises throughout the General courses often lags well behind the introduction of more extensive and complex number systems in other sections of the primary curriculum, and even the High School Service courses are restricted to relatively simple number systems. Negative numbers are rarely used in introductory graphical exercises.

The *introductory vocabulary* of specific graphical terms (such as "horizontal", "vertical", "co-ordinate" and "axis") is generally fairly limited in the primary courses, particularly at the lower grades, but seems to be more extensive in the High School Service courses, which often cover more sophisticated skills. In contrast with this, the *range of information models* either given as examples or suggested for later exercises is usually much more extensive in the General courses than in the Service courses, as is the range of graphical forms, and in fact the Service courses are concerned only with the two-dimensional line graph, according to Weintraub (14), is "one of the most difficult of all graphic forms to interpret". It is also, however, by far the most common and versatile form.

In spite of various broad similarities emphasised in this analysis, there are as many different intuitive approaches to the teaching of graphical skills as the total number of courses under study. It is therefore obvious that the intuitive approach is inconsistent, and most likely that it is also inefficient. The *purpose of this project* is to validate empirically a hierarchical network of basic graphical skills, and to determine the influence of various factors on the structure of this hierarchy.

There is undoubtedly a large range of important skills associated with the construction and interpretation of graphs, particularly if all the common graphical forms are considered, but the complex procedures involved in hierarchical validation, and the necessity to control where possible the known and likely variables, have imposed certain restrictions on the *scope of this investigation.* With this in mind, it was decided to restrict the range of postulated skills to those directly associated with graphical interpretation, rather than with construction, and to limit the format to line graphs based on a two-dimensional grid.

The *initial outline of the hierarchy* was determined by defining a number of terminal skills, then specifying the set of subordinate skills thought to be necessary for their achievement. This process was progressively repeated to produce an overlapping set of hierarchically arranged skills, each logically dependent on those below. This method, which follows the procedure outlined by Gagné and Paradise (6), is consistent with general practice in the field, though the intuitive gap or leap between successive skills is, initially at least, a matter for arbitrary definition. An alternative, though less reliable procedure adopted by Smith (13) involves the collection of a pool of miscellaneous items, which are arbitrarily grouped according to conceptual similarities, then hierarchically ordered by analysis of difficulty levels. The *final postulated hierarchy* for this project is outlined in Appendix I, slightly modified from its initial form on the basis of preliminary testing results. It is divided for convenience into six sections, dealing respectively with positional aspects, displacement, gradient and area.

TABLE III

SERVICE COURSES

CHARACTERISTICS	B.S.C.S.	I.S.C.S.	A.S.E.P. (National Trial Unit)
Stage of Introduction	middle High School	Junior High School (11-12 years)	early High School
General Approach	Hierarchical	Hierarchical	Hierarchical
Skill Class Emphasis	interpretation leading to construction	construction leading to interpretation	construction leading to interpretation
Numerical Range	positive integers	positive rational numbers (including fractions and decimals)	positive rational numbers (including fractions and decimals)
Specific Vocabulary	Moderate (8 terms)	Extensive (13 terms)	Extensive (14 terms)
Informational Models	Physiological and Population model	limited range of models	limited range of models
Graphical Forms	Line Graph only	Line Graph only	Line Graph only

In order to define more precisely the range of possible examples associated with each of the skills, more limited *subdivisions* were defined, then tested with two different questions from each subdivision over a sample of at least one hundred subjects at suitable High School grades. Subsequent comparative inspection analysis was used to differentiate between pairs of subdivisions where several subjects were able to answer one set of questions but not the other. This subdivision analysis was then used to define more precisely the limits, such as numerical range, of each skill, so that cross-questioning between different subdivisions of sequential skills, which might invalidate any hierarchical connection, could be easily avoided. An example of this subdivision analysis is shown in Appendix III.

The *learning and testing programme* designed to validate the postulated hierarchy was simplified on the basis of the analysis above by reducing the number of experimental subdivisions in each of the terminal skills. Thus the numerical range was limited to the positive integral numbers from one to ten, and all construction skills excluded unless necessary for subsequent interpretative questions. On the other hand, since both the Horizontal and Vertical co-ordinates were necessary for later displacement and gradient calculations, these were incorporated as separate divisions in each of the appropriate subordinate skills.

The *general presentation procedure* was to teach, then immediately test each skill in turn (as shown in Appendix II) along linear sections of the hierarchy, and to retest previous skills where non-linear connections were involved. Two questions were used to test each skill, in order to account for any errors or successes by chance. An alternative procedure adopted by White (15) involves an immediate additional retest of the first skill in each pair of skills presented, to eliminate those subjects who learn the first skill only in the process of doing the second. This step was considered unnecessary here, since it was argued that learning a subordinate skill in the process of doing another logically precludes any absolute hierarchical dependence of the second skill on the first, and in fact the number of subjects excluded by White on the basis of this additional procedure was generally insignificant.

A *trial learning programme* containing all 24 postulated skills and involving 73 associated questions, was prepared in accordance with the presentation procedure outline above. This was then divided into three sections of approximately equal length, dealing respectively with position, displacement/gradient, and area, thus corresponding with the general conceptual areas defined in the hierarchy. The trial programme was then administered to one class of at least 30 students at each of three academic levels (grade 6, form 1 and form 2) in Melbourne metropolitan schools. Single programme sections were administered in alternate periods, in order to avoid as far as possible any general lapse in interest or concentration.

The *purpose of this trial* was to eliminate any major errors from the hierarchy, to select the most appropriate academic level for the major validation experiment, and to determine the instructional requirements for each skill, and the total time needed to complete each section at the chosen level. The instructional section of the programme was written in the simplest possible terms, and although symbols were often used for convenience, both these and any specific graphical terms were always explained when introduced. Apart from any intuitive grounds, this approach was necessitated by the relatively poor understanding of certain graphical terms revealed in recent studies by Gardner (7, 8) with High School students in both Victoria and T.P.N.G., the latter being relevant to a subsequent aspect of this study. In contrast with the earlier approach by White (15), specific terminology was not included here as part of the hierarchical structure, being defined as a single element of knowledge, rather than representative of a general skill.

As a result of the trial above, several minor *modifications* were found necessary in the major validation programme. Additional examples were needed for some of the more complex skills, and clarification occasionally required for questions or instructions. No alterations were necessary in general format, however, since each of the programme sections was completed by every student in the course of a single period (35 - 40 minutes), and there were no obvious or general vocabulary difficulties. The *statistical analysis* for the major validation programme will follow the procedure outlined by White (15), and the *testing population* of 200 students chosen at form 1 level from a random selection of Melbourne metropolitan High Schools, using one half class of students from each selected school.

The final, and perhaps the most important feature of this project involves the *effects of certain variables* on the hierarchical structure of the postulated network of graphical skills. These variables include numerical range, which largely determines the computational complexity of the skills, the type of informational model, which almost certainly contributes to their conceptual difficulty, and the background of the subjects with respect to both specific mathematical experience and general cultural environment.

In order to determine the effect of *numerical range,* a second programme was prepared, parallel to the first in every respect but involving positive decimal fractions, rather than integral numbers. This will be tested with an analogous population of students, to determine whether the same hierarchical structure is validated.

The effect of *different informational models* will be determined by the application, also to a similar population of form 1 students, of a third programme, differing only from the first with respect to the type of informational or conceptual model. In contrast with the symbolic (X/Y) model presented in the first programme, this one involves a more realistic population model, with Annual Birth Rate plotted on the Vertical axis of the graph, and Time on the Horizontal axis. This model obviously involves a completely different set of specific terminology, since the explanation of every skill is related to the concepts of Time and Annual Birth Rate, rather than to non-specific symbols. It was necessary in choosing the model for this programme to give particular consideration to the meaning of area, since in many other common models this is not a meaningful quantity, and would therefore be inappropriate for comparison.

It is fundamental to the theoretical basis of logical hierarchies that these structures should be largely independent of *background experience* or environment. The possible effects of such a variable will be determined in this project by testing the first programme with an analagous population of students in other states of Australia, where there are readily definable differences in background mathematical experience (at primary level), and in the Territory of Papua and New Guinea, where in addition to this there is a completely different cultural environment. It may be necessary in the latter case, because of certain vocabulary difficulties encountered by others in former research (8), to test the programme at a higher academic level, though fair consideration was given to these prospective problems in the construction of the major validation programme.

With respect to *future research* on graphical skills not covered in this project, the importance of constructional techniques, and their obvious association with interpretative skills, would undoubtedly make this a promising area for extending the present studies. Further extensions to cover both constructional and interpretative aspects related to other graphical forms would also provide a valuable basis for prospective research.

REFERENCES:

1. Australian Science Education Project. "Graphing, Averaging, Reading a Scale". National Trial Unit. *ASEP* (1972).

2. Biological Science Curriculum Study. "How to Make and Read a Graph" *in* "Biological Science: Patterns and Processes". (Revised edition, Student Manual). Holt, Rinehart & Winston, New York (1970).

3. Education Department (Queensland). "Program in Mathematics for Primary Schools". Stages 1 - 4 (1966), Stages 5 - 6 (1967), Stages 7 - 8 (1968).

4. Education Department (South Australia). "Course in Mathematics". Grade I (1971), Grade II (1971), Grades II - IV (1970), Grade V (1969), Grade VI (1970), Grades VI - VII (1971).

5. Education Department (Victoria). "Curriculum Guide - Pure Number". Section A (1967), B (1965), C (1967), D (1969), E (1965), F (1965). "Applied Number" - Sections A - C (1969), D - F (1968). "Mathematics" Section G (1968), H & I (1969).

6. Gagné, R.M. & Paradise, N.E. "Abilities and Learning Sets in Knowledge Acquisition". *Psychological Monographs* 76:7 (whole number 526) 1962.

7. Gardner, P.L. "Words in Science". Report prepared for the Australian Science Education Project. Melbourne 1972. (in print).

8. Gardner, P.L. "Project S.W.N.G. - Scientific Words New Guinea". Report to the Department of Education, T.P.N.G. (1971).

9. Intermediate Science Curriculum Study. "Probing the Natural World". Vol. 1. Silver Burdett Co., New Jersey, 1970.

10. Nuffield Mathematics Project 1. - "Pictorial Representation" Nuffield/Chambers & Murray (U.K.) 1969.

11. Nuffield Mathematics Project 2. - "Graphs Leading to Algebra" Nuffield/Chambers & Murray (U.K.) 1969.

12. School Mathematics Study Group. "Mathematics for the Elementary School". Grade 6 (part 2) and Grade 6 (part 3) Teachers' Commentary (Revised edition) S.M.S.G. (U.S.A.) 1962.

13. Smith, J. "Graph Skills Program". Australian Council for Educational Research - Confidential Report, 1970.

14. Weintraub, S. "Reading Graphs, Charts and Diagrams". *Reading Teacher. 20* : 345 - 349 (1967).

15. White, R.T. "Learning Graphical Skills in Kinematics". Ph.D. dissertation. Monash University (1971).

APPENDIX I

BASIC SKILLS OF GRAPHICAL INTERPRETATION
FINAL POSTUALTED HIERARCHY

Section 1 - Position
2 - Position
3 - Position
4 - Displacement
5 - Gradient
6 - Area

Numerical Code

2/1(A) = Section 2/Level 1 (terminal)/Track (A)

POSITION (1)

1/1(A) Calculate the Horizontal or Vertical position of a point, specified by one co-ordinate, on a two-dimensional line graph.

1/1(B) Mark the position of a point, specified by one co-ordinate (Horizontal or Vertical) on a two-dimensional line graph.

1/2 Calculate the Horizontal or Vertical position of a given point on a two-dimensional grid.

1/3 Calculate the position of a given point on a Horizontal or Vertical number line.

POSITION (2)

2/1(A) Calculate the Horizontal or Vertical position of a point, specified by one co-ordinate, interpolated between a given row of points on a two-dimensional grid.

2/1(B) Calculate the Horizontal or Vertical position of a point, specified by one co-ordinate, extrapolated beyond a given line segment or row of points on a two-dimensional grid.

1/1(A)

POSITION (3)

3/1 Identify from a mixed sample of curves those with a Maximum or Minimum Turning Point.

3/2(A) Calculate the Maximum or Minimum value of a curve on a two-dimensional grid.

3/2(B) Identify and mark the Turning Point on a given curve.

1/1(A)

DISPLACEMENT

4/1 Calculate the Horizontal or Vertical displacement between two given points on a two-dimensional grid.

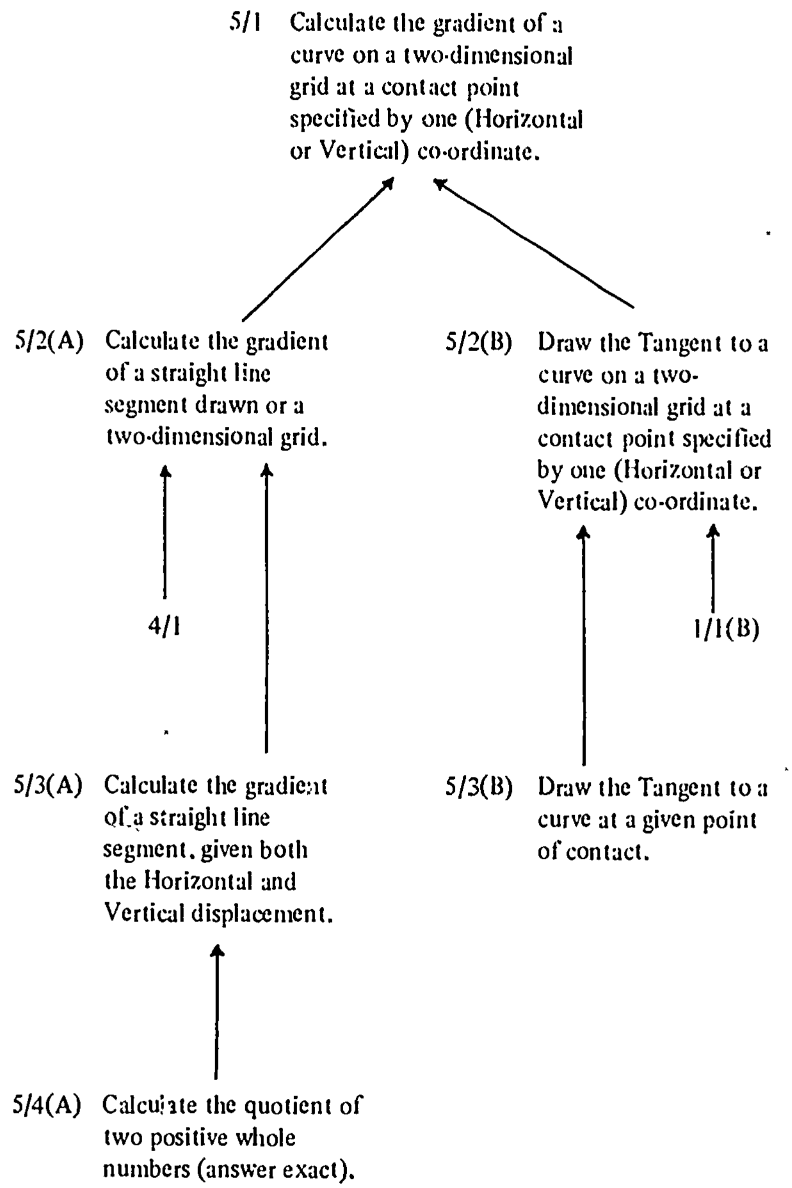
1/2

4/2 Calculate the displacement between two given points on a Horizontal or Vertical number line.

1/3

4/3 Calculate the difference between two positive whole numbers (answer positive).

GRADIENT


5/1 Calculate the gradient of a curve on a two-dimensional grid at a contact point specified by one (Horizontal or Vertical) co-ordinate.
5/2(A) Calculate the gradient of a straight line segment drawn or a two-dimensional grid.
5/2(B) Draw the Tangent to a curve on a two-dimensional grid at a contact point specified by one (Horizontal or Vertical) co-ordinate.
4/1
1/1(B)
5/3(A) Calculate the gradient of a straight line segment, given both the Horizontal and Vertical displacement.
5/3(B) Draw the Tangent to a curve at a given point of contact.
5/4(A) Calculate the quotient of two positive whole numbers (answer exact).

AREA

6/1 Calculate the approximate area (by the method of counting squares) enclosed between two points of a given line segment, each specified by one co-ordinate, and the Horizontal axis of a two-dimensional grid.

1/1(B)

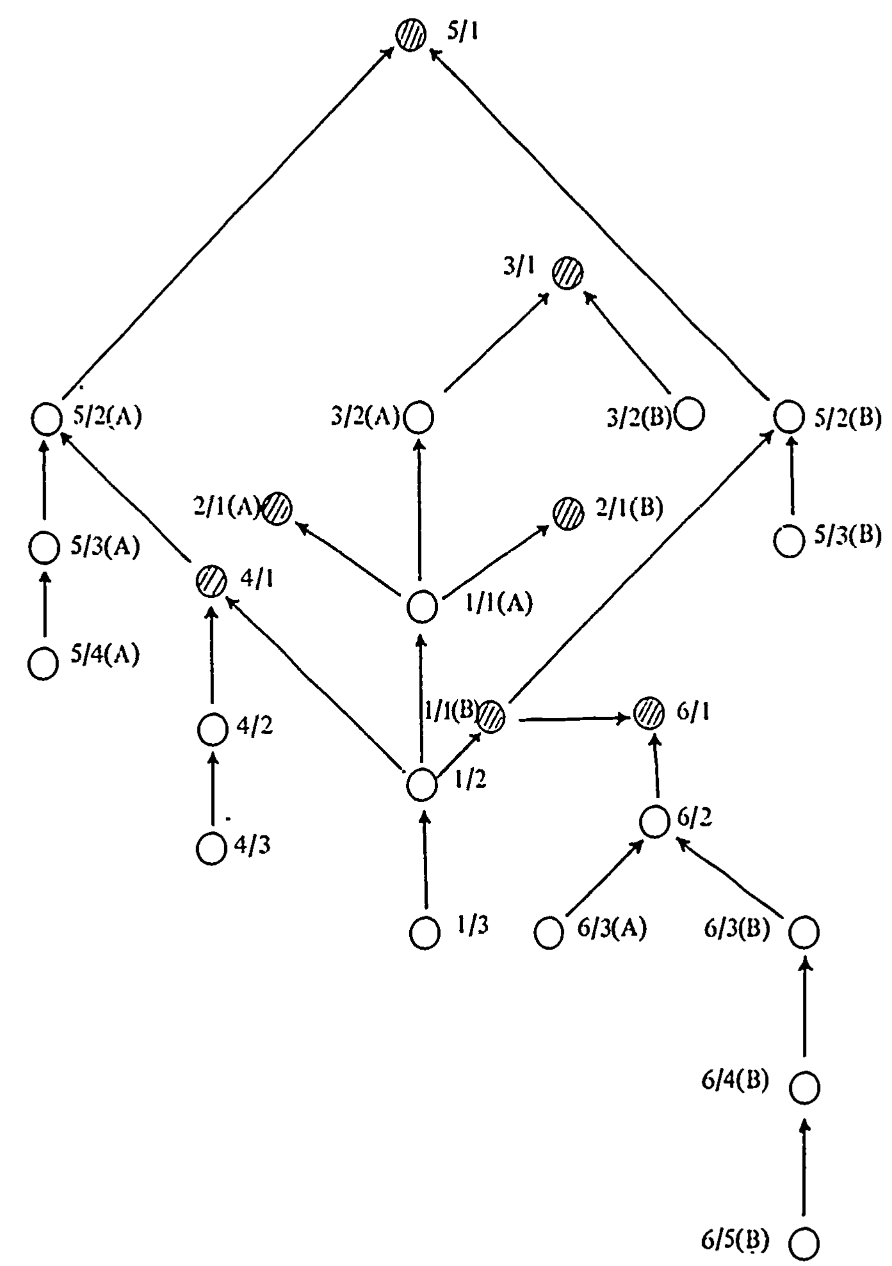
6/2 Calculate the approximate area (by the method of counting squares) enclosed between two marked points on a given line segment and the Horizontal axis of a two-dimensional grid.

6/3(A) Classify the blocks to be counted in order to calculate the area of a specified section on a two-dimensional line graph, where some of the blocks are cut by the given line.

6/3(B) Calculate the area of a single block on a two-dimensional grid.

6/4(B) Calculate the area of a rectangle, given both its length and height.

6/5(B) Calculate the product of two positive whole numbers.



Terminal Skills

Subordinate Skills

APPENDIX II

TESTING PROGRAMME - Sequence of Skill Presentation.

SECTION 1 (Position)

1/3 – 1/2 - 1/1(B) - 1/1(A) - 2/1(A) – 2/1(B) - 1/2(Retest)

1/1(A)(Retest) – 3/2(A) – 3/2(B) - 3/1

SECTION 2 (Displacement and Gradient)

1/3(Retest) 4/3 4/2 - 1/2(Retest) – 4/1 – 5/4(A) – 5/3(A)

4/1(Retest) 5/2(A) - 5/3(B) – 1/1(B)(Retest) - 5/2(B)

5/2(A)(Retest) - 5/1

SECTION 3 (Area)

6/5(B) - 6/4(B) - 6/3(B) – 6/3(A) – 6/2 – 1/1(B)(Retest) – 6/1

TESTING PROCEDURE (for determination of the hierarchical connection between skills 1/3 and 1/2 in Section 1)

Teach 1/3 - Test 1/3 – Teach 1/2 – Test 1/2

APPENDIX III

SUBDIVISION ANALYSIS (Skill No. 1/2)

QUESTIONS

1. Calculate the Horizontal (or Vertical) position of a given point on a two-dimensional grid. (Interpretation)

2. Mark a position, specified by both Horizontal and Vertical co-ordinates, on a two-dimensional grid. (Construction)

Question Group	Subdivisions (Required Co-ordinate)				Other Co-ordinate
	Horizontal or Vertical	Integral No. or Decimal	Number Pos. Neg. or 0.	Construction or Interpretation	
1	H	I	+	I*	V/I/+
2	H	I	+	C	V/I/+
3	H	I	-	I*	V/I/+
4	H	I	+	I*	V/I/+
5	H	D	+	I*	V/I/+
6	V	I	+	I*	H/I/+

COMPARATIVE DIVISIONS (Required Co-ordinate)

H/V = 1/6

I/D = 1/5

+/- = 1/3

+/0 = 1/4

C/I* = 1/2

RESULT Each question group was found to represent a separate subdivision of the general skill.

EVALUATION INSTRUMENTS FOR THE RESEARCHER

The following statement represents a summary of tests submitted in answer to a mailed invitation late in 1970. No attempt has been made to evaluate the quality of the evaluation instruments. Because of the incompleteness of this listing and the variable quality of the instruments listed, it was considered undesirable to produce a more complete inventory of tests at this stage.

Comments and advice of interested people would be most welcome. These comments should be mailed to N.L. Baumgart or David Cohen.

A SURVEY OF EVALUATION INSTRUMENTS

N.L. Baumgart and D. Cohen

Title	Author	Source	Description
W.A. Board of Secondary Education Comparability Tests - Series 1, 2, 3	Officers of Board of Secondary Education and W.A.Educ.Dept.	W.A. Board of Sec.Educ. 3 Ord Street, West Perth	60 item multiple choice measuring achievement of concepts, understandings, and some lower order areas of affective domain. Series 1 is adapted from Form 3A Sequential Test of Educational Progress. Series 2, 3 developed locally. Used in first 3 years of high school in W.A. for moderating science achievement.
How Creative Are You?	David Cohen	Author, Macquarie University	Creativity tests for junior secondary level in two parallel Forms (A and B), each in two parts: designed to evaluate fluency, flexibility, originality in production of words, ideas, devices, designs.
Sydney Teachers' College Standardized Test (X) - Science	K. Dawes, A.A. Hukins, R.W. Maclay	Prof. A. Hukins, University of N.S.W.	Group pencil-and-paper multiple choice achievement tests, designed for use in Form 4, N.S.W. Schools. Part A (50 items) is for O-level, Part B (15 items) is an extension for Credit-level, and Part C (15 items) for A-level.
Science Activities Inventory	L.L. Foster	Author, W.A. Institute of Technology	56-item five-point rating inventory booklet to identify manifest science interests and activities (i.e. verbal reports of voluntary participation in scientific activities).

Title	Author	Source	Description
Biology Readiness Test Materials	P.L. Gardner *et al.*	ACER	Six tests of biology achievement related to prerequisite concepts and principles for "Web of Life" Course.
Scientific Words - New Guinea (now, Words in Science)	P.L. Gardner	Author, Faculty of Education, Monash Univ. or, A.S.E.P.	Approx. 16 different tests, each of 40 items, measuring scientific, non-technical vocabulary. Multiple-choice format.
Student Goals and Objectives in Chemistry	N. Henry	Author, Royal Melbourne Institute of Technology.	A listing intended for use by teachers of chemistry at senior secondary level of 113 chemistry goals (á la Bloom) plus a set of 149 brief descriptions of chemistry lessons which may be rated according to the list of goals.
A Scale to Measure Attitude to Investigation and Discovery of Knowledge	A.A. Hukins	Author, School of Education, Univ. of N.S.W.	A group test of 21 items. Each item is a sentence expressing an attitude; the student selects those with which he agrees. Suitable for use at 10th grade level.
IEA Science Project Tests and Questionnaires	International Assoc. for Evaluation of Education	ACER	Group tests of achievement, understanding science, science opinions, word knowledge, achievement motivation and the like for 14 year olds and for final year secondary students. Also Biology, Chemistry and Physics achievement for final year secondary and teacher questionnaires.

Title	Author	Source	Description
A Test of Interests	G.R. Meyer	Published by Jacaranda Press Pty. Ltd., 46 Douglas St., Milton, Q.	Interest and attitude statements rated 4,3,2,1 or 0. Five sub-scales with 30, 40 60, 20 and 20 items respectively. Battery requires 40 mins.
N.S.W. School Certificate Practical Science Examination	R.W. de M. Maclay	Author, Sydney Teachers' College, Newtown, 2042.	16 items testing laboratory performance: some open-ended, most observational. See also "Educational Measurement and Assessment", The Australian College of Education, A.H. Massina & Co., Carlton : 1969.
Queensland Grade 8 Science Achievement	C.N. Power	Author, University of Queensland	72-item multiple-choice items related to content of Queensland Grade 8 Science Course.
Intermediate Chemistry Test	R.W. Stanhope	ACER (1935)	Forms A and B each comprising 153 completion type items related to subject matter and application of knowledge of laboratory procedures.
A Performance Test - A Problem in Chemical Analysis for Use in Science Classes.	I.D. Thomas	Author, Healesville H.S. P.O. Box 224, Healesville 3777.	Test of practical work in chemical analysis suitable for grades 7 to 11.
Understanding in Science	R. Tisher	Author, University of Queensland. or, A.S.E.P.	Total of 24 multiple-choice items; developed in a study replicating some of Piaget's studies at University of New England. Scores used to classify respondents into stages of mental development.

Title	Author	Source	Description
Questionnaire A	ASEP	ASEP, Glenbervie Rd., Toorak, Vic.	A test of 33 items (5 point rating scale) to obtain teachers' opinions about the values which should underlie the production of a science course for Grades 7 - 10.
Questionnaire B	ASEP	ASEP	A 12 item test with multiple choice responses; used to obtain teachers' opinions about teaching science in the classroom.
Test of Educational Set	N. Baumgart	Author, School of Education. Macquarie Univ.	26 item. forced choice format; gives indication for preference for learning factual material or conceptual material; sub-scales relate to physical sciences and to humanities. Appropriate for 11 -12th grades.
Views of Science Scale	Effie Best	Author, Adelaide Teachers' College.	Designed to measure teacher antecedents and also usable as an outcome measure for methodology.
Science Classroom Activities Checklist (SCACL)	Effie Best	Author	Form 1 – measures teachers' general intentions Form 2 – measures general classroom practices as perceived by students (i.e. transactions, not outcomes).
Biology Exercise Report	Effie Best	Author	More specific than SCACL, this measures transactions as perceived by students.
Teacher-Pupil Relationship Scale	Effie Best	Author	Measures student perceptions of the teacher on a student-to-subject-centred dimension. ("Appears to be a rather powerful predictor of effective outcomes in Science").

Title	Author	Source	Description
. . . .	Effie Best	Author	Open-ended items for students about how they profited from biology, with three-dimension category system to classify responses.
Educational Set (Physics)	L.D. Blazely	Author	A test of 24 items in which respondents order 3 alternatives. The test attempts to measure the extent to which students value facts, applications, mathematical generalizations, verbal generalizations, and questioning in physics.
A Questionnaire for Chemistry Teachers	N.A. Broadhurst	Author, Bedford Park Teachers' College, S.A.	18-item questionnaire to obtain background information and teaching practices for chemistry teachers.

THE CONTRIBUTORS

N. L. BAUMGART is a Lecturer in Education within the School of Education at Macquarie University. He was for some time a teacher and deputy principal in Queensland schools. After completing a science degree and a masters programme in education at the University of Queensland, he joined the staff of Macquarie. He is currently on leave while acting as a resident Educational Consultant in the Philippines.

D. COHEN is Senior Lecturer in Education at Macquarie University. From 1951-62 he taught for the Education Department of Victoria. In 1964 he completed his Ph.D. at Michigan State University in the field of curriculum development. He has lectured in science education at the Technical Teachers' College, and was Research Officer for the Education Department of Victoria (1966-68). His main interests are in curriculum development and evaluation, science education and educational media. He is the author of a number of books and published papers, a member of the Editorial Board of the National Science Curriculum Materials Project, and was editor of the Australian Science Teachers Journal.

P. J. FENSHAM is currently Professor of Science Education at Monash University after an earlier career that included work in physical chemistry and the social sciences. Before coming to Monash he was Reader in Physical Chemistry at Melbourne University. At present he is engaged in innovating new patterns of science teacher education.

P. L. GARDNER a Senior Lecturer in Education, joined the staff of Monash University after some years teaching in Victorian High Schools. His particular interests are in curriculum issues. He has engaged in research in concept attainment and attitude objectives in science education. He has also developed a word test for New Guinea and some readiness tests for *WEB OF LIFE* biology.

R. D. LINKE Research Scholar in Education at Monash University is currently completing a Ph.D. in science education. He came to Victoria from South Australia where he completed a B.Sc. (Hons) in Developmental Biology at Flinders University. He has lectured, part-time, at Toorak Teachers, developed an experimental course in Environmental Science and is currently involved in a study for a prospective National Survey of Environmental Education in Australia – a project sponsored jointly by the Australian Conservation Foundation and Monash University.

K. W. MORITZ has for the past three years been an area specialist in evaluation for A.S.E.P. He is a graduate, B.A., B.Sc., and B.Ed., from Melbourne University and was for some time a science master at Ringwood High, a science co-ordinator at Mitcham High and a science master at Cargilfield School, Edinburgh, Scotland. He has also spent some time in test development with A.C.E.R., and was an evaluation officer for the Junior Secondary Science Project.

C. N. POWER Lecturer in Education at the University of Queensland, has completed a Ph.D. study on the effects of communication patterns and feedback in science lessons. Dr. Power has served as a science teacher in various state high schools in Queensland, as a Research Officer for the Research and Curriculum Branch, State Education Department, Queensland, and as a Lecturer in Education, University of Queensland. At the University, he has continued his interest in micro-teaching, test construction, implementation of curricula, science education and evaluation, and has developed a modified Bellack Scheme for the observation and analysis of classroom behaviours. He is currently preparing various manuscripts for Australian and overseas books.

S. R. SHEPHERD B.Sc., B.Ed., T.P.T.C., is an Area Specialist in Biology at A.S.E.P. He was formerly a teacher in Victorian High Schools and Lecturer in Method of Science at the Faculty of Education, Melbourne University.

R. P. TISHER Reader in Education at the University of Queensland, was, for some time, a science teacher in secondary schools in New South Wales, and a lecturer in Teachers Colleges. He joined the staff of the University of Queensland in 1963. Dr. Tisher's research activities have been associated with science education, the effects of teaching strategies and teacher education. At present he is co-ordinator for the Diploma in Education programme at Queensland and is interested in developing micro-teaching techniques for research projects. During 1969-70, he was a Senior Scholar under the Australian-American Foundation Scheme and a Visiting Professor at the Science Education Centre, University of Iowa.

R. T. WHITE was a teacher in Victorian High Schools for ten years, and also taught for a year in New Zealand and for another year at the Melbourne Secondary Teachers College. A year as Senior Master and Acting Headmaster of a high school encouraged him to become Physics Executive Officer at the Victorian Universities and Schools Examinations Board in 1967. After two years in that post he began full-time study in the Faculty of Education, Monash University. He became a member of staff in March, 1971. He is now a Senior Lecturer, specializing in applied statistics and experimental design.